

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 Genève, Geneva, 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

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

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

Pekka Lahti
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

1	Introduction	1
1.1	Background and Content	1
1.2	Statistical Duality—an Outline	5
	References	9
 Part I Mathematics		
2	Rudiments of Hilbert Space Theory	13
2.1	Basic Notions and the Projection Theorem	13
2.2	The Fréchet–Riesz Theorem and Bounded Linear Operators	17
2.3	Strong, Weak, and Monotone Convergence of Nets of Operators	20
2.4	The Projection Lattice $\mathcal{P}(\mathcal{H})$	22
2.5	The Square Root of a Positive Operator	24
2.6	The Polar Decomposition of a Bounded Operator	26
2.7	Orthonormal Sets	27
2.8	Direct Sums of Hilbert Spaces	31
2.9	Tensor Products of Hilbert Spaces	32
2.10	Exercises	34
	Reference	36
3	Classes of Compact Operators	37
3.1	Compact and Finite Rank Operators	37
3.2	The Spectral Representation of Compact Selfadjoint Operators	40
3.3	The Hilbert–Schmidt Operator Class $\mathcal{HS}(\mathcal{H})$	45
3.4	The Trace Class $\mathcal{T}(\mathcal{H})$	47
3.5	Connection of the Ideals $\mathcal{T}(\mathcal{H})$ and $\mathcal{HS}(\mathcal{H})$ with the Sequence Spaces ℓ^1 and ℓ^2	49
3.6	The Dualities $\mathcal{C}(\mathcal{H})^* = \mathcal{T}(\mathcal{H})$ and $\mathcal{T}(\mathcal{H})^* = \mathcal{L}(\mathcal{H})$	52

3.7	Linear Operators on Hilbert Tensor Products and the Partial Trace.	55
3.8	The Schmidt Decomposition of an Element of $\mathcal{H}_1 \otimes \mathcal{H}_2$	59
3.9	Exercises.	61
4	Operator Integrals and Spectral Representations:	
	The Bounded Case	63
4.1	Classes of Sets and Positive Measures	63
4.2	Measurable Functions	65
4.3	Integration with Respect to a Positive Measure	65
4.4	The Hilbert Space $L^2(\Omega, \mathcal{A}, \mu)$	68
4.5	Complex Measures and Integration.	69
4.6	Positive Operator Measures	71
4.7	Positive Operator Bimeasures	75
4.8	Integration of Bounded Functions with Respect to a Positive Operator Measure.	80
4.9	The Connection Between (Semi)Spectral Measures and (Semi)Spectral Functions	83
4.10	A Riesz–Markov–Kakutani Type Representation Theorem for Positive Operator Measures.	85
4.11	The Spectral Representation of Bounded Selfadjoint Operators	86
4.12	The Spectrum of a Bounded Operator	93
4.13	The Spectral Representations of Unitary and Other Normal Operators	94
4.14	Exercises.	96
	References.	99
5	Operator Integrals and Spectral Representations:	
	The Unbounded Case	101
5.1	Elementary Notes on Unbounded Operators	101
5.2	Integration of Unbounded Functions with Respect to Positive Operator Measures	104
5.3	Integration of Unbounded Functions with Respect to Projection Valued Measures	107
5.4	The Cayley Transform	110
5.5	The Spectral Representation of an Unbounded Selfadjoint Operator	113
5.6	The Support of the Spectral Measure of a Selfadjoint Operator	115
5.7	Applying a Borel Function to a Selfadjoint Operator	117
5.8	One-Parameter Unitary Groups and Stone’s Theorem.	119
5.9	Taking Stock: Hilbert Space Theory and Its Use in Quantum Mechanics.	123

5.10	Exercises.	125
	References.	126
6	Miscellaneous Algebraic and Functional Analytic Techniques	127
6.1	Normal and Positive Linear Maps on $\mathcal{L}(\mathcal{H})$	127
6.2	Basic Notions of the Theory of C^* -algebras and Their Representations.	130
6.3	Algebraic Tensor Products of Vector Spaces	134
6.4	Completions	135
6.5	Exercises.	135
	References.	136
7	Dilation Theory	137
7.1	Completely Positive Linear Maps	137
7.2	A Bilinear Dilation Theorem	139
7.3	The Stinespring and Naimark Dilation Theorems	144
7.4	Normal Completely Positive Operators from $\mathcal{L}(\mathcal{H})$ into $\mathcal{L}(\mathcal{K})$	147
7.5	Naimark Projections of Operator Integrals	152
7.6	Operations and Instruments.	153
7.7	Measurement Dilation.	159
7.8	Exercises.	161
	References.	162
8	Positive Operator Measures: Examples	163
8.1	The Canonical Spectral Measure and Its Fourier-Plancherel Transform.	163
8.2	Restrictions of Spectral Measures	166
8.3	Smearings and Convolutions.	168
8.4	Phase Space Operator Measures	172
8.5	Moment Operators and Spectral Measures	175
8.6	Semispectral Measures and Direct Integral Hilbert Spaces	178
8.7	A Dirac Type Formalism: An Elementary Approach	182
8.8	Exercises.	185
	References.	186

Part II Elements

9	States, Effects and Observables	191
9.1	States	192
9.2	Effects.	196
9.3	Observables	200
9.4	State Changes	208
9.5	Compound Systems	213
9.6	Exercises.	221
	References.	223

10	Measurement	225
10.1	Measurement Schemes	226
10.2	Instruments	230
10.3	Sequential, Joint and Mixed Measurements	232
10.4	Examples of Measurement Schemes	236
10.5	Repeatable Measurements	247
10.6	Ideal Measurements	248
10.7	Correlations, Disturbance and Entanglement	252
10.8	Appendix	255
10.9	Exercises	258
	References	259
11	Joint Measurability	261
11.1	Definitions and Basic Results	261
11.2	Alternative Definitions	265
11.3	Regular Observables	266
11.4	Sharp Observables	269
11.5	Compatibility, Convexity, and Coarse-Graining	271
11.6	Exercises	273
	References	274
12	Preparation Uncertainty	275
12.1	Indeterminate Values of Observables	276
12.2	Measures of Uncertainty	276
12.3	Examples of Preparation Uncertainty Relations	280
12.4	Exercises	284
	References	285
13	Measurement Uncertainty	287
13.1	Conceptualising Error and Disturbance	288
13.2	Comparing Distributions	290
13.3	Error Bar Width	294
13.4	Value Comparison Error	299
13.5	Connections	302
13.6	Unsharpness	303
13.7	Finite Outcome Observables	307
13.8	Appendix	312
13.9	Exercises	314
	References	314

Part III Realisations

14	Qubits	319
14.1	Qubit States and Observables	319
14.2	Preparation Uncertainty Relations for Qubits	322
14.3	Compatibility of a Pair of Qubit Effects	324

14.4	Excursion: Compatibility of Three Qubit Effects	329
14.5	Approximate Joint Measurements of Qubit Observables.	331
14.6	Appendix	340
14.7	Exercises.	342
	References.	343
15	Position and Momentum	345
15.1	The Weyl Pairs.	345
15.2	Preparation Uncertainty Relations for $\mathbf{Q}$ and $\mathbf{P}$	350
15.3	Approximate Joint Measurements of $\mathbf{Q}$ and $\mathbf{P}$	353
15.4	Measuring $\mathbf{Q}$ and $\mathbf{P}$ with a Single Measurement Scheme.	355
15.5	Appendix	358
15.6	Exercises.	363
	References.	364
16	Number and Phase	367
16.1	Covariant Observables	368
16.2	Canonical Phase	372
16.3	Phase Space Phase Observables	379
16.4	Number-Phase Complementarity.	381
16.5	Other Phase Theories	384
16.6	Exercises.	386
	References.	387
17	Time and Energy	389
17.1	The Concept of Time in Quantum Mechanics	389
17.2	Time in Nonrelativistic Classical Mechanics.	391
17.3	Covariant Time Observables in Nonrelativistic Quantum Mechanics.	394
17.4	Exercises.	402
	References.	402
18	State Reconstruction	405
18.1	Informational Completeness	406
18.2	The Pauli Problem	408
18.3	State Reconstruction	413
18.4	Exercises.	421
	References.	423
19	Measurement Implementations	425
19.1	Arthurs–Kelly Model	426
19.2	Photon Detection, Phase Shifters and Beam Splitters	433
19.3	Balanced Homodyne Detection and Quadrature Observables . . .	439
19.4	Eight-Port Homodyne Detection and Phase Space Observables	443

19.5	Eight-Port Homodyne Detection and Phase Observables	447
19.6	Mach–Zehnder Interferometer.	452
19.7	Exercises.	461
	References.	461
 Part IV Foundations		
20	Bell Inequalities and Incompatibility	465
20.1	Bell Inequalities and Compatibility: General Observations	465
20.2	Bell Inequalities and Joint Probabilities	467
20.3	Bell Inequality Violation and Nonlocality.	470
20.4	Bell Inequality Violation and Incompatibility	471
20.5	Exercises.	475
	References.	475
21	Measurement Limitations Due to Conservation Laws	477
21.1	Measurement of Spin Versus Angular Momentum Conservation.	477
21.2	The Yanase Condition	479
21.3	The Wigner–Araki–Yanase Theorem	479
21.4	A Quantitative Version of the WAY Theorem	481
21.5	Position Measurements Obeying Momentum Conservation	482
21.6	A Measurement-Theoretic Interpretation of Superselection Rules	485
21.7	Exercises.	487
	References.	487
22	Measurement Problem	489
22.1	Preliminaries.	489
22.2	Reading of Pointer Values	491
22.3	The Problem of Objectification.	493
22.4	Exercises.	496
	References.	496
23	Axioms for Quantum Mechanics	499
23.1	Statistical Duality and Its Representation	500
23.2	Quantum Logic.	507
23.3	Filters and the Projection Postulate.	514
23.4	Hilbert Space Coordinatisation	518
23.5	The Role of Symmetries in the Representation Theorem	524
23.6	The Case of the Complex Field	530
23.7	Exercises.	532
	References.	534
Index		537

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:

It is assumed that the description of a physical system is based on a complex separable Hilbert space $\mathcal{H}$ with inner product $\langle \cdot | \cdot \rangle$ 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 $\langle \varphi | E^A(X) \varphi \rangle$, where $E^A(X)$ 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 $\text{tr}[\rho E^A(X)]$ 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 .

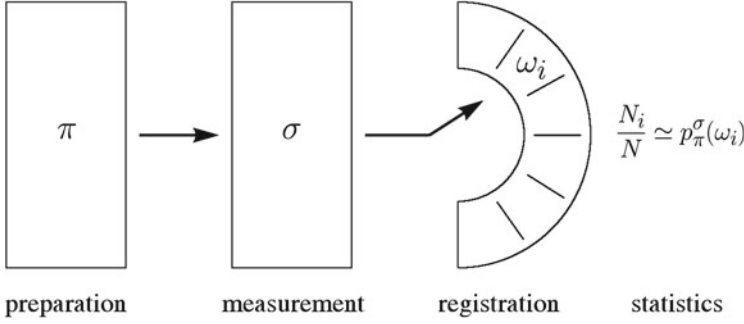


Fig. 1.1 Scheme of a physical experiment

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 $p_\pi^\sigma : \mathcal{A} \rightarrow [0, 1]$, 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 , then¹

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

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 $\mathcal{S}$ 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 $\mathcal{O}$ 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 $(\mathcal{S}, \mathcal{O})$ as a *statistical duality*.

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

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

¹We 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 (S, O) may always be assumed to possess. First, since a convex combination of two or more probability measures is a probability measure, the set S of states can be equipped with a convex structure. Indeed, if $s_1, s_2 \in S$ and $0 \leq \lambda \leq 1$, then for any $O \in O$, 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) $s \in S$ such that

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

for all $O \in O$. The assumption that 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 structure² of the set of states 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.)

²The 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 $(\mathcal{S}, \mathcal{O})$. 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 $f_i : \Omega \rightarrow \Omega_i$, $i = 1, 2$, such that for any state $s \in \mathcal{S}$,

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

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

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 $(\tilde{O}, \tilde{\Omega}, \tilde{\mathcal{A}})$, together with measurable functions $f_i : \tilde{\Omega} \rightarrow \Omega_i$, $i = 1, 2$, such that for any state $s \in \mathcal{S}$, 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 $(\mathcal{S}, \mathbf{O})$. 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 $(\mathcal{S}, \mathbf{O})$ can be considered to be composed of two (or more) subsystems with the dualities $(\mathcal{S}_i, \mathbf{O}_i)$. 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. $\mathbb{N} = \{1, 2, 3, \dots\}$, and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Definition 2.1 Let E be a (complex) vector space. We say that a mapping $h : E \times E \rightarrow \mathbb{C}$ 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

$$(IP1) \quad h(\varphi, \alpha\psi + \beta\eta) = \alpha h(\varphi, \psi) + \beta h(\varphi, \eta),$$

$$(IP2) \quad h(\varphi, \psi) = \overline{h(\psi, \varphi)},$$

$$(IP3) \quad h(\varphi, \varphi) \geq 0,$$

$$(IP4) \quad h(\varphi, \varphi) > 0 \text{ if } \varphi \neq 0.$$

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*.

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

- (a) $\langle \alpha\varphi + \beta\psi | \eta \rangle = \bar{\alpha} \langle \varphi | \eta \rangle + \bar{\beta} \langle \psi | \eta \rangle$;
- (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

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

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 $\langle \varphi | \psi \rangle = 0$. $\square$

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

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

and substitute first $\alpha = \beta = 1$ and then $\alpha = 1, \beta = i$ to see that $\text{Im } h(\varphi, \psi) = -\text{Im } h(\psi, \varphi)$ and $\text{Re } h(\varphi, \psi) = \text{Re } h(\psi, \varphi)$. 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. $\triangleleft$

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 $\varphi \mapsto \sqrt{\langle \varphi | \varphi \rangle}$ is a *norm* on E . Unless otherwise stated, an inner product space will be equipped with this norm. $\triangleleft$

Example 2.1 The set $\mathbb{C}^n = \{x = (x_1, \dots, x_n) \mid x_k \in \mathbb{C}, k = 1, \dots, n\}$ (where $n \in \mathbb{N}$) is an inner product space with respect to its usual operations $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ and $\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$ and the

inner product $\langle x | y \rangle = \sum_{k=1}^n \bar{x}_k y_k$, $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$. The Cauchy-Schwarz inequality acquires the form

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

◁

Example 2.2 Denote $\ell^2 = \ell_{\mathbb{C}}^2 = \{f : \mathbb{N} \rightarrow \mathbb{C} \mid \sum_{n=1}^{\infty} |f(n)|^2 < \infty\}$ 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

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

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

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

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

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 $\varphi_1, \dots, \varphi_n \in E$ are vectors satisfying $\langle \varphi_i | \varphi_j \rangle = 0$, 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 $B : F \times F \rightarrow \mathbb{C}$ 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 $\lim_{n \rightarrow \infty} \|\varphi_n\| = d$. The parallelogram law shows that

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

when $m, n \rightarrow \infty$. Thus (φ_n) is a Cauchy sequence and hence converges to some $\varphi \in E$ (as E is closed), and $\|\varphi\| = \lim_{n \rightarrow \infty} \|\varphi_n\| = d$ 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 $P_M : \mathcal{H} \rightarrow M$, 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 $\|P_M \varphi\| = \|\psi\| \leq \|\varphi\|$.

2.2 The Fréchet–Riesz Theorem and Bounded Linear Operators

Let $\mathcal{H}$ be a Hilbert space. A linear map $T : \mathcal{H} \rightarrow \mathcal{H}$ is called a *bounded (linear) operator* or a *bounded linear map* if there is a constant $C \in [0, \infty)$ satisfying $\|T\varphi\| \leq C \|\varphi\|$ 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 $T : \mathcal{H} \rightarrow \mathcal{H}$. If $T \in \mathcal{L}(\mathcal{H})$ we write $\|T\| = \sup\{\|T\varphi\| \mid \varphi \in \mathcal{H}, \|\varphi\| \leq 1\}$. It is easy to see that $\mathcal{L}(\mathcal{H})$ is a vector space, and $\|\cdot\|$ is a norm. Moreover, $\|T\varphi\| \leq \|T\| \|\varphi\|$, which implies that $\|ST\| \leq \|S\| \|T\|$ 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 $f : G \rightarrow \mathbb{C}$. Here continuity is equivalent to the condition

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

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 $f \neq 0$, 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 $f(\varphi - (f(\varphi)/f(\xi))\xi) = f(\varphi) - (f(\varphi)/f(\xi))f(\xi) = 0$. This means that $\langle \xi \mid \varphi - (f(\varphi)/f(\xi))\xi \rangle = 0$, 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 $\langle \psi \mid \psi - \psi' \rangle = \langle \psi' \mid \psi - \psi' \rangle$, then $\|\psi - \psi'\|^2 = 0$. As $|f(\varphi)| \leq \|\psi\| \|\varphi\|$, we have $\|f\| \leq \|\psi\|$, and on the other hand $\|\psi\|^2 = f(\psi) \leq \|f\| \|\psi\|$, so that $\|\psi\| \leq \|f\|$. $\square$

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 $B : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ 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) $\sup \{|B(\varphi, \psi)| \mid \|\varphi\| \leq 1, \|\psi\| \leq 1\} < \infty$.

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

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 $|f_\varphi(\psi)| \leq C \|\varphi\| \|\psi\|$, f_φ is continuous. Theorem 2.5 yields a unique $\xi_\varphi \in \mathcal{H}$ such that $f_\varphi(\psi) = \langle \xi_\varphi \mid \psi \rangle$ for all $\psi \in \mathcal{H}$. We define $S\varphi = \xi_\varphi$. Since $B(\alpha\varphi_1 + \beta\varphi_2, \psi) = \overline{\alpha}B(\varphi_1, \psi) + \overline{\beta}B(\varphi_2, \psi) = \overline{\alpha} \langle S\varphi_1 \mid \psi \rangle + \overline{\beta} \langle S\varphi_2 \mid \psi \rangle = \langle \alpha S\varphi_1 + \beta S\varphi_2 \mid \psi \rangle$, S is linear. Since

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

we have $\|S\varphi\| \leq C \|\varphi\|$, 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. $\square$

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 $\langle \varphi \mid T\psi \rangle = \langle T^*\varphi \mid \psi \rangle$ for all $\varphi, \psi \in \mathcal{L}(\mathcal{H})$: we simply take $B(\varphi, \psi) = \langle \varphi \mid T\psi \rangle$ in Proposition 2.1. Since $\|T^*\varphi\|^2 \leq \langle \varphi \mid T T^* \varphi \rangle \leq \|\varphi\| \|T\| \|T^*\varphi\|$, it is clear that $\|T^*\| \leq \|T\|$. Using (a) in the next theorem, we see that on the other hand $\|T\| = \|T^{**}\| \leq \|T^*\|$, 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 $T^* T = T T^*$, T is *normal*. If $T^* T = T T^* = I$, 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*

$$\|T\| = \sup_{\|\varphi\| \leq 1} |\langle \varphi \mid T\varphi \rangle|.$$

Proof Using the polarisation identity and the parallelogram law we obtain for $\varphi, \psi \in \mathcal{H}$ with $\|\varphi\| \leq 1, \|\psi\| \leq 1$,

$$\begin{aligned} |\operatorname{Re} \langle \varphi | T\psi \rangle| &= \frac{1}{4} |\langle \psi + \varphi | T(\psi + \varphi) \rangle - \langle \psi - \varphi | T(\psi - \varphi) \rangle| \\ &\leq \frac{1}{4} M(\|\psi + \varphi\|^2 + \|\psi - \varphi\|^2) = \frac{1}{2} M(\|\varphi\|^2 + \|\psi\|^2) \leq M, \end{aligned}$$

where $M = \sup_{\|\varphi\| \leq 1} |\langle \varphi | T\varphi \rangle|$. (Note that, e.g., $\langle \psi + \varphi | T(\psi + \varphi) \rangle \in \mathbb{R}$.) Suppose that $\|\varphi\| \leq 1$ and $\|\psi\| \leq 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $|\langle \varphi | T\psi \rangle| = \alpha \langle \varphi | T\psi \rangle = \langle \varphi | T\alpha\psi \rangle$. The first part of the proof (applied to the vectors $\alpha\psi$ and φ) yields $|\langle \varphi | T\psi \rangle| = \operatorname{Re} \langle \varphi | T\alpha\psi \rangle \leq M$, so that

$$\|T\| = \sup \{ |\langle \varphi | T\psi \rangle| \mid \|\varphi\| \leq 1, \|\psi\| \leq 1 \} \leq M.$$

Conversely, $|\langle \varphi | T\varphi \rangle| \leq \|T\|$ when $\|\varphi\| \leq 1$, and so $M \leq \|T\|$. $\square$

We end this section with two useful decomposition results.

Proposition 2.3 *If $T \in \mathcal{L}(\mathcal{H})$ then T can be written in a unique way as $T = A + iB$ where $A, B \in \mathcal{L}(\mathcal{H})$ are selfadjoint. The operator T is normal if and only if $AB = BA$.*

Proof If we have $T = A + iB$, then necessarily $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$, and conversely. A simple calculation proves the second claim. $\square$

For the set of the selfadjoint operators in $\mathcal{L}(\mathcal{H})$ we use the notation $\mathcal{L}_s(\mathcal{H})$. We now consider a natural partial order in $\mathcal{L}_s(\mathcal{H})$. We say that $T \in \mathcal{L}(\mathcal{H})$ is *positive*, and write $T \geq 0$, if $\langle \varphi | T\varphi \rangle \geq 0$ for all $\varphi \in \mathcal{H}$. We denote the set of the positive operators $T \in \mathcal{L}(\mathcal{H})$ by $\mathcal{L}_s(\mathcal{H})^+$. It follows from Remark 2.1 that for $T \in \mathcal{L}(\mathcal{H})$ we have $T \in \mathcal{L}_s(\mathcal{H})$ if and only if $\langle \varphi | T\varphi \rangle \in \mathbb{R}$ for all $\varphi \in \mathcal{H}$. In particular, $\mathcal{L}_s(\mathcal{H})^+ \subset \mathcal{L}_s(\mathcal{H})$. For $S, T \in \mathcal{L}_s(\mathcal{H})$ we write $S \leq T$ if and only if $T - S \geq 0$. Clearly $T \leq T$, and if $R \leq S$ and $S \leq T$, then $R \leq T$. Moreover, the conditions $S \leq T$ and $T \leq S$ together imply $\langle \varphi | (T - S)\varphi \rangle = 0$ for all $\varphi \in \mathcal{H}$, and so by the polarisation identity (or Proposition 2.2) $S = T$. Thus we have a partial order in $\mathcal{L}_s(\mathcal{H})$. In the next decomposition result there is no uniqueness claim.

Proposition 2.4 *If $A \in \mathcal{L}_s(\mathcal{H})$ then A can be written as $A = A_1 - A_2$ where $A_1, A_2 \in \mathcal{L}_s(\mathcal{H})^+$.*

Proof We may choose $A_1 = \frac{1}{2}(\|A\| I + A)$ and $A_2 = \frac{1}{2}(\|A\| I - A)$. $\square$

2.3 Strong, Weak, and Monotone Convergence of Nets of Operators

The usual norm on the space $\mathcal{L}(\mathcal{H})$ of bounded linear operators on $\mathcal{H}$ determines its canonical Banach space structure. There are several other important locally convex topologies on $\mathcal{L}(\mathcal{H})$. We postpone the discussion of some of them to later chapters. In this section two notions of convergence in $\mathcal{L}(\mathcal{H})$ are introduced: strong and weak. They are related to the so-called strong and weak operator topologies, but here we avoid the explicit use of these topologies.

Let $(\mathcal{I}, \geq)$ be a *directed set*. This means that “ $\geq$ ” is a binary relation on the set $\mathcal{I}$ satisfying the following conditions:

- (D1) $m \geq p$ whenever $m \geq n$ and $n \geq p$;
- (D2) $m \geq m$ whenever $m \in \mathcal{I}$;
- (D3) whenever $m, n \in \mathcal{I}$ there is some $p \in \mathcal{I}$ satisfying $p \geq m$ and $p \geq n$.

A mapping $i \mapsto x_i$ from $\mathcal{I}$ into a set X is then called a *net* or a *generalised sequence* (in X). Such a net is often denoted by $(x_i)_{i \in \mathcal{I}}$, generalising the notation for a sequence. If X here is a topological space, the net $(x_i)_{i \in \mathcal{I}}$ *converges* to a point $x \in X$ if for every neighbourhood U of x there is some $i_0 \in \mathcal{I}$ such that $x_i \in U$ whenever $i \geq i_0$. If X is a Hausdorff space, e.g., a metrisable space, condition (D3) implies that x , the *limit* of the net $(x_i)_{i \in \mathcal{I}}$, is uniquely determined. We use the notations $\lim_{i \in \mathcal{I}} x_i$, $\lim_i x_i$, $\lim x_i$ for this limit.

Definition 2.3 Let $(\mathcal{I}, \geq)$ be a directed set and $T_i \in \mathcal{L}(\mathcal{H})$ for all $i \in \mathcal{I}$.

- (a) The net $(T_i)_{i \in \mathcal{I}}$ *converges strongly* to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim T_i \varphi = T \varphi$ for all $\varphi \in \mathcal{H}$. We then denote $T_i \rightarrow^s T$ or $T = \text{s-lim } T_i$.
- (b) The net $(T_i)_{i \in \mathcal{I}}$ *converges weakly* to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim \langle \varphi | T_i \psi \rangle = \langle \varphi | T \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$. We then denote $T_i \rightarrow^w T$ or $T = \text{w-lim } T_i$.

Since the inner products $\langle \varphi | T \psi \rangle$ completely determine T , the limit operator in (b) is also completely determined. The polarisation identity of Theorem 2.3 (c) shows that the condition $\lim \langle \varphi | T_i \varphi \rangle = \langle \varphi | T \varphi \rangle$ for all $\varphi \in \mathcal{H}$ already guarantees that $T_i \rightarrow^w T$. It is easy to see that a norm convergent net converges strongly and a strongly convergent net converges weakly. If the dimension of $\mathcal{H}$ is infinite, then in general neither implication can be reversed.

Let $\mathcal{F} \subset \mathcal{L}_s(\mathcal{H})$. The set $\mathcal{F}$ is *bounded above* if it has an *upper bound*, say $T \in \mathcal{L}_s(\mathcal{H})$, satisfying $S \leq T$ for all $S \in \mathcal{F}$. If T_0 is an upper bound of $\mathcal{F}$ satisfying $T_0 \leq T$ for every upper bound T of $\mathcal{F}$, then T_0 is the (clearly uniquely determined) *least upper bound (supremum)* of $\mathcal{F}$, and we denote $T_0 = \sup \mathcal{F}$. A *lower bound* and the *greatest lower bound (infimum)* $\inf \mathcal{F}$ are analogously defined. The same terminology is used for any partially ordered set.

Theorem 2.7 Let $(\mathcal{I}, \geq)$ be a directed set and $(T_i)_{i \in \mathcal{I}}$ an increasing net in $\mathcal{L}_s(\mathcal{H})$ (i.e. $T_i \geq T_j$ whenever $i \geq j$). If the set $\{T_i \mid i \in \mathcal{I}\}$ is bounded above, then it has the least upper bound, say T . Moreover, $T_i \rightarrow^s T$ and $T_i \rightarrow^w T$. The similar statement concerning the greatest lower bounds of decreasing nets bounded below is also valid.

Proof For each $\varphi \in \mathcal{H}$, the net $(\langle \varphi | T_i \varphi \rangle)_{i \in \mathcal{I}}$ in $\mathbb{R}$ is increasing and bounded above by $\langle \varphi | S_0 \varphi \rangle$ where $S_0 \in \mathcal{L}_s(\mathcal{H})$ is some upper bound of $\{T_i \mid i \in \mathcal{I}\}$ and so it has a limit which we denote by $f(\varphi)$. The polarisation identity shows that we can also define $B(\varphi, \psi) = \lim_{i \in \mathcal{I}} \langle \varphi | T_i \psi \rangle = \frac{1}{4} \sum_{n=0}^3 i^n f(\psi + i^n \varphi)$ for all $\varphi, \psi \in \mathcal{H}$. The usual limit rules (valid also for nets) show that B satisfies the conditions (i) and (ii) in Proposition 2.1. We show that its boundedness condition (iii) also holds. Without loss of generality we may assume that $\mathcal{I}$ has a smallest element i_0 , and since $\langle \xi | T_{i_0} \xi \rangle \leq f(\xi) \leq \langle \xi | S_0 \xi \rangle$ for all $\xi \in \mathcal{H}$, we get $|B(\varphi, \psi)| \leq |f(\psi + i^n \varphi)| \leq \|\psi + i^n \varphi\|^2 \max\{\|T_{i_0}\|, \|S_0\|\} \leq 4 \max\{\|T_{i_0}\|, \|S_0\|\}$ whenever $\|\varphi\| \leq 1$ and $\|\psi\| \leq 1$. Using Proposition 2.1 we thus get a unique $T \in \mathcal{L}(\mathcal{H})$ such that $B(\varphi, \psi) = \langle T \varphi | \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$. One immediately verifies that $T \in \mathcal{L}_s(\mathcal{H})$ and $\langle \varphi | T \psi \rangle = \lim_{i \in \mathcal{I}} \langle \varphi | T_i \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$. By definition, $T_i \leq T$, and if $S \in \mathcal{L}_s(\mathcal{H})$ satisfies $T_i \leq S$ for all $i \in \mathcal{I}$, then $\langle \varphi | T \varphi \rangle = \lim_{i \in \mathcal{I}} \langle \varphi | T_i \varphi \rangle \leq \langle \varphi | S \varphi \rangle$. Thus $T = \sup_{i \in \mathcal{I}} T_i$. We have also seen that $T = \text{w-lim } T_i$. We still show that $T = \text{s-lim } T_i$. The mapping $(\xi, \eta) \mapsto \langle \xi | (T - T_i) \eta \rangle$ is a positive sesquilinear form, and so it satisfies the Cauchy–Schwarz inequality (see Remark 2.1). Therefore, if $\varphi \in \mathcal{H}$, then

$$\begin{aligned} |\langle \xi | (T - T_i) \varphi \rangle|^2 &\leq \langle \xi | (T - T_i) \xi \rangle \langle \varphi | (T - T_i) \varphi \rangle \\ &\leq \langle \xi | (T - T_{i_0}) \xi \rangle \langle \varphi | (T - T_i) \varphi \rangle \\ &\leq \|T - T_{i_0}\| \langle \varphi | (T - T_i) \varphi \rangle, \end{aligned}$$

whenever $i \in \mathcal{I}$ and $\|\xi\| \leq 1$, and so

$$\|(T - T_i) \varphi\| = \sup_{\|\xi\| \leq 1} |\langle \xi | (T - T_i) \varphi \rangle| \leq (\|T - T_{i_0}\| \langle \varphi | (T - T_i) \varphi \rangle)^{\frac{1}{2}} \longrightarrow 0.$$

When the operators above are multiplied by -1 we get the claim concerning decreasing nets. $\square$

The following observation will be used later.

Theorem 2.8 (a) Let $(T_i)_{i \in \mathcal{I}}$ be a net in $\mathcal{L}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$ such that $T_i \rightarrow^w T$. Then $T_i^* \rightarrow^w T^*$ and $T_i S \rightarrow^w T S$, $S T_i \rightarrow^w S T$ for all $S \in \mathcal{L}(\mathcal{H})$.

(b) If $(T_i)_{i \in \mathcal{I}}$ is a net in $\mathcal{L}_s(\mathcal{H})^+$ which is increasing and bounded above or decreasing and bounded below, and $T = \text{w-lim } T_i$, then $T^2 = \text{w-lim } T_i^2$.

Proof (a) A straightforward calculation yields this.

(b) From Theorem 2.7 it follows that $T = \text{s-lim } T_i$. Hence for every φ we get $\langle \varphi | T_i^2 \varphi \rangle = \langle T_i \varphi | T_i \varphi \rangle = \|T_i \varphi\|^2 \rightarrow \|T \varphi\|^2 = \langle \varphi | T^2 \varphi \rangle$ implying the claim. $\square$

2.4 The Projection Lattice $\mathcal{P}(\mathcal{H})$

We have defined the (orthogonal) projection of $\mathcal{H}$ onto a closed subspace $M \subset \mathcal{H}$ as the map $P_M (\in \mathcal{L}(\mathcal{H}))$ for which $P_M \varphi = \psi$ when $\varphi = \psi + \xi$ with $\psi \in M$, $\xi \in M^\perp$. We say that P is a *projection* if $P = P_M$ for some closed subspace M of $\mathcal{H}$. We omit the proof of the following list of elementary properties of projections. We generally denote $I = I_{\mathcal{H}} = \text{id}_{\mathcal{H}} = P_{\mathcal{H}}$.

Theorem 2.9 *Let M be a closed subspace of $\mathcal{H}$ and $P = P_M$. Then*

- (a) $M = P(\mathcal{H}) = \{\varphi \mid P\varphi = \varphi\} = \{\varphi \mid \|P\varphi\| = \|\varphi\|\};$
- (b) $M^\perp = \{\varphi \mid P\varphi = 0\};$
- (c) $P = P^2 = P^*;$
- (d) $P_M + P_{M^\perp} = I;$
- (e) $M^{\perp\perp} = M.$

The following characterisation is basic.

Theorem 2.10 *For a linear map $P : \mathcal{H} \rightarrow \mathcal{H}$, the following conditions are equivalent:*

- (i) P is a projection;
- (ii) $P = P^2$ and $\langle \varphi \mid P\psi \rangle = \langle P\varphi \mid \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$;
- (iii) $P = P^2$ and $\|P\| \leq 1$ (in particular, $P \in \mathcal{L}(\mathcal{H})$).

Proof By Theorem 2.9 (i) implies (ii), and (ii) implies (iii) because $\|P\varphi\|^2 = \langle P\varphi \mid P\varphi \rangle = \langle \varphi \mid P^2\varphi \rangle = \langle \varphi \mid P\varphi \rangle \leq \|\varphi\| \|P\varphi\|$ so that $\|P\varphi\| \leq \|\varphi\|$. Now assume (iii). As $P = P^2$, one gets the direct sum representation $\mathcal{H} = M \oplus N$ with the notation $M = P(\mathcal{H})$, $N = (I - P)(\mathcal{H})$. Moreover, N is the kernel of P and M that of $I - P$, and so both are closed. Since also $\mathcal{H} = M \oplus M^\perp$, to see that $P = P_M$ it will suffice to show, e.g., that $M^\perp \subset N$. If $\varphi \in N^\perp$, we have $P\varphi = \varphi + \psi$ where $\psi = P\varphi - \varphi \in N$, and so $\langle \psi \mid \varphi \rangle = 0$, implying $\|\varphi\|^2 \geq \|P\varphi\|^2 = \|\varphi\|^2 + \|\psi\|^2$ so that $\psi = 0$ and $\varphi = P\varphi \in M$. Thus $N^\perp \subset M$, implying $M^\perp \subset N^{\perp\perp} = N$. $\square$

In the rest of this section we use the notation $\mathcal{M}(\mathcal{H})$ for the set of the closed subspaces of $\mathcal{H}$. Theorems 2.9 and 2.10 show that the mapping $P \mapsto P(\mathcal{H})$ is a bijection from the set $\mathcal{P}(\mathcal{H}) = \{P \in \mathcal{L}(\mathcal{H}) \mid P = P^2 = P^*\}$ onto $\mathcal{M}(\mathcal{H})$. We mention some results related to this bijection. The proofs are straightforward and we omit them. In the following three theorems we assume that $P, Q \in \mathcal{P}(\mathcal{H})$ and denote $M = P(\mathcal{H})$, $N = Q(\mathcal{H})$.

Theorem 2.11 *The following conditions are equivalent:*

- (i) $M \subset N;$
- (ii) $QP = P;$
- (iii) $PQ = P;$
- (iv) $\|P\varphi\| \leq \|Q\varphi\|$ for all $\varphi \in \mathcal{H};$
- (v) $P \leq Q$ (i.e. $\langle \varphi \mid P\varphi \rangle \leq \langle \varphi \mid Q\varphi \rangle$ for all $\varphi \in \mathcal{H}$);

(vi) $Q - P \in \mathcal{P}(\mathcal{H})$.

Theorem 2.12 *The following conditions are equivalent:*

- (i) $M \perp N$ (i.e. $\langle \varphi | \psi \rangle = 0$ for all $\varphi \in M, \psi \in N$);
- (ii) $PQ = 0$;
- (iii) $QP = 0$;
- (iv) $Q + P \in \mathcal{P}(\mathcal{H})$.

Theorem 2.13 *The following conditions are equivalent:*

- (i) $PQ = QP$;
- (ii) $PQ \in \mathcal{P}(\mathcal{H})$;
- (iii) $QP \in \mathcal{P}(\mathcal{H})$.

If these conditions hold, then $PQ(\mathcal{H}) = P(\mathcal{H}) \cap Q(\mathcal{H})$.

The next result generalises parts (i) and (v) of Theorem 2.11.

Proposition 2.5 *If $P \in \mathcal{P}(\mathcal{H})$, $M = P(\mathcal{H})$ and $T \in \mathcal{L}_s(\mathcal{H})^+$, then the following conditions are equivalent:*

- (i) $T(\mathcal{H}) \subset M$ and $\|T\| \leq 1$;
- (ii) $T \leq P$.

Proof First assume (i). For all $\varphi \in \mathcal{H}$ we get

$$\langle \varphi | T\varphi \rangle = \langle \varphi | PT\varphi \rangle = \langle P\varphi | TP\varphi \rangle \leq \|P\varphi\| \|T\| \|P\varphi\| \leq \|P\varphi\|^2 = \langle \varphi | P\varphi \rangle,$$

since $PT = T$, so that $T = T^* = T^*P^* = TP$.

Next assume (ii). Since the map $(\varphi, \psi) \mapsto \langle \varphi | T\psi \rangle$ is a positive sesquilinear form, the Cauchy–Schwarz inequality gives $|\langle \xi | T\varphi \rangle|^2 \leq \langle \xi | T\xi \rangle \langle \varphi | T\varphi \rangle \leq \langle \xi | P\xi \rangle \langle \varphi | P\varphi \rangle$ for all $\xi, \varphi \in \mathcal{H}$. In particular, $\|T\varphi\|^2 = \sup_{\|\xi\| \leq 1} |\langle \xi | T\varphi \rangle|^2 \leq 1$ when $\|\varphi\| \leq 1$, and so $\|T\| \leq 1$. Moreover, $T\varphi = 0$ if $P\varphi = 0$, and so $\langle \varphi | T\psi \rangle = \langle T\varphi | \psi \rangle = 0$ whenever $\psi \in \mathcal{H}$ and $\varphi \in M^\perp$, implying $T\psi \in M^{\perp\perp} = M$. $\square$

In the partially ordered set $(\mathcal{L}_s(\mathcal{H}), \leq)$ it is common that even two-element sets fail to have least upper bounds and greatest lower bounds. Thus $(\mathcal{L}_s(\mathcal{H}), \leq)$ is not a lattice. (By definition, a *lattice* is a partially ordered set such that any two-element subset has supremum and infimum.) In fact, it is an *antilattice*, that is, any two elements S and T have greatest lower bound exactly when they are comparable, i.e. $S \leq T$ or $T \leq S$ (and the same is true of least upper bounds). For a proof of this result due to R. Kadison we refer to [1], p. 417. The subset $\mathcal{P}(\mathcal{H})$ of $(\mathcal{L}_s(\mathcal{H}), \leq)$ is, however, even a complete lattice. The next result gives even more information: if $\mathcal{P}(\mathcal{H})$ is regarded as a subset of the larger partially ordered set $\{T \in \mathcal{L}_s(\mathcal{H}) \mid 0 \leq T \leq I\}$ (which we denote by $[0, I]$), the greatest lower bound and the least upper bound of a subset of $\mathcal{P}(\mathcal{H})$ are not affected.

Theorem 2.14 *Let $(P_i)_{i \in \mathcal{I}}$ be any family in $\mathcal{P}(\mathcal{H})$ and denote $M_i = P_i(\mathcal{H})$ for $i \in \mathcal{I}$. Write $M = \bigcap_{i \in \mathcal{I}} M_i$ and let N be the intersection of all the closed subspaces of $\mathcal{H}$ containing every M_i , $i \in \mathcal{I}$. Then P_M is the greatest lower bound and P_N the least upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$ with respect to the set $[0, I] = \{T \in \mathcal{L}_s(\mathcal{H}) \mid 0 \leq T \leq I\}$.*

Proof According to Theorem 2.11, $P_M \leq P_i$ for all $i \in \mathcal{I}$. On the other hand, if $T \in [0, I]$ is a lower bound of the set $\{P_i \mid i \in \mathcal{I}\}$, Proposition 2.5 shows us that $\|T\| \leq 1$ and $T(\mathcal{H}) \subset M_i$ for all $i \in \mathcal{I}$, i.e. $T(\mathcal{H}) \subset M$, so that by the same theorem $T \leq P_M$. This proves the claim about M .

As for N , note that by the above argument, the set of the complementary projections, $\{I - P_i \mid i \in \mathcal{I}\}$, has a greatest lower bound S in the set $[0, I]$. The mapping $T \mapsto I - T$ keeps $[0, I]$ invariant and inverts the order of operators, which implies that $I - S$ is the least upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$ in $[0, I]$. Now S and hence also $I - S$ is a projection. We show that $I - S = P_N$. We have $M_i \subset (I - S)(\mathcal{H})$ for all $i \in \mathcal{I}$, so that by the definition of N we have $I - S \geq P_N$. On the other hand, P_N is an upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$, and so $I - S \leq P_N$. $\square$

In the situation of the above theorem we denote $P_M = \bigwedge_{i \in \mathcal{I}} P_i$ and $P_N = \bigvee_{i \in \mathcal{I}} P_i$. In the case of finitely many projections $P_1, \dots, P_n$ we may also write $\bigwedge_{i \in \{1, \dots, n\}} P_i = P_1 \wedge \dots \wedge P_n$ and $\bigvee_{i \in \{1, \dots, n\}} P_i = P_1 \vee \dots \vee P_n$.

Remark 2.3 Consider on $\mathcal{P}(\mathcal{H})$ the map $P \mapsto P^\perp = I - P$. It is an order-reversing involution and maps each element to a complement, i.e. $P \vee P^\perp = I$ and $P \wedge P^\perp = 0$ (see the above theorem). Such a map is called an *orthocomplementation*. We say that $P, R \in \mathcal{P}(\mathcal{H})$ are *orthogonal* if $P \leq R^\perp$, equivalently, $R \leq P^\perp$, and we denote $P \perp R$. If P and R are orthogonal, they are also *disjoint*, that is, $P \wedge R = 0$. If $\dim(\mathcal{H}) \geq 2$, it is easy to give examples of projections which are disjoint but not orthogonal. $\triangleleft$

2.5 The Square Root of a Positive Operator

We prove that any positive operator has a unique positive square root.

Theorem 2.15 *Suppose that $A \in \mathcal{L}_s(\mathcal{H})^+$. There is a uniquely determined operator $B \in \mathcal{L}_s(\mathcal{H})^+$ (usually denoted by $A^{\frac{1}{2}}$ or $\sqrt{A}$ and called the square root of A) satisfying $B^2 = A$.*

Proof We may assume that $A \in \mathcal{L}_s(\mathcal{H})$ is such that $0 \leq A \leq I$. Define a sequence (B_n) of operators recursively by setting $B_1 = 0$,

$$B_{n+1} = \frac{1}{2}[(I - A) + B_n^2],$$

$n = 1, 2, \dots$ Then

$$B_{n+1} - B_n = \frac{1}{2}[B_n^2 - B_{n-1}^2] = \frac{1}{2}(B_n + B_{n-1})(B_n - B_{n-1}),$$

since the operators B_n ja B_{n-1} , being polynomials of $I - A$, commute. By induction one sees that in fact each B_n is a polynomial of $I - A$ with positive coefficients. After this observation, from the equation obtained before we see by induction that each difference is a polynomial in $I - A$ with positive coefficients, so that this difference is a positive operator, as every power $(I - A)^k \geq 0$. (Observe that $\langle (I - A)^{2m} \varphi | \varphi \rangle = \langle (I - A)^m \varphi | (I - A)^m \varphi \rangle \geq 0$ and $\langle (I - A)^{2m+1} \varphi | \varphi \rangle = \langle (I - A)(I - A)^m \varphi | (I - A)^m \varphi \rangle \geq 0$.) By induction it is seen that $\|B_n\| \leq 1$. Hence the increasing sequence (B_n) has a positive bounded (strong) limit operator $\tilde{B} \leq I$ by Theorem 2.7, and the equality $\tilde{B} = \frac{1}{2}[I - A + \tilde{B}^2]$ follows from Theorem 2.8 (b). Denoting $B = I - \tilde{B}$, we thus have $B^2 = A$.

We next prove the uniqueness claim. Consider the above situation where $0 \leq A \leq I$. Let B be the operator constructed above, satisfying $B^2 = A$. Let also $C \in \mathcal{L}_s(\mathcal{H})^+$ satisfy $C^2 = A$. Then $CA = C^3 = AC$, so that C commutes with every polynomial in A , implying that $CB = BC$. We now use the method described in the first part of the proof to obtain two operators $B_1, C_1 \in \mathcal{L}_s(\mathcal{H})^+$ such that $B_1^2 = B$ and $C_1^2 = C$. Let $\varphi \in \mathcal{H}$ and denote $\psi = (B - C)\varphi$. Then

$$\begin{aligned} \|B_1\psi\|^2 + \|C_1\psi\|^2 &= \langle B_1^2\psi | \psi \rangle + \langle C_1^2\psi | \psi \rangle = \langle B\psi | \psi \rangle + \langle C\psi | \psi \rangle \\ &= \langle B(B - C)\varphi | \psi \rangle + \langle C(B - C)\varphi | \psi \rangle \\ &= \langle (B^2 - C^2)\varphi | \psi \rangle = 0. \end{aligned}$$

It follows that $B_1\psi = C_1\psi = 0$, so that $B\psi = B_1B_1\psi = 0$ and $C\psi = C_1C_1\psi = 0$, implying $\|(B - C)\varphi\|^2 = \langle (B - C)^2\varphi | \varphi \rangle = \langle (B - C)\psi | \varphi \rangle = 0$. Thus $B\varphi = C\varphi$ for all $\varphi \in \mathcal{H}$, i.e. $B = C$. $\square$

The square root gives rise to the following definition.

Definition 2.4 If $T \in \mathcal{L}(\mathcal{H})$, the positive operator $\sqrt{T^*T}$ is denoted by $|T|$ and called the *absolute value* of T .

The absolute value $|T|$ may be characterised as the only positive operator A satisfying $\|A\varphi\| = \|T\varphi\|$ for all $\varphi \in \mathcal{H}$ (exercise).

Remark 2.4 Using the square root we get a quick proof for the implication (ii) $\implies$ (i) in Proposition 2.5: $\|T^{\frac{1}{2}}\varphi\|^2 = \langle \varphi | T\varphi \rangle \leq \langle \varphi | P\varphi \rangle \leq 1$ if $\|\varphi\| \leq 1$, and so $\|T\| \leq \|T^{\frac{1}{2}}\|^2 \leq 1$. Moreover, $P\varphi = 0$ implies $T\varphi = (T^{\frac{1}{2}})^2\varphi = 0$, and the proof is completed as originally. $\triangleleft$

We conclude with another application of the square root.

Proposition 2.6 Any operator $T \in \mathcal{L}(\mathcal{H})$ can be written as a linear combination of four unitary operators.

Proof We first write $T = A + iB$ with A and B selfadjoint. We may assume that $\|A\| \leq 1$ and $\|B\| \leq 1$. Define $U = A + i\sqrt{I - A^2}$. Then $U^* = A - i\sqrt{I - A^2}$ and we have $U^*U = UU^* = I$, $A = \frac{1}{2}(U + U^*)$. Similarly, B is the linear combination of two unitary operators. $\square$

2.6 The Polar Decomposition of a Bounded Operator

For any $T \in \mathcal{L}(\mathcal{H})$, we denote $\ker(T) = \{\varphi \in \mathcal{H} \mid T\varphi = 0\}$.

Lemma 2.1 *If $T \in \mathcal{L}(\mathcal{H})$, then $\overline{T(\mathcal{H})} = \ker(T^*)^\perp$.*

Proof As $T(\mathcal{H})^\perp = \overline{T(\mathcal{H})}^\perp$, it is enough to show that $\ker(T^*) = T(\mathcal{H})^\perp$. The following conditions are equivalent for a vector $\varphi \in \mathcal{H}$: $\varphi \in \ker(T^*)$, $\langle T^*\varphi | \psi \rangle = 0$ for all $\psi \in \mathcal{H}$, $\langle \varphi | T\psi \rangle = 0$ for all $\psi \in \mathcal{H}$. $\square$

For every operator $T \in \mathcal{L}(\mathcal{H})$ we denote by $\text{supp}(T)$ the (orthogonal) projection of $\mathcal{H}$ onto the closed subspace $\ker(T)^\perp$. We say that an operator $V \in \mathcal{L}(\mathcal{H})$ is a *partial isometry* or *partially isometric* if $\|V\text{supp}(V)\varphi\| = \|\text{supp}(V)\varphi\|$ for all $\varphi \in \mathcal{H}$. We then say that $\text{supp}(V)$ is the *initial projection* of V , and the projection P_M onto the (closed since $\text{supp}(V)(\mathcal{H})$ is closed) subspace $M = \{V\varphi \mid \varphi \in \mathcal{H}\} = \{V\text{supp}(V)\varphi \mid \varphi \in \mathcal{H}\}$ of $\mathcal{H}$ is the *final projection* of V .

Theorem 2.16 *Let $V \in \mathcal{L}(\mathcal{H})$ be partially isometric, P the initial projection of V and Q the final projection of V . Then*

- (a) *V^* is partially isometric, Q is the initial projection of V^* , P is the final projection of V^* , and*
- (b) *$V^*V = P$, $VV^* = Q$.*

Proof Let $\varphi \in P(\mathcal{H})$ and $\psi = V\varphi \in Q(\mathcal{H})$. The polarisation identity implies that restricted to the subspace $P(\mathcal{H})$, V preserves the inner product, so that for each $\xi \in \mathcal{H}$ we get

$$\langle \xi | \varphi \rangle = \langle P\xi | \varphi \rangle + \langle (I - P)\xi | \varphi \rangle = \langle P\xi | \varphi \rangle = \langle VP\xi | V\varphi \rangle = \langle V\xi | \psi \rangle = \langle \xi | V^*\psi \rangle.$$

Thus $V^*\psi = \varphi$. It follows that $V^*|Q(\mathcal{H})$ is the inverse of the map $V|P(\mathcal{H}) : P(\mathcal{H}) \rightarrow Q(\mathcal{H})$ and hence isometric. In view of Lemma 2.1, $Q = \text{supp}(V^*)$ and $P(\mathcal{H}) = \ker(V)^\perp = \overline{V^*(\mathcal{H})} = V^*(\mathcal{H})$, and (a) is proved. Clearly also (b) is true, since $V^*|Q(\mathcal{H})$ and $V|P(\mathcal{H})$ are the inverses of each other. $\square$

Theorem 2.17 *If $V \in \mathcal{L}(\mathcal{H})$ is such that V^*V is a projection, then V is partially isometric.*

Proof Let $V^*V = P$ be a projection. If $\varphi \in \mathcal{H}$, then $\|V\varphi\|^2 = \langle \varphi | V^*V\varphi \rangle = \langle \varphi | P\varphi \rangle = \|P\varphi\|^2$, and so $V|P(\mathcal{H})$ is isometric and $V|(P(\mathcal{H}))^\perp = 0$. Here $P = \text{supp}(V)$, because $P(\mathcal{H})^\perp = \ker(V)$. $\square$

Theorem 2.18 *Let $T \in \mathcal{L}(\mathcal{H})$. There is a uniquely determined pair of operators $V, A \in \mathcal{L}(\mathcal{H})$ such that*

- (i) $T = VA$,
- (ii) $A \geq 0$, and
- (iii) V is a partially isometric operator whose initial projection is $\text{supp}(A)$.

Then $A = |T|$ and $\text{supp}(V) = \text{supp}(|T|) = \text{supp}(T)$. Moreover, $|T| = V^*T$.

Proof Let $P = \text{supp}(T)$, $Q = \text{supp}(T^*)$. If $\varphi \in \mathcal{H}$, we have $\|T\varphi\|^2 = \langle \varphi | T^*T \varphi \rangle = \langle \varphi | |T|^2 \varphi \rangle = \| |T| \varphi \|^2$, so that $\text{supp}(|T|) = P$ and hence in view of Lemma 2.1, $|T|(\mathcal{H}) = P(\mathcal{H})$, as $|T|^* = |T|$. This shows, moreover, that the mapping $|T|\varphi \mapsto T\varphi$ is a well-defined (for $T\varphi - T\psi = 0$, when $|T|\varphi - |T|\psi = 0$) linear isometry from $|T|(\mathcal{H})$ onto $T(\mathcal{H})$ and can hence be uniquely extended to a linear isometry $V_0 : P(\mathcal{H}) \rightarrow Q(\mathcal{H})$ (for $Q(\mathcal{H}) = \overline{T(\mathcal{H})}$ by Lemma 2.1). Let now V be a partially isometric operator such that $\text{supp}(V) = P$ and $V|P(\mathcal{H}) = V_0$, i.e. V is the map $\varphi \mapsto V_0 P \varphi$. When we choose $A = |T|$, the requirements (i)–(iii) are fulfilled. From the construction it is clear that $\text{supp}(V) = \text{supp}(|T|) = \text{supp}(T)$ and $|T| = P|T| = V^*V|T| = V^*T$ (see Theorem 2.16 (b)). We next prove the uniqueness part. Let V and A be such that the conditions (i)–(iii) are valid. Then $T^*T = AV^*VA = A \text{supp}(V)A = A^2$ (see Theorem 2.16 (b)), and so by the uniqueness of the positive square root (Theorem 2.15) we get $A = |T|$. From this also the values of V are uniquely determined. $\square$

Definition 2.5 The representation $T = V|T|$ mentioned in Theorem 2.18 is called the *polar decomposition* of T .

2.7 Orthonormal Sets

If $\varphi, \psi \in \mathcal{H}$ and $\langle \varphi | \psi \rangle = 0$, we say that φ and ψ are *mutually orthogonal* and we write $\varphi \perp \psi$. If $K \subset \mathcal{H}$, $\varphi \in \mathcal{H}$ and $\varphi \perp \psi$ for all $\psi \in K$, we denote $\varphi \perp K$. The set $K \subset \mathcal{H}$ is *orthogonal*, if $\varphi \perp \psi$ whenever $\varphi, \psi \in K$, $\varphi \neq \psi$. The family $(\varphi_i)_{i \in \mathcal{I}}$ (and especially a sequence) in $\mathcal{H}$ is *orthogonal* if $\varphi_i \perp \varphi_j$ whenever $i \neq j$. A set or a family is *orthonormal*, if it is orthogonal and in addition every vector in it has norm 1.

A family $(c_i)_{i \in \mathcal{I}}$ of complex numbers is *summable*, if there is a constant $M \in [0, \infty)$ such that $\sum_{i \in F} |c_i| \leq M$ for every finite subset F of $\mathcal{I}$. Then for all $n \in \mathbb{N}$ $|c_i| \geq \frac{1}{n}$ for at most a finite number of $i \in \mathcal{I}$, so that the set $\{i \in \mathcal{I} \mid |c_i| > 0\}$ is at most countable: $\{i_1, i_2, \dots\}$, and the series $\sum_{k=1}^{\infty} c_{i_k}$ is absolutely convergent (in the case of an infinite set), so that independently of the numbering we may define the *sum* of the family $(c_i)_{i \in \mathcal{I}}$ as $\sum_{i \in \mathcal{I}} c_i = \sum_{k=1}^{\infty} c_{i_k}$. If $\{i \in \mathcal{I} \mid |c_i| > 0\}$ is finite, the definition of the sum is obvious (in the case of the empty set the sum is 0). We denote $\sum_{i \in \mathcal{I}} |c_i| = \infty$ if the family is not summable.

Lemma 2.2 *Let $\varphi \in \mathcal{H}$.*

- (a) *If $(\varphi_k)_{k=1}^n$ is a finite orthonormal family in $\mathcal{H}$, then for all $c_1, \dots, c_n \in \mathbb{C}$ we have*

$$\left\| \varphi - \sum_{k=1}^n c_k \varphi_k \right\| \geq \left\| \varphi - \sum_{k=1}^n \langle \varphi_k | \varphi \rangle \varphi_k \right\|,$$

and equality holds only if $c_k = \langle \varphi_k | \varphi \rangle$ for all $k = 1, \dots, n$. Moreover,

$$\sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2 \leq \|\varphi\|^2. \quad (2.1)$$

- (b) *If $K \subset \mathcal{H}$ is an orthonormal set, then the family $(|\langle \psi | \varphi \rangle|^2)_{\psi \in K}$ is summable and*

$$\sum_{\psi \in K} |\langle \psi | \varphi \rangle|^2 \leq \|\varphi\|^2 \quad (\text{Bessel's inequality}).$$

Proof (a) As

$$\begin{aligned} \left\| \varphi - \sum_{k=1}^n c_k \varphi_k \right\|^2 &= \|\varphi\|^2 - \sum_{k=1}^n \overline{c_k} \langle \varphi_k | \varphi \rangle - \sum_{k=1}^n c_k \langle \varphi | \varphi_k \rangle + \sum_{k=1}^n |c_k|^2 \\ &= \|\varphi\|^2 + \sum_{k=1}^n [|\langle \varphi_k | \varphi \rangle|^2 - \overline{c_k} \langle \varphi_k | \varphi \rangle - c_k \overline{\langle \varphi_k | \varphi \rangle} + |c_k|^2] - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2 \\ &= \|\varphi\|^2 + \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle - c_k|^2 - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2, \end{aligned}$$

the choice $c_k = \langle \varphi_k | \varphi \rangle$ yields (2.1). Moreover, $\|\varphi - \sum_{k=1}^n c_k \varphi_k\|^2$ also gets its smallest possible value $\|\varphi\|^2 - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2$ with this and only this choice.

- (b) The claim is an immediate consequence of (a). $\square$

Theorem 2.19 *Let (φ_n) be an orthogonal sequence in $\mathcal{H}$.*

- (a) *The series $\sum_{n=1}^{\infty} \varphi_n$ converges if and only if $\sum_{n=1}^{\infty} \|\varphi_n\|^2 < \infty$. If $\varphi = \sum_{n=1}^{\infty} \varphi_n$, then $\|\varphi\|^2 = \sum_{n=1}^{\infty} \|\varphi_n\|^2$.*
- (b) *If (φ_n) is orthonormal, then for every $\varphi \in \mathcal{H}$ the series $\sum_{k=1}^{\infty} \langle \varphi_k | \varphi \rangle \varphi_k$ converges, and if ψ is its sum, then $(\varphi - \psi) \perp \varphi_n$ for all $n \in \mathbb{N}$.*

Proof (a) If $n > m$, then by orthogonality

$$\begin{aligned} \left\| \sum_{k=1}^n \varphi_k - \sum_{k=1}^m \varphi_k \right\|^2 &= \left\| \sum_{k=m+1}^n \varphi_k \right\|^2 = \sum_{k=m+1}^n \|\varphi_k\|^2 \\ &= \sum_{k=1}^n \|\varphi_k\|^2 - \sum_{k=1}^m \|\varphi_k\|^2. \end{aligned}$$

This implies that the sequence of the partial sums $s_n = \sum_{k=1}^n \varphi_k$ is a Cauchy sequence in $\mathcal{H}$ if and only if the sequence of the partial sums $\sum_{k=1}^n \|\varphi_k\|^2$ is a Cauchy sequence in $\mathbb{R}$. This proves the first claim as $\mathcal{H}$ and $\mathbb{R}$ are complete. Since $\|s_n\|^2 = \sum_{k=1}^n \|\varphi_k\|^2$, the continuity of the norm shows that

$$\|\varphi\|^2 = \lim_{n \rightarrow \infty} \|s_n\|^2 = \lim_{n \rightarrow \infty} \sum_{k=1}^n \|\varphi_k\|^2, \text{ when } \varphi = \sum_{k=1}^{\infty} \varphi_k.$$

(b) The convergence of the series follows from (a) and Lemma 2.2 (b), since $\|\langle \varphi_k | \varphi \rangle \varphi_k\|^2 = |\langle \varphi_k | \varphi \rangle|^2$. Since the map $\xi \mapsto \langle \varphi_n | \xi \rangle$ is continuous, we obtain

$$\langle \varphi - \psi | \varphi_n \rangle = \langle \varphi | \varphi_n \rangle - \sum_{k=1}^{\infty} \langle \varphi | \varphi_k \rangle \langle \varphi_k | \varphi_n \rangle = 0. \quad \square$$

In the next theorem the notation $\varphi = \sum_{\xi \in K} \langle \xi | \varphi \rangle \xi$ means that $\langle \xi | \varphi \rangle \neq 0$ for at most a countable number of vectors ξ : $\xi_1, \xi_2, \dots$, and that independently of the numbering, $\varphi = \sum_n \langle \varphi_n | \varphi \rangle \varphi_n$ (a convergent series or a finite or “empty” sum ($= 0$)). In (iii) and (iv) there is a summable family.

Theorem 2.20 *Let $K \subset \mathcal{H}$ be an orthonormal set. The following conditions are equivalent:*

- (i) *if $\varphi \perp K$, then $\varphi = 0$;*
- (ii) $\varphi = \sum_{\xi \in K} \langle \xi | \varphi \rangle \xi$ *for all $\varphi \in \mathcal{H}$;*
- (iii) $\langle \psi | \varphi \rangle = \sum_{\xi \in K} \langle \xi | \varphi \rangle \overline{\langle \xi | \psi \rangle}$ *for all $\varphi, \psi \in \mathcal{H}$;*
- (iv) $\|\varphi\|^2 = \sum_{\xi \in K} |\langle \xi | \varphi \rangle|^2$ *for all $\varphi \in \mathcal{H}$;*
- (v) *the vector subspace M of $\mathcal{H}$ generated by the set K is dense in $\mathcal{H}$.*

Proof Assume first (i). Let $\varphi \in \mathcal{H}$. According to Lemma 2.2 (b) $\langle \xi | \varphi \rangle \neq 0$ for at most a countable number of the vectors $\xi \in K$; let $\xi_1, \xi_2, \dots$ be their arbitrary numbering. By Theorem 2.19 (b) there is $\psi = \sum_n \langle \xi_n | \varphi \rangle \xi_n$, and $\langle \xi_n | \varphi - \psi \rangle = 0$ for all $n = 1, 2, \dots$. Also, if $\xi \in K \setminus \{\xi_1, \xi_2, \dots\}$ then $\langle \xi | \varphi - \psi \rangle = \langle \xi | \varphi \rangle - \sum_n \langle \xi_n | \varphi \rangle \langle \xi | \xi_n \rangle = 0$. Hence $\varphi - \psi \perp K$, so that by (i) $\varphi = \psi$ and so (ii) holds.

Assume now (ii). By the Cauchy–Schwarz inequality and Lemma 2.2, for every finite set $F \subset K$ we have

$$\sum_{\xi \in F} |\langle \xi | \varphi \rangle \overline{\langle \xi | \psi \rangle}| \leq \left(\sum_{\xi \in F} |\langle \xi | \varphi \rangle|^2 \right)^{\frac{1}{2}} \left(\sum_{\xi \in F} |\langle \xi | \psi \rangle|^2 \right)^{\frac{1}{2}} \leq \|\varphi\| \|\psi\|,$$

so that the right hand side of (iii) is defined. By Lemma 2.2 (b) the set $\{\xi \in K | \langle \xi | \varphi \rangle \neq 0 \text{ or } \langle \xi | \psi \rangle \neq 0\}$ can be enumerated: $\xi_1, \xi_2, \dots$. We may assume that this is an infinite

sequence. Denote

$$\begin{aligned}\varphi_n &= \sum_{k=1}^n \langle \xi_k | \varphi \rangle \xi_k \text{ and } \psi_n = \sum_{k=1}^n \langle \xi_k | \psi \rangle \xi_k, \text{ so that} \\ \langle \psi_n | \varphi_n \rangle &= \sum_{k=1}^n \sum_{p=1}^n \langle \langle \xi_p | \psi \rangle \xi_p | \langle \xi_k | \varphi \rangle \xi_k \rangle = \sum_{k=1}^n \overline{\langle \xi_k | \psi \rangle} \langle \xi_k | \varphi \rangle.\end{aligned}$$

When $n \rightarrow \infty$, by the continuity of the inner product we get (iii). Choosing $\varphi = \psi$ we see at once that (iii) implies (iv). One also sees immediately that (iv) implies (i) so that the conditions (i)–(iv) are equivalent. On the other hand, (ii) implies (v), and (v) implies (i), for if $\varphi \perp K$ then $\varphi \perp M$. $\square$

Definition 2.6 We say that an orthonormal set $K \subset \mathcal{H}$ satisfying the equivalent conditions in the preceding theorem is an *orthonormal basis* or a *complete orthonormal system* in $\mathcal{H}$. The equations appearing in conditions (iii) and (iv) in Theorem 2.20 are called the *Parseval identities*. The numbers $\langle \xi | \varphi \rangle$, $\xi \in K$, are called the *Fourier coefficients* of φ with respect to K .

Every Hilbert space has an orthonormal basis. We prove this using the Zorn lemma. If the space is *separable* (i.e. if it has a countable dense subset), the use of Zorn's lemma (and other methods equivalent to the axiom of choice) could be avoided, but we will not elaborate this question any further.

Theorem 2.21 *If $L \subset \mathcal{H}$ is an orthonormal set, there is an orthonormal basis of $\mathcal{H}$ containing L .*

Proof We equip the set $\mathcal{F}$ of the orthonormal subsets of $\mathcal{H}$ containing L with the inclusion order. Every linearly ordered subset $\mathcal{F}_0$ of $\mathcal{F}$ has an upper bound, namely $\bigcup_{F \in \mathcal{F}_0} F$. According to Zorn's lemma $\mathcal{F}$ has a maximal element K . The maximality means that K satisfies condition (i) in Theorem 2.20. $\square$

Theorem 2.22 *Let K be an orthonormal basis of $\mathcal{H}$. The following conditions are equivalent:*

- (i) *the set K is at most countable;*
- (ii) *the space $\mathcal{H}$ is separable.*

Proof Assume first (i). The set of the linear combinations of the elements of K with coefficients whose real and imaginary parts are rational, is countable. Using condition (v) in Theorem 2.20 and the density of the set of rational numbers in $\mathbb{R}$ we easily see that this set is dense in $\mathcal{H}$, and so (ii) holds. Next assume (ii). If $\varphi, \psi \in K$ and $\varphi \neq \psi$, then $\|\varphi - \psi\|^2 = 2$, so that the open balls $B(\varphi, \frac{1}{2}\sqrt{2})$ and $B(\psi, \frac{1}{2}\sqrt{2})$ are disjoint. Each one of these meets a certain at most countable set, and so (i) holds. $\square$

Remark 2.5 If K and L are orthonormal bases of the same Hilbert space $\mathcal{H}$ then they have the same cardinality (exercise). This cardinality is called the *Hilbert dimension* of $\mathcal{H}$. If M is a closed subspace of $\mathcal{H}$ we call the Hilbert dimension of its orthogonal complement the *Hilbert codimension* of M . If $\mathcal{H}$ is finite-dimensional, then its Hilbert dimension is just its dimension in the usual algebraic sense. (This will follow, e.g., from Theorem 3.3.) However, if $\mathcal{H} = \ell^2$, its Hilbert dimension is the cardinality of $\mathbb{N}$, but it does have an uncountable linearly independent subset. Such a set is, for instance, the set $\left\{ \sum_{n=1}^{\infty} c^n e_n \mid 0 < c < 1 \right\}$ where, for each $n \in \mathbb{N}$, e_n is the function taking the value 1 at n and zero elsewhere (exercise). $\triangleleft$

2.8 Direct Sums of Hilbert Spaces

In this section we assume that $\mathcal{I}$ is a set and $\mathcal{H}_i$ is a Hilbert space for every $i \in \mathcal{I}$. The Cartesian product $\prod_{i \in \mathcal{I}} \mathcal{H}_i$ consists of all vector families $(\varphi_i) = (\varphi_i)_{i \in \mathcal{I}}$ in $\cup_{i \in \mathcal{I}} \mathcal{H}_i$ having the property that $\varphi_i \in \mathcal{H}_i$ for all $i \in \mathcal{I}$. Clearly it is a vector space with respect to the pointwise operations: $\alpha(\varphi_i) = (\alpha\varphi_i)$, $(\varphi_i) + (\psi_i) = (\varphi_i + \psi_i)$. We now consider an important subspace.

Proposition 2.7 *The set $E = \left\{ (\varphi_i)_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} \mathcal{H}_i \mid \sum_{i \in \mathcal{I}} \|\varphi_i\|^2 < \infty \right\}$ is a vector subspace of the Cartesian product $\prod_{i \in \mathcal{I}} \mathcal{H}_i$. If $(\varphi_i), (\psi_i) \in E$, then the family $(\langle \varphi_i \mid \psi_i \rangle)_{i \in \mathcal{I}}$ is summable. Moreover, the mapping $h : E \times E \rightarrow \mathbb{C}$ defined by the formula $h((\varphi_i), (\psi_i)) = \sum_{i \in \mathcal{I}} \langle \varphi_i \mid \psi_i \rangle$ is an inner product and equipped with this inner product E is a Hilbert space.*

We leave the rather straightforward proof as an exercise; the central ideas are already present in the case of the scalar sequence space ℓ^2 . The Hilbert space E in the above proposition is called the *direct sum* or *Hilbert sum* of the family $(\mathcal{H}_i)_{i \in \mathcal{I}}$ of Hilbert spaces. We use the notations $\bigoplus_{i \in \mathcal{I}} \mathcal{H}_i$ and $\sum_{i \in \mathcal{I}}^{\oplus} \mathcal{H}_i$ for it.

Example 2.4 If we take $\mathcal{H}_i = \mathbb{C}$ for all $i \in \mathcal{I}$, then we denote $\bigoplus_{i \in \mathcal{I}} \mathcal{H}_i = \ell^2(\mathcal{I})$. If $\mathcal{I}$ happens to be an orthonormal basis of a Hilbert space $\mathcal{H}$, then the mapping $\varphi \mapsto \chi_{\{\varphi\}}$ from $\mathcal{I}$ to $\ell^2(\mathcal{I})$ (where $\chi_{\{\varphi\}}$ is the function taking the value 1 at φ and zero elsewhere) extends to an isometric isomorphism from $\mathcal{H}$ onto $\ell^2(\mathcal{I})$. This follows readily from Theorem 2.20 and Proposition 2.7. $\triangleleft$

Example 2.5 Let $\mathcal{H}$ be a Hilbert space and $(\mathcal{H}_i)_{i \in \mathcal{I}}$ a family of closed subspaces of $\mathcal{H}$ such that $\mathcal{H}_i \perp \mathcal{H}_j$ whenever $i \neq j$ and $\mathcal{H}$ itself is the only closed subspace containing every $\mathcal{H}_i$. For each $\varphi \in \mathcal{H}$ let $\Phi\varphi$ be the family $(P_i\varphi)_{i \in \mathcal{I}}$ where P_i is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_i$. Then Φ is an isometric isomorphism from $\mathcal{H}$ onto $\bigoplus_{i \in \mathcal{I}} \mathcal{H}_i$ (exercise). For this reason we may sometimes use the identification $\mathcal{H} = \bigoplus_{i \in \mathcal{I}} \mathcal{H}_i$. $\triangleleft$

2.9 Tensor Products of Hilbert Spaces

One can define the general algebraic tensor product of vector spaces or modules, but here we only consider the case of Hilbert spaces, and the result will be a Hilbert space which could be viewed as the completion of the algebraic tensor with respect to a natural inner product. We are not, however, going to develop this latter point of view now. We begin with a lemma showing the existence of the (Hilbert space) tensor product of two Hilbert spaces. To avoid trivialities we assume that our Hilbert spaces contain non-zero elements.

Lemma 2.3 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. There is a Hilbert space $\mathcal{H}_{\otimes}$ with a bilinear map $f : \mathcal{H} \times \mathcal{K} \rightarrow \mathcal{H}_{\otimes}$ such that*

- (i) *the subspace generated by the set $\{f(\varphi, \psi) \mid \varphi \in \mathcal{H}, \psi \in \mathcal{K}\}$ is dense in $\mathcal{H}_{\otimes}$ and*
- (ii) *$\langle f(\varphi_1, \psi_1) \mid f(\varphi_2, \psi_2) \rangle = \langle \varphi_1 \mid \varphi_2 \rangle \langle \psi_1 \mid \psi_2 \rangle$ for all $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$.*

Proof Choose an orthonormal basis H (resp. K) for $\mathcal{H}$ (resp. $\mathcal{K}$). Consider the Hilbert space $\mathcal{H} = \ell^2(H \times K)$ of square-summable families $(c_{\xi, \eta})_{(\xi, \eta) \in H \times K}$. Define a map $f : \mathcal{H} \times \mathcal{K} \rightarrow \ell^2(H \times K)$ by setting

$$f(\varphi, \psi) = (\langle \xi \mid \varphi \rangle \langle \eta \mid \psi \rangle)_{(\xi, \eta) \in H \times K}$$

for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$. The map f is clearly bilinear and it satisfies the condition (i) since $f(\xi, \eta) = \chi_{\{(\xi, \eta)\}}$ whenever $\xi \in H$ and $\eta \in K$. Using Theorem 2.20 we also get

$$\begin{aligned} \langle f(\varphi_1, \psi_1) \mid f(\varphi_2, \psi_2) \rangle &= \sum_{(\xi, \eta) \in H \times K} \overline{\langle \xi \mid \varphi_1 \rangle \langle \eta \mid \psi_1 \rangle} \langle \xi \mid \varphi_2 \rangle \langle \eta \mid \psi_2 \rangle \\ &= \sum_{\xi \in H} \langle \varphi_1 \mid \xi \rangle \langle \xi \mid \varphi_2 \rangle \sum_{\eta \in K} \langle \psi_1 \mid \eta \rangle \langle \eta \mid \psi_2 \rangle \\ &= \langle \varphi_1 \mid \varphi_2 \rangle \langle \psi_1 \mid \psi_2 \rangle \end{aligned}$$

for all $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$, so that the condition (ii) also holds. $\square$

Next we show that the above pair $(\mathcal{H}_{\otimes}, f)$ is essentially unique.

Lemma 2.4 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and suppose that $\mathcal{H}_{\otimes}$ and f are as in Lemma 2.3. Suppose that also the pair $(\mathcal{H}', f')$ satisfies the conditions (i) and (ii) of Lemma 2.3. Then there is an isometric isomorphism $g : \mathcal{H}_{\otimes} \rightarrow \mathcal{H}'$ such that $g \circ f = f'$.*

Proof From the conditions (i) and (ii) of Lemma 2.3 it follows that

$$\begin{aligned} \left\| \sum_{j=1}^n f(\varphi_j, \psi_j) \right\|^2 &= \sum_{j,k=1}^n \langle f(\varphi_j, \psi_j) | f(\varphi_k, \psi_k) \rangle \\ &= \sum_{j,k=1}^n \langle \varphi_j | \varphi_k \rangle \langle \psi_j | \psi_k \rangle = \left\| \sum_{j=1}^n f'(\varphi_j, \psi_j) \right\|^2 \end{aligned}$$

for all $\varphi_j \in \mathcal{H}$ and $\psi_j \in \mathcal{K}$, $j = 1, \dots, n$. Thus the map

$$\sum_{j=1}^n f(\varphi_j, \psi_j) \mapsto \sum_{j=1}^n f'(\varphi_j, \psi_j)$$

is a well-defined isometric isomorphism from a dense subspace of $\mathcal{H}_{\otimes}$ onto a dense subspace of $\mathcal{H}'$. We obtain the isomorphism g by extending the above map by continuity to the whole of $\mathcal{H}_{\otimes}$. $\square$

Definition 2.7 Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. If $\mathcal{H}_{\otimes}$ is a Hilbert space and $f : \mathcal{H} \times \mathcal{K} \rightarrow \mathcal{H}_{\otimes}$ is a map satisfying the conditions (i) and (ii) of Lemma 2.3 we say that the pair $(\mathcal{H}_{\otimes}, f)$ or simply $\mathcal{H}_{\otimes}$ is a *Hilbert tensor product* of the spaces $\mathcal{H}$ and $\mathcal{K}$.

If there seems to be no danger of confusion, we may call the Hilbert tensor product $(\mathcal{H}_{\otimes}, f)$ (or $\mathcal{H}_{\otimes}$) simply the *tensor product* of $\mathcal{H}$ and $\mathcal{K}$. According to Lemma 2.4 the tensor product $(\mathcal{H}_{\otimes}, f)$ of $\mathcal{H}$ and $\mathcal{K}$ is uniquely defined up to an isometric isomorphism. We thus use the notations $\mathcal{H}_{\otimes} = \mathcal{H} \otimes \mathcal{K}$ and $f(\varphi, \psi) = \varphi \otimes \psi$ in the sequel. From the construction of the tensor product in Lemma 2.3 we also see that there is an isometric isomorphism $g : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{K} \otimes \mathcal{H}$ such that $g(\varphi \otimes \psi) = \psi \otimes \varphi$ for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$. If $K \subset \mathcal{H}$ and $L \subset \mathcal{K}$ are orthonormal bases, then according to the conditions (i) and (ii) of Lemma 2.3 the set $\{\xi \otimes \eta \mid \xi \in K, \eta \in L\}$ is an orthonormal basis of the tensor product $\mathcal{H} \otimes \mathcal{K}$. Note that although the so-called *simple tensors* $\varphi \otimes \psi$ (topologically) generate the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$, a vector $\zeta \in \mathcal{H} \otimes \mathcal{K}$ need not be a simple tensor. Any vector $\zeta \in \mathcal{H} \otimes \mathcal{K}$ can be represented as a linear combination of simple tensors or as a limit of such linear combinations. As was already mentioned, the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$ is in fact the Hilbert space completion of the general algebraic tensor product of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$.

One can also form, with essentially the same process as for two Hilbert spaces, Hilbert tensor products $\mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$ for more than two Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_n$. It is also possible to define $\mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$ by induction by using the notion of the tensor product of two Hilbert space and the easily verified associativity result $(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3 = \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$.

Remark 2.6 Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Let K be an orthonormal basis of $\mathcal{K}$ and choose $\mathcal{H}_{\mu} = \mathcal{H}$ for every $\mu \in K$. There is an isometric isomorphism

$\Phi : \sum_{\mu \in K}^{\oplus} \mathcal{H}_{\mu} \rightarrow \mathcal{H} \otimes \mathcal{K}$ satisfying $\Phi(\xi^{(\nu)}) = \xi \otimes \nu$ for all $\nu \in K$ and $\xi \in \mathcal{H}$ where $\xi^{(\nu)} \in \sum_{\mu \in K}^{\oplus} \mathcal{H}_{\mu}$ is the family taking the value ξ at ν and zero elsewhere. We leave the verification as an exercise. Accordingly, we sometimes identify the tensor product $\mathcal{H} \otimes \mathcal{K}$ with the direct sum $\sum_{\mu \in K}^{\oplus} \mathcal{H}_{\mu}$. $\triangleleft$

2.10 Exercises

Unless otherwise stated, $\mathcal{H}$ is an arbitrary Hilbert space.

1. Prove Theorem 2.2.
2. Prove Theorem 2.3.
3. Prove Theorem 2.6.
4. Prove Theorem 2.9.
5. Prove Theorem 2.11.
6. Prove Theorem 2.12.
7. Prove Theorem 2.13.
8. Prove by using the polarisation identity that a linear map $T : \mathcal{H} \rightarrow \mathcal{H}$ is isometric if and only if it preserves the inner products (i.e. $\langle T\varphi | T\psi \rangle = \langle \varphi | \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$).
9. Show that an operator $T \in \mathcal{L}(\mathcal{H})$ is unitary if and only if T is isometric and surjective.
10. Let $T \in \mathcal{L}(\mathcal{H})$. Show that

$$\|T\| = \sup \{ |\langle \varphi | T\psi \rangle| \mid \|\varphi\| \leq 1, \|\psi\| \leq 1 \}.$$

11. Let M be a closed subspace of $\mathcal{H}$. Let $f : M \rightarrow \mathbb{C}$ be a continuous linear functional. Show that there is one and only one continuous linear functional $\tilde{f} : \mathcal{H} \rightarrow \mathbb{C}$ such that $\tilde{f}|_M = f$ and $\|\tilde{f}\| = \|f\|$. Show that $\tilde{f}(\varphi) = 0$ for all $\varphi \in M^{\perp}$.
12. Let $(T_i)_{i \in \mathcal{I}}$ be a net in $\mathcal{L}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$. Prove: If the net $(T_i)_{i \in \mathcal{I}}$ converges strongly to T , then it converges weakly to T .
13. Let $(\mathcal{I}, \geq)$ be a directed set, let $S_i \in \mathcal{L}(\mathcal{H})$ and $T_i \in \mathcal{L}(\mathcal{H})$ for all $i \in \mathcal{I}$, and let $S, T \in \mathcal{L}(\mathcal{H})$. Prove: If $S_i \rightarrow^s S$, $T_i \rightarrow^s T$ and $\sup \{ \|S_i\| \mid i \in \mathcal{I} \} < \infty$, then $S_i T_i \rightarrow^s ST$.
14. Prove: If $T \in \mathcal{L}_s(\mathcal{H})^+$ then $T^n \in \mathcal{L}_s(\mathcal{H})^+$ for all $n \in \mathbb{N}$.
15. Consider the Hilbert space $\ell^2 = \{ f : \mathbb{N} \rightarrow \mathbb{C} \mid \sum_{k=1}^{\infty} |f(k)|^2 < \infty \}$ (equipped with the inner product $\langle f | g \rangle = \sum_{k=1}^{\infty} \overline{f(k)}g(k)$ as usual). Define $T : \ell^2 \rightarrow \ell^2$ by setting $Tf(1) = 0$ and $Tf(k) = f(k-1)$ whenever $k \geq 2$.
 - (a) Show that T is isometric.
 - (b) Is T unitary?

16. (a) What is the adjoint T^* of the operator T defined in the previous exercise?
 (b) Let T be as in the previous exercise. Does the sequence $(T^n)_{n \in \mathbb{N}}$ converge strongly? Does it converge weakly?
17. Define the mapping $T : \ell^2 \rightarrow \ell^2$ via the formula $(Tf)(n) = f(n+1)$. Show that the adjoint of T satisfies the formula $(T^*f)(n) = f(n-1)$ when $n \geq 2$, and $(T^*f)(1) = 0$.
18. We saw in Exercise 15 that the operator T defined there is isometric. Is T^* isometric? Is any one of the formulas $T^*T = I$, $TT^* = I$ true (when I is the identity operator of ℓ^2)? Is T normal?
19. Let $P, Q \in \mathcal{P}(\mathcal{H})$ (i.e. P and Q are projections). Denote $A_n = (PQP)^n$ for all $n \in \mathbb{N}$. Show that (A_n) is a decreasing sequence in $\mathcal{L}_s(\mathcal{H})^+$ and converges weakly and strongly to some $B \in \mathcal{L}_s(\mathcal{H})^+$.
20. We retain the notations and assumptions of the preceding exercise. Prove the following claims:
 - (a) $B \in \mathcal{P}(\mathcal{H})$;
 - (b) $BQP = B$;
 - (c) $B \leq Q$; (hint: show that $B(I - Q)[B(I - Q)]^* = 0$.)
 - (d) $PB = P$, i.e. $B \leq P$;
 - (e) $B = P \wedge Q$.
21. Assume that $(T_i)_{i \in \mathcal{I}}$ is a net of operators in $\mathcal{L}(\mathcal{H})$ and $\text{w-lim } T_i = T \in \mathcal{L}(\mathcal{H})$. Is T necessarily selfadjoint if each T_i is selfadjoint?
22. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded function and $T_g : \ell^2 \rightarrow \ell^2$ the operator defined by the formula $T_g f = gf$. Show that T_g is continuous and $\|T\| = \|g\|_\infty$ where $\|g\|_\infty = \sup_{n \in \mathbb{N}} |g(n)|$.
23. In the situation of the preceding exercise, give necessary and sufficient conditions for T_g to be
 - (a) normal,
 - (b) selfadjoint,
 - (c) unitary,
 - (d) a projection.
24. We use the notation of Exercise 22. Let $(\mathcal{I}, \geq)$ be a directed set and $g_i : \mathbb{N} \rightarrow \mathbb{C}$ a bounded function for each $i \in \mathcal{I}$. Assume that $\sup_{i \in \mathcal{I}} \|g_i\|_\infty < \infty$. Show that for a bounded function $g : \mathbb{N} \rightarrow \mathbb{C}$ we have $\text{s-lim } T_{g_i} = T_g$, if and only if $\lim g_i(n) = g(n)$ for all $n \in \mathbb{N}$.
25. Is the claim in the previous exercise true without the assumption that $\sup_{i \in \mathcal{I}} \|g_i\|_\infty < \infty$?
26. Show that the absolute value $|T|$ of an operator $T \in \mathcal{L}(\mathcal{H})$ is the unique positive operator A satisfying $\|A\varphi\| = \|T\varphi\|$ for all $\varphi \in \mathcal{H}$.
27. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded function and $T_g \in \mathcal{L}(\mathcal{H})$ as in Exercise 22. Describe the polar decomposition of the operator T_g .
28. Describe the polar decomposition of the operator T defined in Exercise 17. Describe also the polar decomposition of T^* .
29. Let $T = VA$ be the polar decomposition of the operator $T \in \mathcal{L}(\mathcal{H})$. Give a necessary and sufficient condition concerning T

- (a) for V to be isometric,
 - (b) for V to be unitary.
30. Show that there is an isometric bijection $J : H \rightarrow H$ which is conjugate linear, i.e. $J(\alpha\varphi + \beta\psi) = \overline{\alpha}J\varphi + \overline{\beta}J\psi$ for all $\varphi, \psi \in H$ and $\alpha, \beta \in \mathbb{C}$. (Hint: “practise” with ℓ^2 .)
 31. Show using the preceding exercise that there is an isometric linear bijection $\Phi : \mathcal{H} \rightarrow \mathcal{H}^*$, where $\mathcal{H}^*$ is the dual of $\mathcal{H}$, i.e. the Banach space of continuous linear functionals $f : \mathcal{H} \rightarrow \mathbb{C}$.
 32. Prove Proposition 2.7 except for the completeness of E .
 33. Prove that the inner product space in Proposition 2.7 is complete.
 34. Prove the claim made in Example 2.4.
 35. Prove the statement left as an exercise in Example 2.5.
 36. Complete the details of Remark 2.6.
 37. Show that if K and L are orthonormal bases of the same Hilbert space $\mathcal{H}$, then they have the same cardinality.
 38. Prove the last statement in Remark 2.5.

Reference

1. Luxemburg, W.A.J., Zaanen, A.C.: Riesz Spaces, vol. I. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., New York (1971) (North-Holland Mathematical Library)

Chapter 3

Classes of Compact Operators

In traditional Hilbert space quantum mechanics, the spectral representation of a (generally unbounded) selfadjoint operator plays a key role. The development leading to that result needs rather advanced measure theoretic techniques, and these will be introduced in the appropriate context in later chapters. There is, however, an important class of operators for which the spectral representation can be reached without any measure theory: the compact selfadjoint operators. In this chapter the main focus is on the class $\mathcal{C}(\mathcal{H})$ of compact operators and its two subsets, the trace class $\mathcal{T}(\mathcal{H})$ and the Hilbert–Schmidt class $\mathcal{HS}(\mathcal{H})$. This theory is not just an elementary introduction to the general case but is actually central e.g. in the study of the states of a quantum mechanical system. The chapter is concluded with applications to tensor products of Hilbert spaces by introducing the notions of a partial trace and the Schmidt decomposition.

3.1 Compact and Finite Rank Operators

We denote by $\mathcal{F}(\mathcal{H})$ the set of the *finite rank* operators $T \in \mathcal{L}(\mathcal{H})$, i.e. those $T \in \mathcal{L}(\mathcal{H})$ for which $\dim T(\mathcal{H}) < \infty$. We now define another class of operators which are particularly well behaved in many respects.

Definition 3.1 A linear map $T : \mathcal{H} \rightarrow \mathcal{H}$ is a *compact operator* if the closure of the set $\{T\varphi \mid \|\varphi\| \leq 1\}$ is compact. We use the notation $\mathcal{C}(\mathcal{H})$ for the set of the compact operators $T : \mathcal{H} \rightarrow \mathcal{H}$.

Theorem 3.1 For a linear map $T : \mathcal{H} \rightarrow \mathcal{H}$ the following conditions are equivalent:

- (i) $T \in \mathcal{C}(\mathcal{H})$;
- (ii) the sequence $(T\varphi_n)$ has a convergent subsequence whenever (φ_n) is a sequence in $\mathcal{H}$ with $\|\varphi_n\| \leq 1$ for all $n \in \mathbb{N}$;

- (iii) the sequence $(T\varphi_n)$ has a convergent subsequence whenever (φ_n) is a bounded sequence in $\mathcal{H}$;
- (iv) the closure of the image $T(B)$ of every bounded set $B \subset \mathcal{H}$ is compact.

Proof As any sequence in a compact set has a convergent subsequence, (i) implies (ii), and it is clear that (ii) implies (iii). Now assume (iii). Let $B \subset \mathcal{H}$ be a bounded set and (ψ_n) a sequence in the closure of $T(B)$. Choose a sequence (φ_n) in B such that $\|T\varphi_n - \psi_n\| < \frac{1}{n}$ for all $n \in \mathbb{N}$. By (iii) there is a subsequence (φ_{n_k}) such that $(T\varphi_{n_k})$ converges to some ψ ($\in \overline{T(B)}$) so that also $\lim_{k \rightarrow \infty} \psi_{n_k} = \psi$. Thus $\overline{T(B)}$ is compact, and (iv) holds. Clearly (i) follows from (iv). $\square$

We say that a set $\mathcal{I} \subset \mathcal{L}(\mathcal{H})$ is a **-ideal* if $\mathcal{I}$ is a vector subspace of $\mathcal{L}(\mathcal{H})$ and the conditions $S \in \mathcal{I}$, $T \in \mathcal{L}(\mathcal{H})$ imply that $ST \in \mathcal{I}$ and $S^* \in \mathcal{I}$ (so that also $TS = (S^*T^*)^* \in \mathcal{I}$).

Theorem 3.2 *The set $\mathcal{C}(\mathcal{H})$ is a closed *-ideal in $\mathcal{L}(\mathcal{H})$.*

Proof It is clear that $\mathcal{C}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$, since every compact set is bounded. We show that $\mathcal{C}(\mathcal{H})$ is closed with respect to the norm topology of $\mathcal{L}(\mathcal{H})$. Let (T_n) be a sequence in $\mathcal{C}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$ such that $\lim_{n \rightarrow \infty} \|T_n - T\| = 0$. Let (φ_k) be a sequence such that $\|\varphi_k\| \leq 1$ whenever $k \in \mathbb{N}$. For some subsequence $(\varphi_{1,k})$ of (φ_k) the sequence $(T_1\varphi_{1,k})$ converges. Next for some subsequence $(\varphi_{2,k})$ of the sequence $(\varphi_{1,k})$ the sequence $(T_2\varphi_{2,k})$ converges. Continuing in this way we get by induction the sequences $(\varphi_{1,k}), \dots, (\varphi_{n,k}), \dots$, where always $(\varphi_{n+1,k})_{k \in \mathbb{N}}$ is a subsequence of the sequence $(\varphi_{n,k})_{k \in \mathbb{N}}$ and the sequence $(T_n\varphi_{n,k})_{k \in \mathbb{N}}$ converges. From the index n on, the diagonal sequence $(\varphi_{k,k})_{k \in \mathbb{N}}$ is always a subsequence of the sequence $(\varphi_{n,k})_{k \in \mathbb{N}}$, so that the sequence $(T_n\psi_k)_{k \in \mathbb{N}}$ where $\psi_k = \varphi_{k,k}$ converges for all $n \in \mathbb{N}$. Let $\varepsilon > 0$ and take $n \in \mathbb{N}$ such that $\|T_n - T\| < \frac{\varepsilon}{3}$. There is $k_0 \in \mathbb{N}$ such that $\|T_n\psi_k - T_n\psi_p\| < \frac{\varepsilon}{3}$ whenever $k, p \geq k_0$. Then $\|T\psi_k - T\psi_p\| \leq \|T\psi_k - T_n\psi_k\| + \|T_n\psi_k - T_n\psi_p\| + \|T_n\psi_p - T\psi_p\| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$, whenever $k, p \geq k_0$. Thus being a Cauchy sequence, the sequence $(T\psi_k)$ converges. Using Theorem 3.1 we see that $T \in \mathcal{C}(\mathcal{H})$.

If $T_1, T_2 \in \mathcal{C}(\mathcal{H})$, $c_1, c_2 \in \mathbb{C}$ and (φ_k) is a sequence in $\mathcal{H}$ such that $\|\varphi_k\| \leq 1$, then for some subsequence $(\varphi_{1,k})$ the sequence $(T_1\varphi_{1,k})$ converges and for some subsequence $(\varphi_{2,k})$ of the sequence $(\varphi_{1,k})$ the sequence $(T_2\varphi_{2,k})$ converges, and then the sequence $((c_1T_1 + c_2T_2)\varphi_{2,k})$ converges. This shows that $\mathcal{C}(\mathcal{H})$ (obviously nonempty) is a vector subspace of $\mathcal{L}(\mathcal{H})$.

If $S \in \mathcal{C}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$ then $ST \in \mathcal{C}(\mathcal{H})$, since T keeps a bounded sequence bounded. Finally let $S \in \mathcal{C}(\mathcal{H})$ and (φ_n) be as above. Since $SS^* \in \mathcal{C}(\mathcal{H})$, for some subsequence (φ_{n_k}) the sequence $(SS^*\varphi_{n_k})$ converges. As

$$\begin{aligned} \|S^*\varphi_{n_k} - S^*\varphi_{n_p}\|^2 &= \langle \varphi_{n_k} - \varphi_{n_p} | SS^*(\varphi_{n_k} - \varphi_{n_p}) \rangle \\ &\leq \|\varphi_{n_k} - \varphi_{n_p}\| \|SS^*(\varphi_{n_k} - \varphi_{n_p})\| \\ &\leq 2\|SS^*\varphi_{n_k} - SS^*\varphi_{n_p}\| \rightarrow 0, \end{aligned}$$

when $k, p \rightarrow \infty$, the sequence $(S^*\varphi_{n_k})$ is a Cauchy sequence, hence convergent. It follows that $S^* \in \mathcal{C}(\mathcal{H})$. $\square$

We often use the following notation introduced by P.A.M. Dirac: If $\varphi, \psi \in \mathcal{H}$, then $|\varphi\rangle\langle\psi|$ stands for the operator $\xi \mapsto \langle\psi|\xi\rangle\varphi$. In connection with this notation it is customary to talk about the *bra* vectors $\langle\psi|$ belonging to the dual of $\mathcal{H}$ and the *ket* vectors $|\psi\rangle$ belonging to $\mathcal{H}$, and then one may conveniently write $|\varphi\rangle\langle\psi| |\xi\rangle = |\varphi\rangle\langle\psi|\xi\rangle (= \langle\psi|\xi\rangle|\varphi\rangle)$.

Theorem 3.3 (a) A subspace $\mathcal{K} \subset \mathcal{H}$, $\mathcal{K} \neq \{0\}$, is finite-dimensional and $\dim(\mathcal{K}) = n$, if and only if in $\mathcal{K}$ there is a finite orthonormal basis $\{\varphi_1, \dots, \varphi_n\}$. Then $\mathcal{K}$ is closed, and $\sum_{k=1}^n |\varphi_k\rangle\langle\varphi_k|$ is the orthogonal projection from $\mathcal{H}$ onto $\mathcal{K}$.

- (b) For all $T \in \mathcal{L}(\mathcal{H})$ and $\varphi, \psi, \xi, \eta \in \mathcal{H}$, we have $T|\varphi\rangle\langle\psi| = |T\varphi\rangle\langle\psi|$, $|\varphi\rangle\langle\psi|T = |\varphi\rangle\langle T^*\psi|$, $|\varphi\rangle\langle\psi| |\xi\rangle\langle\eta| = \langle\psi|\xi\rangle|\varphi\rangle\langle\eta|$ and $|\varphi\rangle\langle\psi|^* = |\psi\rangle\langle\varphi|$.
- (c) An operator $T \in \mathcal{L}(\mathcal{H})$ has finite rank if and only if it is the sum of a finite number of operators of the form $|\varphi\rangle\langle\psi|$.
- (d) The set $\mathcal{F}(\mathcal{H})$ is a $*$ -ideal of $\mathcal{L}(\mathcal{H})$.
- (e) $\mathcal{F}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$.
- (f) $\mathcal{P}(\mathcal{H}) \cap \mathcal{F}(\mathcal{H}) = \mathcal{P}(\mathcal{H}) \cap \mathcal{C}(\mathcal{H})$.

Proof (a) Let $\mathcal{K}$ be finite-dimensional. As every orthonormal set is linearly independent, there is a finite orthonormal set $\{\varphi_1, \dots, \varphi_n\} \subset \mathcal{K}$ having a maximal number of elements. Then the set $\{\varphi_1, \dots, \varphi_n\}$ is a basis of $\mathcal{K}$ (and hence $\dim(\mathcal{K}) = n$). For it is linearly independent, and if there were some $\varphi \in \mathcal{K}$ which is not a linear combination of the vectors $\varphi_1, \dots, \varphi_n$ then $\psi = \varphi - \sum_{k=1}^n \langle\varphi_k|\varphi\rangle\varphi_k \neq 0$ and $\langle\varphi_k|\psi\rangle = 0$ for all $k = 1, \dots, n$. Then $\{\varphi_1, \dots, \varphi_n, \|\psi\|^{-1}\psi\}$ would be an orthonormal set in $\mathcal{K}$ contradicting the maximality of n . The “if” part is clearly true. By mapping φ_k to the vector $(x_1, \dots, x_n) \in \mathbb{C}^n$ where $x_k = 1$ and $x_i = 0$ for $i \neq k$, $k = 1, \dots, n$, and extending the mapping linearly, it is seen that $\mathcal{K}$ is isometrically isomorphic to $\mathbb{C}^n$. As $\mathbb{C}^n$ is complete, $\mathcal{K}$ is therefore closed. Let P be the projection of $\mathcal{H}$ onto $\mathcal{K}$. If $\psi \in \mathcal{H}$, then for all $k = 1, \dots, n$ we get

$$\left\langle \varphi_k \left| \sum_{j=1}^n \langle\varphi_j|\psi\rangle\varphi_j - P\psi \right. \right\rangle = \langle\varphi_k|\psi\rangle - \langle P\varphi_k|\psi\rangle = 0, \text{ implying } P = \sum_{j=1}^n |\varphi_j\rangle\langle\varphi_j|.$$

(b) The formulas follow by simple calculation. For instance, for all $\xi, \eta \in \mathcal{H}$ we get $\langle(|\varphi\rangle\langle\psi|)^*\xi|\eta\rangle = \langle\xi|(|\varphi\rangle\langle\psi|)\eta\rangle = \langle\psi|\eta\rangle\langle\xi|\varphi\rangle$, and on the other hand $\langle(|\psi\rangle\langle\varphi|)\xi|\eta\rangle = \langle\varphi|\xi\rangle\langle\psi|\eta\rangle = \langle\psi|\eta\rangle\langle\xi|\varphi\rangle$, so that $|\varphi\rangle\langle\psi|^* = |\psi\rangle\langle\varphi|$.

(c) Suppose $T \in \mathcal{F}(\mathcal{H})$ and let P be the projection onto $T(\mathcal{H})$ (see (a)). Since by (a) $P = \sum_{k=1}^n |\varphi_k\rangle\langle\varphi_k|$ for some orthonormal set $\{\varphi_1, \dots, \varphi_n\}$, by (b) we get $T = PT = \sum_{k=1}^n |\varphi_k\rangle\langle T^*\varphi_k|$. The converse is clear, since always $\dim(|\varphi\rangle\langle\psi|(\mathcal{H})) \leq 1$.

(d) This follows easily from (a) and (b).

(e) Every $|\varphi\rangle\langle\psi| \in \mathcal{C}(\mathcal{H})$, since the closure of every bounded set in $\mathbb{C}$ is compact. Hence the claim follows from (a) and Theorem 3.2.

(f) From the above we see that $\mathcal{P}(\mathcal{H}) \cap \mathcal{F}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H}) \cap \mathcal{C}(\mathcal{H})$. On the other hand, let $P \in \mathcal{C}(\mathcal{H})$ be a projection. If $P \notin \mathcal{F}(\mathcal{H})$, then by Theorem 2.21 the subspace

$P(\mathcal{H})$ contains an infinite orthonormal set from which we can choose an (infinite) orthonormal sequence (φ_n) , and since

$$\|P\varphi_n - P\varphi_m\|^2 = \|\varphi_n - \varphi_m\|^2 = \|\varphi_n\|^2 + \|\varphi_m\|^2 = 2,$$

the sequence $(P\varphi_n)$ fails to have a convergent subsequence, contradicting Theorem 3.1. $\square$

3.2 The Spectral Representation of Compact Selfadjoint Operators

We first introduce some general spectral notions. Let $T \in \mathcal{L}(\mathcal{H})$. If $\lambda \in \mathbb{C}$, we use the notation T_λ for the element of $\mathcal{L}(\mathcal{H})$ defined by $T_\lambda\varphi = \lambda\varphi - T\varphi$ for all $\varphi \in \mathcal{H}$. Denote $\mathcal{R}(T_\lambda) = T_\lambda(\mathcal{H})$ ($= \{T_\lambda\varphi \mid \varphi \in \mathcal{H}\}$). We define the sets $\rho(T)$, $\sigma(T)$, $\sigma_p(T)$, $\sigma_c(T)$ and $\sigma_r(T)$ as follows: $\rho(T)$, the *resolvent set* of T , is the set of those $\lambda \in \mathbb{C}$ for which $\mathcal{R}(T_\lambda)$ is dense in $\mathcal{H}$ and T_λ has a continuous inverse (defined on $\mathcal{R}(T_\lambda)$); $\sigma(T)$ is the complement of the resolvent $\rho(T)$ and is called the *spectrum* of T ; $\sigma_p(T)$, the *point spectrum* of T , is the set of those $\lambda \in \mathbb{C}$ for which T_λ is not injective, $\sigma_c(T)$, the *continuous spectrum* of T , is the set of those $\lambda \in \mathbb{C}$ for which $\mathcal{R}(T_\lambda)$ is dense in $\mathcal{H}$ and T_λ is injective but its inverse is discontinuous; $\sigma_r(T)$, the *residual spectrum* of T , is the set of those $\lambda \in \mathbb{C}$ for which $\mathcal{R}(T_\lambda)$ is not dense in $\mathcal{H}$ but T_λ is an injection.

Clearly $\rho(T)$, $\sigma_c(T)$, $\sigma_p(T)$, and $\sigma_r(T)$ are disjoint and their union is all of $\mathbb{C}$. Moreover $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

It is worth observing that if $\lambda \in \rho(T)$, then actually $\mathcal{R}(T_\lambda) = \mathcal{H}$. For if $\psi \in \mathcal{H}$, there is a sequence (φ_n) in $\mathcal{H}$ such that $\lim_{n \rightarrow \infty} T_\lambda\varphi_n = \psi$. Since $T_\lambda^{-1} : \mathcal{R}(T) \rightarrow \mathcal{H}$ is continuous, (φ_n) is a Cauchy sequence and hence converges to some $\varphi \in \mathcal{H}$. But then $\psi = \lim_{n \rightarrow \infty} (T_\lambda\varphi_n) = T_\lambda\varphi \in \mathcal{R}(T)$. The elements of the point spectrum $\sigma_p(T)$ are also said to be the *eigenvalues* of T . For $\lambda \in \mathbb{C}$ to be an eigenvalue of T it is thus necessary and sufficient that there be a nonzero vector $\varphi \in \mathcal{H}$ with $\lambda\varphi - T\varphi = 0$. Such a vector is called an *eigenvector* corresponding to the eigenvalue λ . These eigenvectors and 0 form the *eigenspace* corresponding to the eigenvalue λ . This eigenspace is thus the kernel $\ker(T_\lambda)$ of the operator T_λ , hence closed. The dimension of the eigenspace corresponding to an eigenvalue λ is called the *multiplicity* or *degeneracy* of λ . We usually denote $\ker T_\lambda = N_\lambda$ when T is understood, even if $\ker T_\lambda = \{0\}$.

In the general case any one of the three parts of the spectrum may be void (although the spectrum of a bounded operator is always nonempty). In the case of a compact operator the situation, however, is particularly simple, as the entire nonzero part of the spectrum consists of eigenvalues. We do not need or prove the result in this generality, but we do give the spectral representation theorem of a selfadjoint operator from which this fact can also be deduced in this special case.

The following result contains the heart of the title theorem of this section.

Theorem 3.4 *Let $A \in \mathcal{L}(\mathcal{H})$ be a compact selfadjoint operator and denote $S = \sigma_p(A)$.*

- (a) *Every $\lambda \in S$ is real and satisfies $|\lambda| \leq \|A\|$. If $\lambda \neq \mu$ and $\lambda\varphi - A\varphi = \mu\psi - A\psi = 0$, then $\langle \varphi | \psi \rangle = 0$.*
- (b) *Zero is the only possible accumulation point of S .*
- (c) *If $\lambda \in S \setminus \{0\}$, then the corresponding eigenspace N_λ is finite-dimensional.*
- (d) *If $\mathcal{H} \neq \{0\}$, then $S \neq \emptyset$; in particular, there is such a $\lambda \in S$ that $|\lambda| = \|A\|$.*
- (e) *The eigenspaces corresponding to different eigenvalues are pairwise orthogonal and $\mathcal{H}$ is the smallest closed subspace of $\mathcal{H}$ containing all of them. (Notation: $\mathcal{H} = \bigoplus_{\lambda \in S} N_\lambda$.)*

Proof (a) Every eigenvalue λ of A belongs to $\mathbb{R}$, since $\lambda \langle \varphi | \varphi \rangle = \langle \varphi | \lambda\varphi \rangle = \langle \varphi | A\varphi \rangle \in \mathbb{R}$ for some eigenvector φ . Moreover, if φ is a unit eigenvector, we have $\lambda \langle \varphi | \varphi \rangle = \langle \varphi | A\varphi \rangle$, and so $|\lambda| \leq \|\varphi\| \|A\varphi\| \leq \|A\|$. We next show that any two eigenvectors φ and ψ corresponding to different eigenvalues λ and μ are mutually orthogonal. Since $\lambda \langle \varphi | \psi \rangle = \langle \lambda\varphi | \psi \rangle = \langle A\varphi | \psi \rangle = \langle \varphi | A\psi \rangle = \langle \varphi | \mu\psi \rangle = \mu \langle \varphi | \psi \rangle$ and $\lambda \neq \mu$, we must have $\langle \varphi | \psi \rangle = 0$.

(b), (c) Assume that there is a sequence (λ_n) of complex numbers and an orthonormal sequence (φ_n) in $\mathcal{H}$ satisfying $A\varphi_n = \lambda_n\varphi_n$ for all $n \in \mathbb{N}$ and $\lim_{n \rightarrow \infty} \lambda_n = \lambda \neq 0$. Since $\|\lambda\varphi_n - \lambda\varphi_m\|^2 = 2|\lambda|^2$ whenever $n \neq m$, the sequence $(\lambda\varphi_n)$ cannot have a convergent subsequence. Since $\lambda\varphi_n = \lambda_n\varphi_n + (\lambda - \lambda_n)\varphi_n$ and $\lambda - \lambda_n \rightarrow 0$, neither can the sequence $(\lambda_n\varphi_n)$, i.e. the sequence $(A\varphi_n)$, have a convergent subsequence, which is impossible, since A is a compact operator. This argument first implies (b). For if $\lambda \neq 0$, is an accumulation point of S , by (a) we obtain mutually orthogonal unit vectors corresponding to different eigenvalues whose sequence may be assumed to converge to a nonzero limit. To prove (c), note that if N_λ is assumed to be infinite-dimensional for a nonzero λ , we get a contradiction by using a constant sequence (λ_n) with $\lambda_n = \lambda$ and an infinite orthonormal sequence in N_λ .

We next prove (d). We may assume that $A \neq 0$. Using Proposition 2.2 we find a sequence (ψ_n) of unit vectors such that $\lim_{n \rightarrow \infty} |\langle \psi_n | A\psi_n \rangle| = \|A\|$. If necessary, by replacing the sequence (ψ_n) with a suitable subsequence we may assume that $\lim_{n \rightarrow \infty} \langle \psi_n | A\psi_n \rangle = \lambda$ where $\lambda = \|A\|$ or $\lambda = -\|A\|$. Assume that $\lambda = \|A\|$. Then

$$\begin{aligned} 0 &\leq \|A\psi_n - \lambda\psi_n\|^2 = \|A\psi_n\|^2 - 2\lambda \langle \psi_n | A\psi_n \rangle + \lambda^2 \\ &\leq 2\lambda(\lambda - \langle \psi_n | A\psi_n \rangle) \rightarrow 0, \end{aligned}$$

when $n \rightarrow \infty$. As A is a compact operator, the sequence $(A\psi_n)$ has a convergent subsequence $(A\psi_{n_k})$. Since $\lim_{n \rightarrow \infty} (A\psi_n - \lambda\psi_n) = 0$ and $\lambda \neq 0$, the sequence (ψ_{n_k}) converges to some vector ψ_0 . Then $\|\psi_0\| = 1$ and $A\psi_0 - \lambda\psi_0 = \lim_{k \rightarrow \infty} (A\psi_{n_k} - \lambda\psi_{n_k}) = 0$. Thus $\lambda = \|A\|$ is an eigenvalue of the operator A . The case $\lambda = -\|A\|$ is treated analogously.

(e) The first claim is contained in (a). Let $\mathcal{K}$ be the closed subspace of $\mathcal{H}$ generated by $\bigcup_{\lambda \in S} N_\lambda$. We claim that $\mathcal{K} = \mathcal{H}$. Assume, on the contrary, that $\mathcal{K} \neq \mathcal{H}$. Clearly $A(\mathcal{K}) \subset \mathcal{K}$. If $\varphi \in \mathcal{K}^\perp$ and $\psi \in \mathcal{K}$, then $\langle \psi | A\varphi \rangle = \langle A\psi | \varphi \rangle = 0$, which shows that also $A(\mathcal{K}^\perp) \subset \mathcal{K}^\perp$. We now interpret the restriction $A|_{\mathcal{K}^\perp}$ as a mapping into $\mathcal{K}^\perp$ ($\neq \{0\}$). By (c) it has a nonzero eigenvalue, which at the same time is an eigenvalue of A not belonging to S . This contradiction completes the proof of (e). $\square$

We now get the *spectral representation* of a compact selfadjoint operator.

Theorem 3.5 *Let $A \in \mathcal{C}(\mathcal{H})$ be selfadjoint. Then A can be represented as the sum of a norm convergent series or as a finite (or in case $A = 0$ as an empty) sum*

$$A = \sum_n \lambda_n P_n, \quad (3.1)$$

where

- (i) $\lambda_n \neq 0$ for all n ,
- (ii) P_n is a projection, $P_n \neq 0$ and $P_n \in \mathcal{F}(\mathcal{H})$ for all n ,
- (iii) $P_n P_m = 0$ if $n \neq m$,
- (iv) $\lambda_n \neq \lambda_m$ if $n \neq m$,
- (v) the set of the numbers λ_n appearing in (3.1) is $\sigma_p(A) \setminus \{0\}$, and for any n , P_n is the projection onto the eigenspace corresponding to λ_n . Independently of the order the sum in (3.1) is the same, and $\lim_{n \rightarrow \infty} \lambda_n = 0$ if the set $\sigma_p(A) \setminus \{0\}$ is infinite. In particular, $\sigma_p(A) \setminus \{0\}$ is at most countable and zero is its only possible accumulation point.

Proof From Theorem 3.4 it follows that outside any neighbourhood $B(0, \frac{1}{k})$, $k \in \mathbb{N}$, of the origin there are at most a finite number of the eigenvalues of A . Thus the set S' of the nonzero eigenvalues is at most countable, and if S' is infinite, its elements can be arranged into a sequence converging to zero. Let (λ_n) be an arbitrary enumeration of the elements of S' into a sequence. Assuming that S' is infinite we have $\lim_{n \rightarrow \infty} \lambda_n = 0$. (The case of finitely many eigenvalues is simpler and only requires minor changes.) Let P_n be the projection onto the (by Theorem 3.4) finite-dimensional eigenspace corresponding to λ_n . We show that the series $\sum_{n=1}^{\infty} \lambda_n P_n$ converges with respect to the operator norm to the operator A . If $m > n$ and $\varphi \in \mathcal{H}$, $\|\varphi\| \leq 1$, then using Theorem 3.4 (d) (the mutual orthogonality of eigenspaces) and the Pythagorean theorem we get

$$\begin{aligned} \left\| \left(\sum_{k=1}^m \lambda_k P_k - \sum_{k=1}^n \lambda_k P_k \right) \varphi \right\|^2 &= \left\| \left(\sum_{k=n+1}^m \lambda_k P_k \right) \varphi \right\|^2 = \sum_{k=n+1}^m |\lambda_k|^2 \|P_k \varphi\|^2 \\ &\leq \max_{n+1 \leq j \leq m} |\lambda_j|^2 \sum_{k=n+1}^m \|P_k \varphi\|^2 \leq \max_{n+1 \leq k \leq m} |\lambda_k|^2, \end{aligned}$$

which implies that

$$\left\| \sum_{k=1}^m \lambda_k P_k - \sum_{k=1}^n \lambda_k P_k \right\|^2 \leq \max_{n+1 \leq k \leq m} |\lambda_k|^2.$$

Since $\lim_{k \rightarrow \infty} \lambda_k = 0$, we see that the sequence of partial sums $(\sum_{k=1}^n \lambda_k P_k)$ is a Cauchy sequence in the operator norm and hence converges in norm to some $B \in \mathcal{L}(\mathcal{H})$. In view of Theorem 3.4, it is now enough to prove that the restrictions $A|_{N_\lambda}$ and $B|_{N_\lambda}$ are the same for every eigenvalue λ of A . For nonzero eigenvalues this is clear since $B P_n = \lambda_n P_n$ so that $A\varphi = \lambda_n \varphi = B\varphi$ if $\varphi \in N_{\lambda_n}$. On the other hand, if $A\varphi = 0$ for some φ , then by Theorem 3.4 $\varphi \perp N_{\lambda_n}$ for every λ_n in the representation of B , and so $B\varphi = 0$. $\square$

Corollary 3.1 *The set $\mathcal{C}(\mathcal{H})$ is the closure of $\mathcal{F}(\mathcal{H})$ in $\mathcal{L}(\mathcal{H})$.*

Proof Since $\mathcal{C}(\mathcal{H})$ is closed (Theorem 3.2) and $\mathcal{F}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$ (Theorem 3.3 (d)), we have $\overline{\mathcal{F}(\mathcal{H})} \subset \mathcal{C}(\mathcal{H})$. Now let $T \in \mathcal{C}(\mathcal{H})$. Then $T = A + iB$ where $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$ are selfadjoint compact operators (Theorem 3.2). Applying Theorem 3.5 we see that $A, B \in \overline{\mathcal{F}(\mathcal{H})}$, and so $T \in \overline{\mathcal{F}(\mathcal{H})}$. $\square$

Remark 3.1 (a) Using Theorem 3.3 (a) it is seen that in the situation of the preceding theorem we may write

$$A = \sum_n c_n |\varphi_n\rangle \langle \varphi_n|, \quad (3.1')$$

where (φ_n) is a (finite or infinite) orthogonal sequence (and in the case of an infinite sequence the series converges in norm) and the set of the numbers c_n is $\sigma_p(A) \setminus \{0\}$. Each c_n is an eigenvalue of A appearing in the sequence as many times as indicated by its multiplicity, the dimension of the corresponding eigenspace. The representation (3.1) in Theorem 3.5 is clearly unique apart from order, but in (3.1') the vectors φ_n are not uniquely determined. Yet the vectors φ_n with the same value of c_n necessarily form an orthonormal basis of the corresponding eigenspace. If $\dim(\mathcal{H}) = \infty$, by Theorem 2.21 the sequence (φ_n) can be extended to an infinite orthonormal sequence to obtain

$$A = \sum_{n=1}^{\infty} c_n |\varphi_n\rangle \langle \varphi_n|, \quad (3.1'')$$

where it is now possible to have $c_n = 0$.

(b) In the case of a general $T \in \mathcal{C}(\mathcal{H})$ we may use the polar decomposition $T = V|T|$ where $|T| = V^*T \in \mathcal{C}(\mathcal{H})$ (see Theorems 2.18, 3.2). If

$$|T| = \sum_n c_n |\varphi_n\rangle \langle \varphi_n|$$

is the representation corresponding to (3.1'), so that now $c_n > 0$, according to Theorem 3.3 (b)

$$T = \sum_n c_n |V\varphi_n\rangle\langle\varphi_n|,$$

where $(V\varphi_n)_n$ is an orthonormal family, for V preserves the norms and inner products for the vectors of the subspace $\text{supp}(|T|)(\mathcal{H}) = \text{supp}(T)(\mathcal{H})$ (see Theorem 2.18 and its proof). Again using Theorem 2.21 we may write (when $\dim(\mathcal{H}) = \infty$)

$$T = \sum_{n=1}^{\infty} c_n |\psi_n\rangle\langle\varphi_n|, \quad (3.2)$$

where $c_n \geq 0$, (ψ_n) and (φ_n) are orthonormal sequences. We may observe, conversely, that if (ψ_n) and (φ_n) are orthonormal sequences and (c_n) is a sequence with $c_n \geq 0$ and $\lim_{n \rightarrow \infty} c_n = 0$, then

$$\begin{aligned} & \left\| \left(\sum_{n=1}^k c_n |\psi_n\rangle\langle\varphi_n| - \sum_{n=1}^p c_n |\psi_n\rangle\langle\varphi_n| \right) \xi \right\|^2 \\ &= \sum_{n=k+1}^p |c_n \langle\varphi_n|\xi\rangle|^2 \leq \max_{k+1 \leq n \leq p} c_n^2 \|\xi\|^2, \end{aligned}$$

whenever $p > k$, and so the partial sums of the series (3.2) form a Cauchy sequence in $\mathcal{L}(\mathcal{H})$, and so (3.2) determines a compact operator (Theorem 3.2). Moreover, Theorem 3.3 (b) shows that $T^*T = \sum_{n=1}^{\infty} c_n^2 |\varphi_n\rangle\langle\varphi_n|$, and it is easy to see from the uniqueness of the positive square root that

$$|T| = \sum_{n=1}^{\infty} c_n |\varphi_n\rangle\langle\varphi_n|.$$

◁

Remark 3.2 Suppose that $T \in \mathcal{C}(\mathcal{H})$ is nonzero and consider a representation $T = \sum_n c_n |\psi_n\rangle\langle\varphi_n|$ as in Remark 3.1 (b) with $c_n > 0$. Then $|T| = \sum_n c_n |\varphi_n\rangle\langle\varphi_n|$. We assume that $c_1 \geq c_2 \geq \dots$, and then the uniqueness of the spectral representation implies that the coefficients c_n are uniquely determined and for each c_n the linear span of the set $\{\varphi_m \mid c_m = c_n\}$ is unique though the vectors φ_n are not. Since $T\varphi_n = c_n\psi_n$, so that T maps isomorphically the linear span of the orthogonal set $\{\varphi_m \mid c_m = c_n\}$ onto that of the orthogonal set $\{\psi_m \mid c_m = c_n\}$, we see that similar uniqueness statements apply to the vectors ψ_n . In the case of a *nondegenerate eigenvalue* c_n of $|T|$, which by definition occurs only once in the sum for $|T|$, the vectors φ_n are unique up to multiplication by a scalar $\alpha_n \in \mathbb{C}$ with $|\alpha_n| = 1$. ◁

3.3 The Hilbert–Schmidt Operator Class $\mathcal{HS}(\mathcal{H})$

In this section the main object of study is a certain $*$ -ideal in $\mathcal{L}(\mathcal{H})$ having a natural Hilbert space structure. We begin with a calculation on orthonormal bases.

Lemma 3.1 *Let K and L be orthonormal bases in the Hilbert space $\mathcal{H}$. If $T \in \mathcal{L}(\mathcal{H})$, then*

- (a) $\sum_{\xi \in K} \langle \xi | T^* T \xi \rangle = \sum_{\eta \in L} \langle \eta | T T^* \eta \rangle$,
- (b) $\sum_{\xi \in K} \|T\xi\|^2 = \sum_{\eta \in L} \|T\eta\|^2 = \sum_{\eta \in L} \|T^* \eta\|^2$ and
- (c) $\|T\|^2 \leq \sum_{\xi \in K} \|T\xi\|^2 \quad (\leq \infty)$.

Proof (a) Using the Parseval identity in Theorem 2.20 and the fact that the order of summation in double series with positive terms is irrelevant, we get

$$\begin{aligned} \sum_{\xi \in K} \langle \xi | T^* T \xi \rangle &= \sum_{\xi \in K} \|T\xi\|^2 = \sum_{\xi \in K} \sum_{\eta \in L} |\langle \eta | T\xi \rangle|^2 \\ &= \sum_{\eta \in L} \sum_{\xi \in K} |\langle T^* \eta | \xi \rangle|^2 = \sum_{\eta \in L} \|T^* \eta\|^2 = \sum_{\eta \in L} \langle \eta | T T^* \eta \rangle. \end{aligned}$$

(b) The extreme ends were already shown to coincide. With the choice $K = L$ in (a) we obtain the latter equality.

(c) If $\|\varphi\| = 1$ then the set $\{\varphi\}$ can by Theorem 2.21 be extended to an orthonormal basis K' , and then (b) implies $\|T\varphi\|^2 \leq \sum_{\eta \in K'} \|T\eta\|^2 = \sum_{\xi \in K} \|T\xi\|^2$. Taking the supremum with respect to φ we obtain the claim. $\square$

In this section K is always an orthonormal basis of $\mathcal{H}$. The above lemma shows that independently of the choice of K we can make the following definition:

Definition 3.2 If $T \in \mathcal{L}(\mathcal{H})$ and $\sum_{\xi \in K} \|T\xi\|^2 < \infty$ we say that T is a *Hilbert–Schmidt operator*. We use the notation $\mathcal{HS}(\mathcal{H})$ for the set of such operators. We denote $\|T\|_2 = \left(\sum_{\xi \in K} \|T\xi\|^2 \right)^{\frac{1}{2}}$ whenever $T \in \mathcal{HS}(\mathcal{H})$.

Theorem 3.6 *For all $S, T \in \mathcal{HS}(\mathcal{H})$ we may define*

$$\langle S | T \rangle = \sum_{\xi \in K} \langle S\xi | T\xi \rangle \tag{3.3}$$

as the sum of a summable family independently of the choice of K . The set $\mathcal{HS}(\mathcal{H})$ is a vector subspace of $\mathcal{L}(\mathcal{H})$, and the mapping $\langle \cdot | \cdot \rangle : \mathcal{HS}(\mathcal{H}) \times \mathcal{HS}(\mathcal{H}) \rightarrow \mathbb{C}$ defined in (3.3) is an inner product with respect to which $\mathcal{HS}(\mathcal{H})$ is a Hilbert space.

Proof We first show that $\mathcal{HS}(\mathcal{H})$ is a vector subspace. If $S, T \in \mathcal{HS}(\mathcal{H})$ then clearly $cS \in \mathcal{HS}(\mathcal{H})$ for all $c \in \mathbb{C}$, and since

$$\begin{aligned} \|(S + T)\xi\|^2 &= \|S\xi\|^2 + \langle S\xi|T\xi \rangle + \langle T\xi|S\xi \rangle + \|T\xi\|^2 \\ &\leq \|S\xi\|^2 + 2\|S\xi\| \|T\xi\| + \|T\xi\|^2 \leq 2\|S\xi\|^2 + 2\|T\xi\|^2 \end{aligned}$$

for all $\xi \in K$, we have $S + T \in \mathcal{HS}(\mathcal{H})$. For every finite $F \subset K$ we may estimate using the Cauchy–Schwarz inequality

$$\sum_{\xi \in F} |\langle S\xi|T\xi \rangle| \leq \left(\sum_{\xi \in F} \|S\xi\|^2 \right)^{\frac{1}{2}} \left(\sum_{\xi \in F} \|T\xi\|^2 \right)^{\frac{1}{2}} \leq \|S\|_2 \|T\|_2,$$

and it follows that $\langle S|T \rangle$ is defined as the sum of a summable family whenever $S, T \in \mathcal{HS}(\mathcal{H})$. It is easily seen that in this way we get an inner product for the space $\mathcal{HS}(\mathcal{H})$. For example, if $\langle T|T \rangle = 0$, then Lemma 3.1 (c) shows that $T = 0$. As $\langle T|T \rangle = \|T\|_2^2$ is by Lemma 3.1 (b) independent of the choice of K , using the polarisation identity it is seen that $\langle S|T \rangle$ is independent of the choice of K . We still show the completeness of the space $\mathcal{HS}(\mathcal{H})$. Let (T_n) be a Cauchy sequence in it with respect to the norm $\|\cdot\|_2$, so that in view on Lemma 3.1 (c) it is a Cauchy sequence also with respect to the ordinary operator norm of $\mathcal{L}(\mathcal{H})$. Hence there is a $T \in \mathcal{L}(\mathcal{H})$ satisfying $\lim_{n \rightarrow \infty} \|T_n - T\| = 0$. Let $\varepsilon > 0$. Choose a $p \in \mathbb{N}$ such that $\|T_n - T_m\|_2 < \varepsilon$ whenever $n, m \geq p$. Let $\{\xi_1, \dots, \xi_k\}$ be an arbitrary finite subset of K and $n \geq p$. Then

$$\sum_{j=1}^k \|(T_n - T)\xi_j\|^2 = \lim_{m \rightarrow \infty} \sum_{j=1}^k \|(T_n - T_m)\xi_j\|^2 \leq \varepsilon^2.$$

It follows that $T_n - T \in \mathcal{HS}(\mathcal{H})$, hence $T = T_n - (T_n - T) \in \mathcal{HS}(\mathcal{H})$ and $\|T_n - T\|_2 \leq \varepsilon$. Therefore $\lim_{n \rightarrow \infty} \|T_n - T\|_2 = 0$. $\square$

We still prove a theorem showing that $\mathcal{HS}(\mathcal{H})$ is a $*$ -ideal of $\mathcal{L}(\mathcal{H})$.

Theorem 3.7 (a) If $T \in \mathcal{HS}(\mathcal{H})$ then $T^* \in \mathcal{HS}(\mathcal{H})$ and $\|T^*\|_2 = \|T\|_2 \geq \|T\|$.
 (b) If $S \in \mathcal{HS}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$, then $ST, TS \in \mathcal{HS}(\mathcal{H})$ and

$$\begin{aligned} \|ST\|_2 &\leq \|S\|_2 \|T\| \text{ and} \\ \|TS\|_2 &\leq \|T\| \|S\|_2. \end{aligned}$$

- (c) A projection P is in $\mathcal{HS}(\mathcal{H})$ if and only if $P(\mathcal{H})$ is finite-dimensional, and then $\|P\|_2^2 = \dim(P(\mathcal{H}))$.
 (d) $\mathcal{F}(\mathcal{H}) \subset \mathcal{HS}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$.

Proof Item (a) follows from Lemma 3.1 (b), (c). Moreover,

$$\sum_{\xi \in K} \|TS\xi\|^2 \leq \|T\|^2 \sum_{\xi \in K} \|S\xi\|^2 = \|T\|^2 \|S\|_2^2,$$

so that $TS \in \mathcal{HS}(\mathcal{H})$ and $\|TS\|_2 \leq \|T\| \|S\|_2$. By (a) we have $ST = (T^*S^*)^* \in \mathcal{HS}(\mathcal{H})$ and $\|ST\|_2 = \|T^*S^*\|_2 \leq \|T^*\| \|S^*\|_2 = \|T\| \|S\|_2$. Let now $P \in \mathcal{L}(\mathcal{H})$ be a projection. Choose for the Hilbert space $P(\mathcal{H})$ an orthonormal basis L_0 and extend it to an orthonormal basis L of $\mathcal{H}$ (see Theorem 2.21). Since $P(L \setminus L_0) \subset \{0\}$, we see that $\sum_{\xi \in L} \|P\xi\|^2 = \sum_{\xi \in L_0} \|\xi\|^2$ immediately implying (c). (d) Every $|\varphi\rangle\langle\psi|$ is in $\mathcal{HS}(\mathcal{H})$, since $\sum_{\xi \in K} \|\langle\psi|\xi\rangle\varphi\|^2 \leq \|\psi\|^2 \|\varphi\|^2$ (Lemma 2.2 (b)). Hence by Theorems 3.3 (a) and 3.6 we have $\mathcal{F}(\mathcal{H}) \subset \mathcal{HS}(\mathcal{H})$. Now let $T \in \mathcal{HS}(\mathcal{H})$. Since $\sum_{\xi \in K} \|T\xi\|^2 < \infty$, we have $T\xi = 0$ for each $\xi \in K$ with the possible exception of a countable set of vectors $\xi_1, \xi_2, \dots \in K$. Let P_n be the projection onto the (by Theorem 3.3 (a) closed) subspace spanned by $\{\xi_1, \dots, \xi_n\}$. Then by Lemma 3.1 (c)

$$\|T - TP_n\|^2 \leq \|T - TP_n\|_2^2 = \sum_{\substack{\xi \in \\ K \setminus \{\xi_1, \dots, \xi_n\}}} \|T(I - P_n)\xi\|^2 = \sum_{j=n+1}^{\infty} \|T\xi_j\|^2 \rightarrow 0,$$

when $n \rightarrow \infty$. Since always $TP_n \in \mathcal{F}(\mathcal{H})$, we have $T \in \mathcal{C}(\mathcal{H})$ by Theorem 3.2. $\square$

3.4 The Trace Class $\mathcal{T}(\mathcal{H})$

In linear algebra the trace of a matrix is defined as the sum of its main diagonal elements. Even in the case of infinite-dimensional Hilbert spaces the notion of a trace can be extended to certain operators, those belonging to the so-called trace class which we now define. As in the previous section, in this section K is an arbitrary orthonormal basis of $\mathcal{H}$. In the next definition the expression $\sum_{\xi \in K} \langle \xi | T \xi \rangle$ is in view of Lemma 3.1 independent of the choice of K , since $T = S^*S = SS^*$, where $S = T^{\frac{1}{2}}$.

Definition 3.3 For each $T \in \mathcal{L}_s(\mathcal{H})^+$ we denote

$$\mathrm{tr}[T] = \sum_{\xi \in K} \langle \xi | T \xi \rangle \quad (\leq \infty). \quad (3.4)$$

We write $\mathcal{T}(\mathcal{H}) = \{T \in \mathcal{L}(\mathcal{H}) | \mathrm{tr}[|T|] < \infty\}$ and call the set $\mathcal{T}(\mathcal{H})$ the *trace class*.

The trace class has a close connection with the Hilbert–Schmidt operators:

Lemma 3.2 (a) $\mathcal{T}(\mathcal{H}) = \{ST | S, T \in \mathcal{HS}(\mathcal{H})\}$.

(b) $\mathcal{T}(\mathcal{H}) \subset \mathcal{HS}(\mathcal{H})$.

Proof (a) Assume first that $S, T \in \mathcal{HS}(\mathcal{H})$. Let $ST = U|ST|$ be the polar decomposition of ST . Then

$$\sum_{\xi \in K} \langle \xi | |ST| \xi \rangle = \sum_{\xi \in K} \langle \xi | U^*ST\xi \rangle = \sum_{\xi \in K} \langle (U^*S)^*\xi | T\xi \rangle \leq \|U^*S\|_2 \|T\|_2 < \infty$$

by the Cauchy–Schwarz inequality in $\mathcal{HS}(\mathcal{H})$, so that $ST \in \mathcal{T}(\mathcal{H})$. Next suppose that $V \in \mathcal{T}(\mathcal{H})$ and let $V = W|V|$ be its polar decomposition. Then $V = W|V|^{\frac{1}{2}}|V|^{\frac{1}{2}}$, where by definition $|V|^{\frac{1}{2}} \in \mathcal{HS}(\mathcal{H})$, and according to Theorem 3.7 (b) also $W|V|^{\frac{1}{2}} \in \mathcal{HS}(\mathcal{H})$.

(b) This follows from (a) and Theorem 3.7 (b). $\square$

Before defining the trace we prove another auxiliary result.

Lemma 3.3 *If $T \in \mathcal{T}(\mathcal{H})$ and K, L are orthonormal bases in $\mathcal{H}$, then the families $(\langle \xi | T \xi \rangle)_{\xi \in K}$ and $(\langle \eta | T \eta \rangle)_{\eta \in L}$ are summable and*

$$\sum_{\xi \in K} \langle \xi | T \xi \rangle = \sum_{\eta \in L} \langle \eta | T \eta \rangle. \quad (3.5)$$

Proof By Lemma 3.2 $T = AB$ where $A, B \in \mathcal{HS}(\mathcal{H})$. Since $\langle \xi | T \xi \rangle = \langle A^* \xi | B \xi \rangle$, the claim follows from Theorem 3.6, for also $A^* \in \mathcal{HS}(\mathcal{H})$ by Theorem 3.7. $\square$

Definition 3.4 If $T \in \mathcal{T}(\mathcal{H})$ the number $\text{tr}[T] = \sum_{\xi \in K} \langle \xi | T \xi \rangle$ (which according to the preceding lemma is independent of the choice of K) is called the *trace* of T .

Theorem 3.8 *The trace class $\mathcal{T}(\mathcal{H})$ is a $*$ -ideal of $\mathcal{L}(\mathcal{H})$, and the trace $\text{tr}[\cdot]$ is a linear functional on it. The trace is unitarily invariant, i.e.*

$$\text{tr}[T] = \text{tr}[U^* T U] \quad (3.6)$$

whenever $T \in \mathcal{T}(\mathcal{H})$ and $U \in \mathcal{L}(\mathcal{H})$ is unitary. Moreover,

$$\text{tr}[ST] = \text{tr}[TS] \quad (3.7)$$

whenever $S \in \mathcal{L}(\mathcal{H})$ and $T \in \mathcal{T}(\mathcal{H})$.

Proof If $T = AB$ with $A, B \in \mathcal{HS}(\mathcal{H})$ (see Lemma 3.2), then $T^* = B^* A^* \in \mathcal{T}(\mathcal{H})$ by Theorem 3.7 (a) and Lemma 3.2. From Lemma 3.2 and Theorem 3.7 (b) it also follows that $\mathcal{T}(\mathcal{H})$ is closed with respect to multiplication by elements of $\mathcal{L}(\mathcal{H})$. If $S, T \in \mathcal{T}(\mathcal{H})$ and $T + S = U|S + T|$ is the polar decomposition of the sum $S + T$. Then by Theorem 2.18 $|S + T| = U^*(S + T)$, and since $U^* S, U^* T \in \mathcal{T}(\mathcal{H})$, we have

$$\sum_{\xi \in K} \langle \xi | |S + T| \xi \rangle \leq \sum_{\xi \in K} |\langle \xi | U^* S \xi \rangle| + \sum_{\xi \in K} |\langle \xi | U^* T \xi \rangle| < \infty.$$

Hence $\mathcal{T}(\mathcal{H})$ is a $*$ -ideal. From Definition 3.4 one sees immediately the linearity of the trace. If $U \in \mathcal{L}(\mathcal{H})$ is unitary, then $\{U\xi \mid \xi \in K\}$ is an orthonormal basis of $\mathcal{H}$, and so by Lemma 3.3

$$\text{tr}[U^* T U] = \sum_{\xi \in K} \langle U\xi | T U\xi \rangle = \text{tr}[T] \quad \text{when } T \in \mathcal{T}(\mathcal{H}).$$

Thus (3.6) holds and $\operatorname{tr}[UT] = \operatorname{tr}[U^*UTU] = \operatorname{tr}[TU]$. From this (3.7) follows, since every $S \in \mathcal{L}(\mathcal{H})$ is a linear combination of four unitary operators by Proposition 2.6. $\square$

Also $\mathcal{T}(\mathcal{H})$ turns out to be a Banach space with respect to a suitable norm (which is in general greater than the ordinary operator norm). We next define this norm and prove some of its basic properties. The completeness will be proved later.

Definition 3.5 For each $T \in \mathcal{T}(\mathcal{H})$ denote

$$\|T\|_1 = \operatorname{tr}[|T|] \quad (= \| |T|^{\frac{1}{2}} \|_2^2).$$

Theorem 3.9 The mapping $T \mapsto \|T\|_1$ is a norm in $\mathcal{T}(\mathcal{H})$, and we have

$$|\operatorname{tr}[ST]| \leq \|S\| \|T\|_1 \quad \text{for all } S \in \mathcal{L}(\mathcal{H}), T \in \mathcal{T}(\mathcal{H}). \quad (3.8)$$

Proof Let $T, T_0 \in \mathcal{T}(\mathcal{H})$. We first prove (3.8). If $T = U|T|$ is the polar decomposition of T and $S \in \mathcal{L}(\mathcal{H})$, we have

$$\begin{aligned} |\operatorname{tr}[ST]| &= |\operatorname{tr}[SU|T|]| = |\operatorname{tr}[SU|T|^{\frac{1}{2}}|T|^{\frac{1}{2}}]| = \left| \sum_{\xi \in K} \langle |T|^{\frac{1}{2}} U^* S^* \xi | |T|^{\frac{1}{2}} \xi \rangle \right| \\ &= |\langle |T|^{\frac{1}{2}} U^* S^* | |T|^{\frac{1}{2}} \rangle| \leq \| |T|^{\frac{1}{2}} U^* S^* \|_2 \| |T|^{\frac{1}{2}} \|_2 \\ &\leq \| |T|^{\frac{1}{2}} \|_2 \|U^*\| \|S^*\| \| |T|^{\frac{1}{2}} \|_2 = \|S\| \|T\|_1 \end{aligned}$$

by using the Cauchy–Schwarz inequality and Theorem 3.7 (b). If $\lambda \in \mathbb{C}$, then $\|\lambda T\|_1 = \| |\lambda T|^{\frac{1}{2}} \|_2^2 = |\lambda| \|T\|_1$. If $\|T\|_1 = \| |T|^{\frac{1}{2}} \|_2^2 = 0$, then $|T| = 0$, so that $\|T\|^2 = \|T^*T\| = 0$, i.e. $T = 0$. Let now $T + T_0 = U|T + T_0|$ be the polar decomposition of the sum $T + T_0$. Then by Theorem 2.18 we get

$$\begin{aligned} \|T + T_0\|_1 &= \operatorname{tr}[|T + T_0|] = \operatorname{tr}[U^*T + U^*T_0] = \operatorname{tr}[U^*T] + \operatorname{tr}[U^*T_0] \\ &\leq |\operatorname{tr}[U^*T]| + |\operatorname{tr}[U^*T_0]| \leq \|U^*\| \|T\|_1 + \|U^*\| \|T_0\|_1 \\ &\leq \|T\|_1 + \|T_0\|_1. \end{aligned}$$

Thus the mapping $T \mapsto \|T\|_1$ is a norm. $\square$

3.5 Connection of the Ideals $\mathcal{T}(\mathcal{H})$ and $\mathcal{HS}(\mathcal{H})$ with the Sequence Spaces ℓ^1 and ℓ^2

We know that $\mathcal{T}(\mathcal{H}) \subset \mathcal{HS}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$ (Lemma 3.2, Theorem 3.7). In view of Remark 3.1 the compact operators, on the other hand are precisely those having the representation

$$T = \sum_{n=1}^{\infty} c_n |\psi_n\rangle\langle\varphi_n| \quad (3.9)$$

where $c_n \geq 0$, $\lim_{n \rightarrow \infty} c_n = 0$ and (φ_n) and (ψ_n) are orthonormal sequences in $\mathcal{H}$. (We assume in this and the next section that $\mathcal{H}$ is infinite-dimensional; the simpler finite-dimensional case would only require minor changes in notation.) We denote as usual

$$\ell^1 = \left\{ (c_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |c_n| < \infty \right\} \text{ and } \ell^2 = \left\{ (c_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |c_n|^2 < \infty \right\}.$$

The following theorem explains the connections mentioned in the title of this section.

Theorem 3.10 *In the situation of formula (3.9) we have*

- (a) $T \in \mathcal{HS}(\mathcal{H})$, if and only if $(c_n) \in \ell^2$, and then $\|T\|_2 = \left(\sum_{n=1}^{\infty} c_n^2 \right)^{\frac{1}{2}}$;
- (b) $T \in \mathcal{T}(\mathcal{H})$, if and only if $(c_n) \in \ell^1$, and then $\|T\|_1 = \sum_{n=1}^{\infty} c_n$ and $\text{tr}[T] = \sum_{n=1}^{\infty} c_n \langle \varphi_n | \psi_n \rangle$.

Proof First, suppose that $T \in \mathcal{HS}(\mathcal{H})$. Choose for $\mathcal{H}$ an orthonormal basis K containing the set $\{\varphi_n | n \in \mathbb{N}\}$ (see Theorem 2.21). Then

$$\sum_{\xi \in K} \|T\xi\|^2 = \sum_{n=1}^{\infty} \|c_n \psi_n\|^2 = \sum_{n=1}^{\infty} c_n^2, \text{ for } T\xi = \sum_{m=1}^{\infty} c_m \langle \varphi_m | \xi \rangle \psi_m = c_n \psi_n,$$

if $\xi = \varphi_n$, and $T\xi = 0$, if $\xi \in K \setminus \{\varphi_n\}$. This implies (a). Since

$$|T|^{\frac{1}{2}} = \sum_{n=1}^{\infty} c_n^{\frac{1}{2}} |\varphi_n\rangle\langle\varphi_n|$$

(see Remark 3.1), from the above we see that $\sum_{\xi \in K} \langle \xi | |T| \xi \rangle = \sum_{n=1}^{\infty} c_n$, implying the first claim in (b) and the equality $\|T\|_1 = \sum_{n=1}^{\infty} c_n$. If $T \in \mathcal{T}(\mathcal{H})$, we further get

$$\text{tr}[T] = \sum_{\xi \in K} \langle \xi | T \xi \rangle = \sum_{n=1}^{\infty} \langle \varphi_n | c_n \psi_n \rangle = \sum_{n=1}^{\infty} c_n \langle \varphi_n | \psi_n \rangle. \quad \square$$

We now list some immediate consequences of the preceding theorem.

Theorem 3.11 (a) *If $\varphi, \psi \in \mathcal{H}$ then $|\varphi\rangle\langle\psi| \in \mathcal{T}(\mathcal{H}) \subset \mathcal{HS}(\mathcal{H})$ and*

$$\| |\varphi\rangle\langle\psi| \|_1 = \| |\varphi\rangle\langle\psi| \|_2 = \|\varphi\| \|\psi\|.$$

- (b) *If $T \in \mathcal{F}(\mathcal{H})$ then $T \in \mathcal{T}(\mathcal{H})$. If $T = \sum_{i=1}^n |\varphi_i\rangle\langle\psi_i|$ then $\text{tr}[T] = \sum_{i=1}^n \langle\psi_i|\varphi_i\rangle$.*
- (c) *If $P \in \mathcal{L}(\mathcal{H})$ is a projection then $P \in \mathcal{T}(\mathcal{H})$ if and only if $P \in \mathcal{F}(\mathcal{H})$, and then $\text{tr}[P] = \dim P(\mathcal{H})$.*
- (d) *If $T \in \mathcal{T}(\mathcal{H})$, then the series $\sum_{n=1}^{\infty} c_n |\psi_n\rangle\langle\varphi_n|$ in (3.9) converges with respect to the norm $\|\cdot\|_1$, and if $T \in \mathcal{HS}(\mathcal{H})$, then it converges with respect to the norm $\|\cdot\|_2$.*
- (e) *The set $\mathcal{F}(\mathcal{H})$ is dense in $\mathcal{T}(\mathcal{H})$ with respect to the norm $\|\cdot\|_1$ and in the set $\mathcal{HS}(\mathcal{H})$ with respect to the norm $\|\cdot\|_2$.*
- (f) *If $T \in \mathcal{T}(\mathcal{H})$ then $\|T\| \leq \|T\|_2 \leq \|T\|_1$.*

Proof (a) We may assume that $\varphi \neq 0, \psi \neq 0$. As

$$|\varphi\rangle\langle\psi| = \|\varphi\| \|\psi\| \left| \|\varphi\|^{-1}\varphi \right\rangle \left\langle \|\psi\|^{-1}\psi \right|,$$

the claim follows from Theorem 3.10.

(b) If $\varphi_i \neq 0, \psi_i \neq 0$, then by Theorem 3.10 (b)

$$\text{tr}[|\varphi_i\rangle\langle\psi_i|] = \|\varphi_i\| \|\psi_i\| \langle \|\psi_i\|^{-1}\psi_i | \|\varphi_i\|^{-1}\varphi_i \rangle = \langle\psi_i|\varphi_i\rangle,$$

and the claim follows from the linearity of the trace.

(c) If $P \in \mathcal{T}(\mathcal{H})$, then $P \in \mathcal{C}(\mathcal{H})$ (see Lemma 3.2 (b), Theorem 3.7), so that $P \in \mathcal{F}(\mathcal{H})$ (Theorem 3.3 (f)). The equality $\text{tr}[P] = \dim(P(\mathcal{H}))$ follows from (b) and Theorem 3.3 (a).

(d) The claims follow from Theorem 3.10, since for example

$$\left\| T - \sum_{n=1}^p c_n |\psi_n\rangle\langle\varphi_n| \right\|_2 = \left(\sum_{n=p+1}^{\infty} c_n^2 \right)^{\frac{1}{2}} \rightarrow 0, \text{ when } p \rightarrow \infty.$$

(e) This is a direct consequence of (d).

(f) The inequality $\|T\|_2 \leq \|T\|_1$ follows from Theorem 3.10, for if $0 \neq c = \sum_{n=1}^{\infty} c_n$

then $\sum_{n=1}^{\infty} (c_n c^{-1})^2 \leq \sum_{n=1}^{\infty} c_n c^{-1} = 1$, so that $\left(\sum_{n=1}^{\infty} (c_n c^{-1})^2 \right)^{\frac{1}{2}} \leq 1$, and so

$$\left(\sum_{n=1}^{\infty} c_n^2 \right)^{\frac{1}{2}} \leq c.$$

The inequality $\|T\| \leq \|T\|_2$ was already seen in Theorem 3.7 (a). $\square$

3.6 The Dualities $\mathcal{C}(\mathcal{H})^* = \mathcal{T}(\mathcal{H})$ and $\mathcal{T}(\mathcal{H})^* = \mathcal{L}(\mathcal{H})$

The trace class of the Hilbert space $\mathcal{H}$ has an important role in operator theory. For example, equipped with the norm introduced in Definition 3.5, as a normed space $\mathcal{T}(\mathcal{H})$ can be identified with the dual of the space $\mathcal{C}(\mathcal{H})$ of compact operators, and the dual of $\mathcal{T}(\mathcal{H})$ in turn with $\mathcal{L}(\mathcal{H})$. We prove these results in this section.

Theorem 3.12 *Define for each $T \in \mathcal{T}(\mathcal{H})$ the mapping $f_T : \mathcal{C}(\mathcal{H}) \rightarrow \mathbb{C}$ via the formula $f_T(S) = \text{tr}[TS]$, $S \in \mathcal{C}(\mathcal{H})$. In this way we obtain a linear isometric bijection $T \mapsto f_T$ from the space $\mathcal{T}(\mathcal{H})$ (equipped with the norm $\|\cdot\|_1$) onto the dual of $\mathcal{C}(\mathcal{H})$.*

Proof By Theorem 3.8 f_T is a linear functional defined on $\mathcal{C}(\mathcal{H})$. It follows from Theorems 3.9 and 3.8 that f_T is continuous and $\|f_T\| \leq \|T\|_1$. Let now $f \in \mathcal{C}(\mathcal{H})^*$. Since by Theorem 3.7 $\mathcal{HS}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$ and $\|S\|_2 \geq \|S\|$ for all $S \in \mathcal{HS}(\mathcal{H})$, the restriction $f|_{\mathcal{HS}(\mathcal{H})}$ is a continuous linear functional on the Hilbert space $\mathcal{HS}(\mathcal{H})$, so that by the Fréchet–Riesz theorem there is a $T \in \mathcal{HS}(\mathcal{H})$ satisfying $f(S) = \langle T^* | S \rangle = \sum_{\xi \in K} \langle T^* \xi | S \xi \rangle = \sum_{\xi \in K} \langle \xi | T S \xi \rangle = \text{tr}[TS]$ for all $S \in \mathcal{HS}(\mathcal{H})$ (here $TS \in \mathcal{T}(\mathcal{H})$ by Lemma 3.2 (a)). We prove that $T \in \mathcal{T}(\mathcal{H})$, $f = f_T$ and $\|f_T\| \geq \|T\|_1$. Let $T = V|T|$ be the polar decomposition of T . Then $|T| = V^*T$ and hence by Theorem 3.8

$$|\text{tr}[S|T|]| = |\text{tr}[SV^*T]| = |\text{tr}[TSV^*]| = |f(SV^*)| \leq \|f\| \|S\| \quad (3.10)$$

whenever $S \in \mathcal{F}(\mathcal{H})$ (implying $SV^* \in \mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$). Let now

$$|T| = \sum_{n=1}^{\infty} c_n |\varphi_n\rangle \langle \varphi_n|$$

be the representation given by (3.1'') (so that $c_n \geq 0$). If P_k is the projection onto the subspace spanned by $\varphi_1, \dots, \varphi_k$, then by (3.10) and Theorem 3.10 (b)

$$\sum_{n=1}^k c_n = \text{tr} \left[\sum_{n=1}^k c_n |\varphi_n\rangle \langle \varphi_n| \right] = \text{tr}[P_k |T|] \leq \|f\|,$$

so that $\sum_{k=1}^{\infty} c_k < \infty$, and by Theorem 3.10 (b) $|T| \in \mathcal{T}(\mathcal{H})$. Hence $T \in \mathcal{T}(\mathcal{H})$ and $\|T\|_1 = \text{tr}[|T|] \leq \|f\|$. The functional f and f_T , which are continuous functions on $\mathcal{C}(\mathcal{H})$, agree on the dense subspace $\mathcal{F}(\mathcal{H})$, and so $f = f_T$. The mapping $T \mapsto f_T$ is clearly linear, and when it is shown to be an injection, the theorem is proved. But if $T \in \mathcal{T}(\mathcal{H})$ is such that $f_T = 0$, then $\langle \varphi | T \psi \rangle = \text{tr}[T|\psi\rangle\langle\varphi|] = f_T(|\psi\rangle\langle\varphi|) = 0$ for all $\varphi, \psi \in \mathcal{H}$ (see Theorems 3.3 (b), 3.11 (b)), so that $T = 0$. $\square$

Since $\mathcal{C}(\mathcal{H})^*$ is a Banach space, we immediately obtain the following corollary:

Corollary 3.2 *The space $(\mathcal{T}(\mathcal{H}), \|\cdot\|_1)$ is complete.*

Theorem 3.13 *For each $S \in \mathcal{L}(\mathcal{H})$ define the mapping $f_S : \mathcal{T}(\mathcal{H}) \rightarrow \mathbb{C}$ via the formula $f_S(T) = \text{tr}[ST]$ for all $T \in \mathcal{T}(\mathcal{H})$. In this way we get an isometric linear bijection $S \mapsto f_S$ from the space $\mathcal{L}(\mathcal{H})$ onto the dual of the Banach space $(\mathcal{T}(\mathcal{H}), \|\cdot\|_1)$.*

Proof From Theorem 3.8 it follows that each f_S is defined on $\mathcal{T}(\mathcal{H})$ and linear. By Theorem 3.9 $|f_S(T)| = |\text{tr}[ST]| \leq \|S\| \|T\|_1$ for all $T \in \mathcal{T}(\mathcal{H})$, so that f_S is continuous and $\|f_S\| \leq \|S\|$. The mapping $S \mapsto f_S$ is clearly linear. If $\varphi, \psi \in \mathcal{H}$, then by using Theorems 3.3 (b) and 3.11 (b) we see that

$$\begin{aligned} |\langle \varphi | S \psi \rangle| &= |\text{tr}[|S\psi\rangle\langle\varphi|]| = |f_S(|\psi\rangle\langle\varphi|)| \\ &\leq \|f_S\| \|\psi\rangle\langle\varphi|\|_1 = \|f_S\| \|\psi\| \|\varphi\| \end{aligned}$$

so that

$$\|S\| = \sup \{ |\langle \varphi | S \psi \rangle| \mid \|\varphi\| \leq 1, \|\psi\| \leq 1 \} \leq \|f_S\|.$$

Thus the mapping $S \mapsto f_S$ is isometric and in particular injective. We still show that it is a surjection onto the dual of $\mathcal{T}(\mathcal{H})$. Let $f \in \mathcal{T}(\mathcal{H})^*$. We define a mapping $B : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ via the formula $B(\varphi, \psi) = f(|\psi\rangle\langle\varphi|)$. Then $|B(\varphi, \psi)| \leq \|f\| \|\psi\rangle\langle\varphi|\|_1 = \|f\| \|\varphi\| \|\psi\|$ by Theorem 3.11 (a). Since B is also conjugate linear with respect to the first and linear with respect to the second argument, by Proposition 2.1 there is $S \in \mathcal{L}(\mathcal{H})$ such that $B(\varphi, \psi) = \langle \varphi | S \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$. We show that $f = f_S$, i.e.

$$\text{tr}[ST] = f(T) \text{ for all } T \in \mathcal{T}(\mathcal{H}). \quad (3.11)$$

Both sides of (3.11) are linear functions of T that are continuous on $\mathcal{T}(\mathcal{H})$ with respect to the norm $\|\cdot\|_1$ (see Theorem 3.9). Since $\mathcal{F}(\mathcal{H})$ is dense in $\mathcal{T}(\mathcal{H})$ (see Theorem 3.11 (e)), in view of Theorem 3.3 (c) it is enough to show that (3.11) holds whenever $T = |\psi\rangle\langle\varphi|$ for some $\varphi, \psi \in \mathcal{H}$. But according to Theorems 3.3 (b) and 3.11 (b) we get

$$\text{tr}[S|\psi\rangle\langle\varphi|] = \text{tr}[|S\psi\rangle\langle\varphi|] = \langle \varphi | S \psi \rangle = B(\varphi, \psi) = f(|\psi\rangle\langle\varphi|). \quad \square$$

Remark 3.3 Let us recall the notation $\mathcal{L}_s(\mathcal{H}) = \{T \in \mathcal{L}(\mathcal{H}) | T = T^*\}$. In the sequel we write $\mathcal{T}_s(\mathcal{H}) = \{T \in \mathcal{T}(\mathcal{H}) | T = T^*\}$, and $\mathcal{T}_s(\mathcal{H})_1^+ = \{T \in \mathcal{T}(\mathcal{H}) | T \geq 0, \text{tr}[T] = 1\}$. Both $\mathcal{L}_s(\mathcal{H})$ and $\mathcal{T}_s(\mathcal{H})$ are in a natural way real Banach spaces, for they are closed with respect to addition and multiplication by a real number, and in addition norm closed because e.g. $\|T^*\|_1 = \|T\|_1$ for all $T \in \mathcal{T}(\mathcal{H})$ (a consequence of Theorems 3.3 (b) and 3.10 (b)). We now prove a result which shows that the dual $\mathcal{T}_s(\mathcal{H})^*$ of the real Banach space $\mathcal{T}_s(\mathcal{H})$ (i.e. the space of the continuous $\mathbb{R}$ -linear functionals on $\mathcal{T}_s(\mathcal{H})$) and $\mathcal{L}_s(\mathcal{H})$ can be identified. $\triangleleft$

Theorem 3.14 *Let $S \in \mathcal{L}(\mathcal{H})$ and f_S be as in Theorem 3.13.*

- (a) $S \geq 0$, if and only if $f_S(T) \geq 0$ whenever $T \in \mathcal{T}_s(\mathcal{H})$, $T \geq 0$.
- (b) $S = S^*$, if and only if $f_S(T) \in \mathbb{R}$ whenever $T \in \mathcal{T}_s(\mathcal{H})$, and then

$$\|S\| = \|f_S\| = \sup\{|f_S(T)| \mid T \in \mathcal{T}_s(\mathcal{H}), \|T\|_1 \leq 1\}.$$

- (c) If $g : \mathcal{T}_s(\mathcal{H}) \rightarrow \mathbb{R}$ is a continuous $\mathbb{R}$ -linear functional, there is a uniquely determined $S \in \mathcal{L}_s(\mathcal{H})$ such that $g = f_S|_{\mathcal{L}_s(\mathcal{H})}$.
- (d) The mapping $S \mapsto g_S$, where $g_S(T) = \text{tr}[ST]$, is an isometric linear bijection from $\mathcal{L}_s(\mathcal{H})$ onto the real Banach space $\mathcal{T}_s(\mathcal{H})^*$.

Proof (a) This follows from the fact that $f_S(|\varphi\rangle\langle\varphi|) = \text{tr}[S|\varphi\rangle\langle\varphi|] = \langle\varphi|S\varphi\rangle$ and each $T \in \mathcal{T}(\mathcal{H})$, $T \geq 0$, is of the form $\sum t_i |\varphi_i\rangle\langle\varphi_i|$, $t_i \geq 0$ (see Theorem 3.10).

(b) The “if” part is obtained from (a) by representing a selfadjoint operator as the difference of two positive operators. (In the case of $\mathcal{L}(\mathcal{H})$ see Proposition 2.4). If $D \in \mathcal{T}_s(\mathcal{H})$, then $D = D_1 - D_2$ with both $D_i \in \mathcal{T}(\mathcal{H})_1^+$ as can be seen by applying Theorem 3.10. The equality $\langle\varphi|S\varphi\rangle = \text{tr}[S|\varphi\rangle\langle\varphi|]$ for all $\varphi \in \mathcal{H}$ proves the “if” part. The first norm equality was noted in Theorem 3.13, and in the latter case obviously “ $\geq$ ” holds. On the other hand, for each $\varepsilon > 0$ there is some $T \in \mathcal{T}(\mathcal{H})$ satisfying $\|T\|_1 \leq 1$ and $|f_S(T)| \geq \|f_S\| - \varepsilon$, and by multiplying T with a suitable complex number c with $|c| = 1$ we may assume that $f_S(T) \geq 0$. Then

$$|f_S((T + T^*))| = \frac{1}{2}|f_S(T) + f_S(T^*)| = f_S(T) \geq \|f_S\| - \varepsilon$$

(for $\overline{f_S(T)} = f_S(T^*)$, since $f_S(T) + f_S(T^*) \in \mathbb{R}$ and $if_S(T) - if_S(T^*) = f_S(iT - iT^*) \in \mathbb{R}$).

(c) Define $f(T) = g(\frac{1}{2}(T + T^*)) + ig(\frac{1}{2i}(T - T^*))$, $T \in \mathcal{T}(\mathcal{H})$. Then in view of Theorem 3.13 f has the form $f = f_S$ for some $S \in \mathcal{L}_s(\mathcal{H})$. This S is uniquely determined e.g. by the norm formula in (b).

(d) This follows from (b) and (c). $\square$

3.7 Linear Operators on Hilbert Tensor Products and the Partial Trace

In this section we give some results concerning the linear operators on the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$ of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. Here and elsewhere, when needed, to avoid trivialities we may assume that our Hilbert spaces contain nonzero elements. All the familiar operator algebras such as the set $\mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ of bounded operators, the trace class $\mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ and the set $\mathcal{HS}(\mathcal{H} \otimes \mathcal{K})$ of Hilbert–Schmidt operators on $\mathcal{H} \otimes \mathcal{K}$ are defined in the usual way. We also use the notation $\mathcal{P}(\mathcal{H} \otimes \mathcal{K})$ for the projection lattice of the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$. In what follows, we show that one may construct operators on the tensor product space $\mathcal{H} \otimes \mathcal{K}$ out of operators of the component spaces $\mathcal{H}$ and $\mathcal{K}$.

Lemma 3.4 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and $S \in \mathcal{L}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{K})$. Then*

$$\left\| \sum_{j=1}^n S\varphi_j \otimes T\psi_j \right\| \leq \|S\| \|T\| \left\| \sum_{j=1}^n \varphi_j \otimes \psi_j \right\| \quad (3.12)$$

for all $\varphi_j \in \mathcal{H}$ and $\psi_j \in \mathcal{K}$, $j = 1, \dots, n$, when n is an arbitrary natural number.

Proof First, suppose that $T = I_{\mathcal{K}}$. Let $\varphi_j \in \mathcal{H}$ and $\psi_j \in \mathcal{K}$, $j = 1, \dots, n$. Consider first the case when the set $\{\psi_1, \dots, \psi_n\}$ is orthonormal. We now have

$$\begin{aligned} \left\| \sum_{j=1}^n S\varphi_j \otimes \psi_j \right\|^2 &= \sum_{j,k=1}^n \langle S\varphi_j | S\varphi_k \rangle \langle \psi_j | \psi_k \rangle = \sum_{j=1}^n \|S\varphi_j\|^2 \\ &\leq \|S\|^2 \sum_{j,k=1}^n \langle \varphi_j | \varphi_k \rangle \langle \psi_j | \psi_k \rangle = \|S\|^2 \left\| \sum_{j=1}^n \varphi_j \otimes \psi_j \right\|^2. \end{aligned}$$

In the general case of a not necessarily orthonormal set $\{\psi_1, \dots, \psi_n\}$ choose a finite orthonormal set $\{\eta_1, \dots, \eta_m\} \subset \mathcal{K}$ such that $\psi_j = \sum_{k=1}^m c_{jk} \eta_k$ for all $j = 1, \dots, n$ with some complex numbers c_{jk} , $j = 1, \dots, n, k = 1, \dots, m$. Then using the above calculation we get

$$\begin{aligned} \left\| \sum_{j=1}^n S\varphi_j \otimes \psi_j \right\| &= \left\| \sum_{k=1}^m \left(\sum_{j=1}^n S(c_{jk} \varphi_j) \right) \otimes \eta_k \right\| \\ &\leq \|S\| \left\| \sum_{j=1}^n \sum_{k=1}^m c_{jk} \varphi_j \otimes \eta_k \right\| = \|S\| \left\| \sum_{j=1}^n \varphi_j \otimes \psi_j \right\|. \end{aligned}$$

Assuming $S = I_{\mathcal{H}}$ we obtain the same result for a general $T \in \mathcal{L}(\mathcal{K})$. Combining these results gives us the inequality (3.12). $\square$

Lemma 3.4 shows that whenever $\sum_{j=1}^n \varphi_j \otimes \psi_j = \sum_{k=1}^m \xi_k \otimes \eta_k$, then

$$\left\| \sum_{j=1}^n S\varphi_j \otimes T\psi_j - \sum_{k=1}^m S\xi_k \otimes T\eta_k \right\| \leq \|S\| \|T\| \left\| \sum_{j=1}^n \varphi_j \otimes \psi_j - \sum_{k=1}^m \xi_k \otimes \eta_k \right\| = 0.$$

Thus the map $\sum_{j=1}^n \varphi_j \otimes \psi_j \mapsto \sum_{j=1}^n S\varphi_j \otimes T\psi_j$ is well defined on a dense subspace of $\mathcal{H} \otimes \mathcal{K}$ and since it is clearly linear, we may extend it to a bounded linear map on the whole of $\mathcal{H} \otimes \mathcal{K}$. Let us formulate this result in a proposition.

Proposition 3.1 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and $S \in \mathcal{L}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{K})$. There is a unique operator $S \otimes T \in \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ such that*

$$(S \otimes T)(\varphi \otimes \psi) = S\varphi \otimes T\psi$$

for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$.

The linear operator $S \otimes T$ of the above proposition is called the (*Hilbert*) *tensor product* of the operators S and T . Next we list some basic properties of the tensor product operators.

Proposition 3.2 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and $S, S_1, S_2 \in \mathcal{L}(\mathcal{H})$ and $T, T_1, T_2 \in \mathcal{L}(\mathcal{K})$. Then*

- (a) $\alpha(S \otimes T) = (\alpha S) \otimes T = S \otimes (\alpha T)$ for all $\alpha \in \mathbb{C}$;
- (b) $(S_1 + S_2) \otimes T = S_1 \otimes T + S_2 \otimes T$;
- (c) $(S_1 \otimes T_1)(S_2 \otimes T_2) = S_1 S_2 \otimes T_1 T_2$;
- (d) $(S \otimes T)^* = S^* \otimes T^*$;
- (e) if S and T are selfadjoint, then also $S \otimes T$ is selfadjoint;
- (f) if S and T are unitary, then also $S \otimes T$ is unitary;
- (g) if $S \in \mathcal{P}(\mathcal{H})$ and $T \in \mathcal{P}(\mathcal{K})$, then $S \otimes T \in \mathcal{P}(\mathcal{H} \otimes \mathcal{K})$;
- (h) if $S \in \mathcal{T}(\mathcal{H})$ and $T \in \mathcal{T}(\mathcal{K})$, then $S \otimes T \in \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ and

$$\text{tr}[S \otimes T] = \text{tr}[S]\text{tr}[T];$$

- (i) if $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$, then

$$|\varphi_1 \otimes \psi_1\rangle \langle \varphi_2 \otimes \psi_2| = |\varphi_1\rangle \langle \varphi_2| \otimes |\psi_1\rangle \langle \psi_2|.$$

Proof (a), (b) For all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$ we have

$$(\alpha(S \otimes T))(\varphi \otimes \psi) = \alpha(S\varphi \otimes T\psi) = \alpha S\varphi \otimes T\psi = ((\alpha S) \otimes T)(\varphi \otimes \psi)$$

so that $\alpha(S \otimes T) = (\alpha S) \otimes T$ since finite sums of simple tensors form a dense subspace of $\mathcal{H} \otimes \mathcal{K}$. In the same way one sees that $\alpha(S \otimes T) = S \otimes (\alpha T)$ and $(S_1 + S_2) \otimes T = S_1 \otimes T + S_2 \otimes T$.

(c) Again, for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$ we have $(S_1 \otimes T_1)(S_2 \otimes T_2)(\varphi \otimes \psi) = (S_1 \otimes T_1)(S_2 \varphi \otimes T_2 \psi) = S_1 S_2 \varphi \otimes T_1 T_2 \psi$ so that the argument is concluded as above.

(d) Let $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$. One has

$$\begin{aligned} \langle \varphi_1 \otimes \psi_1 | (S \otimes T) \varphi_2 \otimes \psi_2 \rangle &= \langle \varphi_1 \otimes \psi_1 | S \varphi_2 \otimes T \psi_2 \rangle = \langle \varphi_1 | S \varphi_2 \rangle \langle \psi_1 | T \psi_2 \rangle \\ &= \langle S^* \varphi_1 | \varphi_2 \rangle \langle T^* \psi_1 | \psi_2 \rangle = \langle S^* \varphi_1 \otimes T^* \psi_1 | \varphi_2 \otimes \psi_2 \rangle \\ &= \langle (S^* \otimes T^*)(\varphi_1 \otimes \psi_1) | \varphi_2 \otimes \psi_2 \rangle. \end{aligned}$$

Since the linear combinations of the simple tensors are dense in the Hilbert tensor product the claim follows.

The claim in (e) is an immediate consequence of (d).

(f) Clearly $I_{\mathcal{H} \otimes \mathcal{K}} = I_{\mathcal{H}} \otimes I_{\mathcal{K}}$. Suppose that S and T are unitary. Then $(S \otimes T)^*(S \otimes T) = (S^* \otimes T^*)(S \otimes T) = S^* S \otimes T^* T = I_{\mathcal{H}} \otimes I_{\mathcal{K}} = (S \otimes T)(S \otimes T)^*$ implying the claim.

(g) $(S \otimes T)^2 = S^2 \otimes T^2 = S \otimes T = S^* \otimes T^* = (S \otimes T)^*$ whenever $S = S^* = S^2$ and $T = T^* = T^2$.

(h) Noting that $|S \otimes T|^2 = (S \otimes T)^*(S \otimes T) = S^* S \otimes T^* T = |S|^2 \otimes |T|^2 = (|S| \otimes |T|)^2$, we have $|S \otimes T| = |S| \otimes |T|$ by the uniqueness of the positive square root. Let $K \subset \mathcal{H}$ and $L \subset \mathcal{K}$ be orthonormal bases. As $\sum_{\xi \in K} \langle \xi | |S| \xi \rangle$ and $\sum_{\eta \in L} \langle \eta | |T| \eta \rangle < \infty$, we have

$$\begin{aligned} \sum_{\xi \in K, \eta \in L} \langle \xi \otimes \eta | |S \otimes T| \xi \otimes \eta \rangle &= \sum_{\xi \in K, \eta \in L} \langle \xi | |S| \xi \rangle \langle \eta | |T| \eta \rangle \\ &= \sum_{\xi \in K} \langle \xi | |S| \xi \rangle \sum_{\eta \in L} \langle \eta | |T| \eta \rangle < \infty, \end{aligned}$$

so that $S \otimes T \in \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$. Since the sets $(\langle \xi | S \xi \rangle)_{\xi \in K}$ and $(\langle \eta | T \eta \rangle)_{\eta \in L}$ are summable, we get

$$\begin{aligned} \text{tr}[S \otimes T] &= \sum_{\xi \in K, \eta \in L} \langle \xi \otimes \eta | (S \otimes T)(\xi \otimes \eta) \rangle = \sum_{\xi \in K, \eta \in L} \langle \xi | S \xi \rangle \langle \eta | T \eta \rangle \\ &= \sum_{\xi \in K} \langle \xi | S \xi \rangle \sum_{\eta \in L} \langle \eta | T \eta \rangle = \text{tr}[S] \text{tr}[T]. \end{aligned}$$

(i) It suffices to observe that either side applied to $\xi \otimes \eta$ for any $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$ yields $\langle \varphi_2 | \xi \rangle \langle \psi_2 | \eta \rangle (\varphi_1 \otimes \psi_1)$. $\square$

Remark 3.4 All the above properties of tensor product operators $S \otimes T$ can be generalised in a straightforward way to the case of general tensor product operators $S_1 \otimes \cdots \otimes S_n$ on Hilbert tensor products $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$. We may also use obvious associativity properties without explicit justification. $\triangleleft$

We conclude this section by examining the trace class operators of the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$. We see that any operator $T \in \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ defines two trace class operators $T_I \in \mathcal{T}(\mathcal{H})$ and $T_{II} \in \mathcal{T}(\mathcal{K})$ through a linear map called the partial trace.

Proposition 3.3 *If $T \in \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$, there is a unique $T_I \in \mathcal{T}(\mathcal{H})$ such that*

$$\mathrm{tr}[T_I A] = \mathrm{tr}[T(A \otimes I_{\mathcal{K}})] \quad (3.13)$$

for all $A \in \mathcal{L}(\mathcal{H})$. In particular, if $T \in \mathcal{T}_s(\mathcal{H} \otimes \mathcal{K})_1^+$, then $T_I \in \mathcal{T}_s(\mathcal{H})_1^+$.

Proof First we show that if there is a trace class operator T_I satisfying the condition (3.13), it is unique. Suppose that T'_I is another such operator. The substitution $A = |\psi\rangle\langle\varphi|$ in (3.13) yields $\langle\varphi|T_I\psi\rangle = \langle\varphi|T'_I\psi\rangle$. Since this is valid for any $\varphi \in \mathcal{H}$, we have $T'_I = T_I$. Next we show that such an operator exists. According to Remark 3.1 we may write $T = \sum_{n=1}^{\infty} c_n |\psi_n\rangle\langle\varphi_n|$ where $c_n \geq 0$ and (ψ_n) and (φ_n) are orthonormal sequences in $\mathcal{H} \otimes \mathcal{K}$. If for each n there is an operator $S_n \in \mathcal{T}(\mathcal{H})$ such that $\mathrm{tr}[S_n A] = \mathrm{tr}[|\psi_n\rangle\langle\varphi_n|(A \otimes I_{\mathcal{K}})]$ for all $A \in \mathcal{L}(\mathcal{H})$, then we can take $T_I = \sum_{n=1}^{\infty} c_n S_n$. We may thus assume that $T = |\psi\rangle\langle\varphi|$ for some $\varphi, \psi \in \mathcal{H} \otimes \mathcal{K}$. Choose orthonormal bases $K \subset \mathcal{H}$, $L \subset \mathcal{K}$. Denoting $\varphi_{\eta} = \sum_{\xi \in K} \langle\xi \otimes \eta|\varphi\rangle \xi$ and $\psi_{\eta} = \sum_{\xi \in K} \langle\xi \otimes \eta|\psi\rangle \xi$ for all $\eta \in L$, one immediately sees that $\varphi = \sum_{\eta \in L} \varphi_{\eta} \otimes \eta$, $\psi = \sum_{\eta \in L} \psi_{\eta} \otimes \eta$, $\sum_{\eta \in L} \|\varphi_{\eta}\|^2 = \|\varphi\|^2$ and $\sum_{\eta \in L} \|\psi_{\eta}\|^2 = \|\psi\|^2$. Define the operator $T_I = \sum_{\eta \in L} |\psi_{\eta}\rangle\langle\varphi_{\eta}|$. This operator belongs to the trace class since $\| |\psi_{\eta}\rangle\langle\varphi_{\eta}| \|_1 = \|\psi_{\eta}\| \|\varphi_{\eta}\|$ and

$$\sum_{\eta \in L} \|\psi_{\eta}\| \|\varphi_{\eta}\| \leq \sqrt{\sum_{\eta \in L} \|\psi_{\eta}\|^2} \sqrt{\sum_{\eta \in L} \|\varphi_{\eta}\|^2} = \|\varphi\| \|\psi\| < \infty.$$

We now have

$$\begin{aligned} \mathrm{tr}[|\psi\rangle\langle\varphi|(A \otimes I_{\mathcal{K}})] &= \langle\varphi|(A \otimes I_{\mathcal{K}})\psi\rangle = \sum_{\eta_1, \eta_2 \in L} \langle\varphi_{\eta_1} \otimes \eta_1|A\psi_{\eta_2} \otimes \eta_2\rangle \\ &= \sum_{\eta \in L} \langle\varphi_{\eta}|A\psi_{\eta}\rangle = \mathrm{tr}[T_I A]. \end{aligned}$$

Thus T_I satisfies the condition (3.13). Suppose now that T is positive and choose any $\varphi_I \in \mathcal{H}$. One finds that

$$\langle\varphi_I|T_I\varphi_I\rangle = \mathrm{tr}[T(|\varphi_I\rangle\langle\varphi_I| \otimes I_{\mathcal{K}})] = \sum_{\eta \in L} \langle\varphi_I \otimes \eta|T\varphi_I \otimes \eta\rangle \geq 0$$

implying that $T_I \geq 0$. Substituting $A = I_{\mathcal{H}}$ in (3.13) one sees that $\mathrm{tr}[T] = \mathrm{tr}[T_I]$ and thus especially when T is of trace 1, then T_I is also of trace 1. $\square$

Remark 3.5 The map $T \mapsto T_I$ from $\mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ to $\mathcal{T}(\mathcal{H})$ defined through Eq. (3.13) is easily seen to be linear. Proposition 3.3 also shows that this map is trace preserving (i.e. $\mathrm{tr}[T_I] = \mathrm{tr}[T]$) and positive (i.e. $T_I \geq 0$ if $T \geq 0$). The map $T \mapsto T_I$ of Proposition 3.3 is called *the partial trace* (over the Hilbert space $\mathcal{K}$) and it is denoted $T_I = \mathrm{tr}_{\mathcal{K}}[T] = \mathrm{tr}_{II}[T]$. It is obvious that we may define the partial trace $\mathrm{tr}_{\mathcal{H}} : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}) \rightarrow \mathcal{T}(\mathcal{K})$ over $\mathcal{H}$ in a completely analogous manner. $\triangleleft$

3.8 The Schmidt Decomposition of an Element of $\mathcal{H}_1 \otimes \mathcal{H}_2$

The formula in the next theorem gives the *Schmidt decomposition* (also known as the *polar* or *biorthogonal decomposition*) of an element in the Hilbert tensor product of two Hilbert spaces. Its proof involves in a crucial way some of the most central techniques developed above, like the polar decomposition of an operator and the spectral theory of compact operators. In a sense it wraps up the theory we have seen so far. We let $\mathcal{H}_1$ and $\mathcal{H}_2$ denote two Hilbert spaces.

Theorem 3.15 (a) Any nonzero $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be expressed as the sum of a norm convergent series

$$\Psi = \sum_i \lambda_i \varphi_i \otimes \psi_i$$

where (φ_i) resp. (ψ_i) is an orthonormal sequence (finite or infinite) in $\mathcal{H}_1$ resp. $\mathcal{H}_2$, and each $\lambda_i > 0$. In this kind of representation we always have $\sum_i \lambda_i^2 = \|\Psi\|^2$.

(b) Suppose that

$$\Psi = \sum_i \mu_i \xi_i \otimes \eta_i$$

is another representation having the properties mentioned in (a). Assume that $\lambda_1 \geq \lambda_2 \geq \dots$ and $\mu_1 \geq \mu_2 \geq \dots$. Then $\lambda_i = \mu_i$ for each i . For any i the sets $\{\varphi_j \mid \lambda_j = \lambda_i\}$ and $\{\xi_j \mid \mu_j = \mu_i\}$ span the same (finite-dimensional) subspace of $\mathcal{H}_1$, and similarly the sets $\{\psi_j \mid \lambda_j = \lambda_i\}$ and $\{\eta_j \mid \mu_j = \mu_i\}$ span the same subspace of $\mathcal{H}_2$. In case some number λ_i (and hence also μ_i) occurs only once, we must have $\xi_i = c_i \varphi_i$ and $\eta_i = c_i^{-1} \psi_i$ for some $c_i \in \mathbb{C}$ with $|c_i| = 1$.

Proof We first consider the special case where $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. Choose an orthonormal basis K for $\mathcal{H}$ and let $J : \mathcal{H} \rightarrow \mathcal{H}$ be defined as the conjugate-linear isometric map which changes the coefficients of any $\varphi \in \mathcal{H}$ in the expansion with respect to K to their complex conjugates. Define the bilinear map $f : \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{HS}(\mathcal{H})$ via $f(\varphi, \psi) = |\varphi\rangle\langle J\psi|$. Using Theorems 3.6, 3.3 (b), 3.11 and 2.20 we get, for any $\varphi, \psi, \xi, \eta \in \mathcal{H}$,

$$\begin{aligned} \langle f(\varphi, \psi) \mid f(\xi, \eta) \rangle &= \text{tr}[f(\varphi, \psi)^* f(\xi, \eta)] = \text{tr}[|J\psi\rangle\langle\varphi||\xi\rangle\langle J\eta|] \\ &= \langle \varphi \mid \xi \rangle \langle J\eta \mid J\psi \rangle = \langle \varphi \mid \xi \rangle \langle \psi \mid \eta \rangle. \end{aligned}$$

Since the finite rank operators are dense in $\mathcal{HS}(\mathcal{H})$, we are thus allowed to use Lemma 2.4 to conclude that there is an isometric isomorphism $g : \mathcal{H} \otimes \mathcal{H} \rightarrow \mathcal{HS}(\mathcal{H})$ satisfying $g(\varphi \otimes \psi) = f(\varphi, \psi)$. Using Remark 3.1 (b) we find some positive numbers $\lambda_1 \geq \lambda_2 \geq \dots$ and some (finite or infinite) orthonormal sequences (φ_i) and $(J\psi_i)$ such that

$$g(\Psi) = \sum_i \lambda_i |\varphi_i\rangle\langle J\psi_i|$$

(convergence in the operator norm and also in the Hilbert–Schmidt norm by Theorem 3.11 (d)). As $J = J^{-1}$ preserves orthogonality, (ψ_i) is also an orthonormal sequence. We have $\Psi = g^{-1}[\sum_i \lambda_i |\varphi_i\rangle\langle J\psi_i|] = \sum_i \lambda_i g^{-1}(|\varphi_i\rangle\langle J\psi_i|) = \sum_i \lambda_i \varphi_i \otimes \psi_i$, the last sum converging in the norm of $\mathcal{H} \otimes \mathcal{H}$. We have thus proved the first claim of (a) in this special case. The equation involving the norm also follows for this particular construction. After we have proved (b), we see it generally (or else prove it by a direct calculation, exercise).

We now consider the uniqueness part (b), still with $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. Since g is an isometry, we get

$$\sum_i \lambda_i |\varphi_i\rangle\langle J\psi_i| = g(\Psi) = g\left(\sum_i \mu_i \xi_i \otimes \eta_i\right) = \sum_i \mu_i |\xi_i\rangle\langle J\eta_i|.$$

From Remark 3.2 we see at once the equalities $\lambda_i = \mu_i$. Moreover, the linear spans $\text{lin}\{J\psi_j \mid \lambda_j = \lambda_i\}$ and $\text{lin}\{J\eta_j \mid \mu_j = \mu_i\}$ are the same, and so are $\text{lin}\{\varphi_j \mid \lambda_j = \lambda_i\}$ and $\text{lin}\{\xi_j \mid \lambda_j = \lambda_i\}$. Since $J(\text{lin}\{\psi_j \mid \lambda_j = \lambda_i\}) = \text{lin}\{J\psi_j \mid \lambda_j = \lambda_i\}$ and similarly $J(\text{lin}\{\eta_j \mid \mu_j = \mu_i\}) = \text{lin}\{J\eta_j \mid \mu_j = \mu_i\}$, we have $\text{lin}\{\psi_j \mid \lambda_j = \lambda_i\} = \text{lin}\{\eta_j \mid \mu_j = \mu_i\}$. Suppose now that some λ_i occurs only once. In this case we get $\xi_i = \alpha\varphi_i$ and $\eta_i = \beta\psi_i$ for some $\alpha, \beta \in \mathbb{C}$. The maps $|\varphi_i\rangle\langle J\psi_i|$ and $|\xi_i\rangle\langle J\eta_i|$ take the same value ($= \lambda_i^{-1}g(\Psi)(J\psi_i)$) at $J\psi_i$, and so $\langle J\psi_i \mid J\psi_i\rangle\varphi_i = \langle J\beta\psi_i \mid J\psi_i\rangle\alpha\varphi_i$ implying $1 = \beta\alpha$ and hence the last claim in (b), since all the vectors $\varphi_i, \psi_i, \xi_i, \eta_i$ have norm one.

In the general case we may clearly assume that $\mathcal{H}_1 \subset \mathcal{H}_2$ or $\mathcal{H}_2 \subset \mathcal{H}_1$. As our claims are invariant under the isometric isomorphism from $\mathcal{H}_1 \otimes \mathcal{H}_2$ onto $\mathcal{H}_2 \otimes \mathcal{H}_1$ mapping $\varphi \otimes \psi$ to $\psi \otimes \varphi$, we may in fact assume that $\mathcal{H}_1$ is just a closed subspace of $\mathcal{H}_2$. Take $\mathcal{H} = \mathcal{H}_2$ in the first part of the proof. Let D denote the closed linear span of the set $\{\xi \otimes \eta \mid \xi \in \mathcal{H}_1, \eta \in \mathcal{H}\}$. The map $(\xi, \eta) \mapsto \xi \otimes \eta$ from $\mathcal{H}_1 \times \mathcal{H}$ to D clearly satisfies the hypotheses of Lemma 2.4, and so it determines an isometric isomorphism from $\mathcal{H}_1 \otimes \mathcal{H}$ onto D . We identify $\mathcal{H}_1 \otimes \mathcal{H}$ with D via this isomorphism. The whole theorem will be proved if we show that whenever $\Psi \in D$ with $\Psi = \sum_i \lambda_i \varphi_i \otimes \psi_i$ as in the statement of (a), then necessarily every φ_i belongs to $\mathcal{H}_1$. We observe that the range of $g(\Psi)$ is contained in $\mathcal{H}_1$ for every $\Psi \in D$, since this is clearly true for $g(\varphi \otimes \psi) = |\varphi\rangle\langle J\psi|$ where $\varphi \in \mathcal{H}_1$, and an arbitrary $g(\Psi)$ with $\Psi \in D$ can be approximated in (the Hilbert–Schmidt norm and hence also in) the operator norm by finite sums of operators of the type $g(\varphi \otimes \psi)$. But in the representation $g(\Psi) = \sum_i \lambda_i \varphi_i \otimes \psi_i$ each $\varphi_i = \lambda_i^{-1}g(\Psi)(J\psi_i)$ is in the range of $g(\Psi)$. $\square$

We have the following useful application.

Proposition 3.4 *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and*

$$\Psi = \sum_i \lambda_i \varphi_i \otimes \psi_i$$

the Schmidt decomposition of a unit vector $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Then for the partial trace operators of the rank one projection $|\Psi\rangle\langle\Psi|$ we have the formulas

$$\mathrm{tr}_{\mathcal{H}_2}[|\Psi\rangle\langle\Psi|] = \sum_i \lambda_i^2 |\varphi_i\rangle\langle\varphi_i|$$

and

$$\mathrm{tr}_{\mathcal{H}_1}[|\Psi\rangle\langle\Psi|] = \sum_i \lambda_i^2 |\psi_i\rangle\langle\psi_i|.$$

Proof Denote $\Psi_n = \sum_{i=1}^n \lambda_i \varphi_i \otimes \psi_i$. Then

$$|\Psi_n\rangle\langle\Psi_n| = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j |\varphi_i \otimes \psi_i\rangle\langle\varphi_j \otimes \psi_j| = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j |\varphi_i\rangle\langle\varphi_j| \otimes |\psi_i\rangle\langle\psi_j|$$

by Proposition 3.2 (i). For any $A \in \mathcal{L}(\mathcal{H}_1)$ we get

$$\begin{aligned} \mathrm{tr}[(|\varphi_i\rangle\langle\varphi_j| \otimes |\psi_i\rangle\langle\psi_j|)(A \otimes I_{\mathcal{H}_2})] &= \mathrm{tr}[|\varphi_i\rangle\langle\varphi_j|A] \mathrm{tr}[|\psi_i\rangle\langle\psi_j|] \\ &= \mathrm{tr}[|\varphi_i\rangle\langle\varphi_j|A] \langle\psi_j|\psi_i\rangle \end{aligned}$$

(see Proposition 3.2 (h) and Theorem 3.11 (b)), and since the vectors ψ_j are orthonormal, we obtain

$$\mathrm{tr}[|\Psi_n\rangle\langle\Psi_n|(A \otimes I_{\mathcal{H}_2})] = \sum_{i=1}^n \lambda_i^2 \mathrm{tr}[|\varphi_i\rangle\langle\varphi_i|A] = \mathrm{tr}\left[\left(\sum_{i=1}^n \lambda_i^2 |\varphi_i\rangle\langle\varphi_i|\right)A\right].$$

Since the series $\sum_i \lambda_i^2 |\varphi_i\rangle\langle\varphi_i|$ converges in the trace norm (see Theorem 3.10 (b) and Theorem 3.11 (d)) and $|\Psi_n\rangle\langle\Psi_n| \rightarrow |\Psi\rangle\langle\Psi|$ in the trace norm (as can be seen using Theorem 3.11 (a)), from Theorem 3.11 (a) it follows that

$$\mathrm{tr}[|\Psi\rangle\langle\Psi|(A \otimes I_{\mathcal{H}_2})] = \mathrm{tr}\left[\left(\sum_i \lambda_i^2 |\varphi_i\rangle\langle\varphi_i|\right)A\right].$$

This proves the claim for $\mathrm{tr}_{\mathcal{H}_2}[|\Psi\rangle\langle\Psi|]$, and the proof for $\mathrm{tr}_{\mathcal{H}_1}[|\Psi\rangle\langle\Psi|]$ is similar. $\square$

Remark 3.6 The above proposition can be used to give an alternative proof for Proposition 3.3, thus reducing the partial trace result to the Schmidt decomposition. $\triangleleft$

3.9 Exercises

1. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded function and $T_g : \ell^2 \rightarrow \ell^2$ the bounded linear operator defined by the formula $T_g f = g f$ (see exercise 22 in Chap. 2). Show that T_g is a compact operator if $\lim_{n \rightarrow \infty} g(n) = 0$.
2. Is the condition $\lim_{n \rightarrow \infty} g(n) = 0$ also necessary for the operator T_g in the preceding exercise to be compact?

3. Let $T = VA$ be the polar decomposition of a compact operator T . Is
 - (a) the operator V ,
 - (b) the operator A
 necessarily compact?
4. Does there exist any injective compact operator $T \in \mathcal{L}(\ell^2)$?
5. Does there exist any surjective compact operator $T \in \mathcal{L}(\ell^2)$?
6. Let $T \in \mathcal{L}(\mathcal{H})$ be a positive compact operator. Show that the square root of T is compact. (Hint: use the spectral representation.)
7. Let $\mathcal{I} \neq \{0\}$ be a (not necessarily closed) two-sided ideal of $\mathcal{L}(\mathcal{H})$. Show that $\mathcal{I}$ contains every finite rank operator.
8. Using the fact that a subset of a metric space is compact if and only if it is precompact and complete, show that the space $\mathcal{C}(\mathcal{H})$ of the compact operators on $\mathcal{H}$ is norm closed in $\mathcal{L}(\mathcal{H})$.
9. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded function and $T_g : \ell^2 \rightarrow \ell^2$ the bounded linear map defined via the formula $T_g f = gf$. Show that T_g is a Hilbert–Schmidt operator if and only if $g \in \ell^2$.
10. Show that, in the situation of the preceding exercise, T_g belongs to the trace class $\mathcal{T}(\ell^2)$ if and only if $g \in \ell^1$, i.e. $\sum_{n=1}^{\infty} |g(n)| < \infty$.
11. Find some isometric isomorphism between the Hilbert space $\mathcal{HS}(\mathcal{H})$ (of the Hilbert–Schmidt operators on $\mathcal{H}$) and a suitable Hilbert sum $\sum_{x \in X}^{\oplus} \mathcal{H}_x$ where always $\mathcal{H}_x = \mathcal{H}$.
12. Let $\mathcal{I} \subset \mathcal{L}(\mathcal{H})$ be a vector subspace such that $ST \in \mathcal{I}$ whenever $S \in \mathcal{I}$ and $T \in \mathcal{L}(\mathcal{H})$. (Thus $\mathcal{I}$ is a right ideal.) Show that $\mathcal{I}$ is a two-sided ideal if and only if $S^* \in \mathcal{I}$ whenever $S \in \mathcal{I}$.
13. Let $S, T \in \mathcal{L}(\mathcal{H})$ be compact operators. Show that the map $X \mapsto SXT$ from the Banach space $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H})$ is a compact operator. (The definition of a compact operator is the same as in the case of a Hilbert space.)
14. Let $S, T \in \mathcal{L}(\mathcal{H}) \setminus \{0\}$ be such that the map $X \mapsto SXT$ from the Banach space $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H})$ is a compact operator. Show that S and T are compact operators.
15. Prove the formula $\langle J\eta | J\psi \rangle = \langle \psi | \eta \rangle$ which was used in the proof of Theorem 3.15.
16. Prove by a direct calculation the formula $\sum_i \lambda_i^2 = \|\Psi\|^2$ in Theorem 3.15.
17. Suppose that the Schmidt decomposition of $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ has only a finite number p of terms. Is it possible that Ψ might be written as a sum of fewer than p terms of the form $\varphi \otimes \psi$ with $\varphi \in \mathcal{H}_1, \psi \in \mathcal{H}_2$ (with no requirement of orthogonality)?
18. Let $\mathcal{H}$ be a finite-dimensional Hilbert space with an orthonormal basis $\{\epsilon_1, \dots, \epsilon_n\}$. Show that there is an inner product preserving isomorphism from $\mathcal{H}$ onto $\mathbb{C}^n$ mapping each ϵ_i to $e_i = (0, \dots, 0, 1, 0, \dots, 0)$. For simplicity, consider the tensor product $\mathcal{H} \otimes \mathcal{H}$. Show that $\mathcal{H} \otimes \mathcal{H}$ is isomorphic as a Hilbert space to $M_n(\mathbb{C})$, the Hilbert space of $n \times n$ matrices equipped with the inner product $\langle A | B \rangle = \text{tr}[A^*B]$ where $\text{tr}[C]$ is the sum of the main diagonal elements of the matrix C . Interpret and reprove the content of Theorem 3.15 in this situation by considering the polar decomposition of a matrix. Does one avoid invoking the map J used in the proof of Theorem 3.15 or is it lurking somewhere?

Chapter 4

Operator Integrals and Spectral Representations: The Bounded Case

In the preceding chapter the spectral theory of compact selfadjoint operators played a key role. One purpose of this chapter is to present the corresponding theory for arbitrary bounded selfadjoint operators. We begin with auxiliary techniques from the (scalar) theory of measure and integration. Then the notion of positive operator measure is introduced and studied. This is a key concept for the rest of the book, but in this chapter its special case, spectral measure, is needed for the spectral representation theory alluded to above. We also consider the two-variable case of positive operator bimeasures, an important notion for our physical applications, but at the same time a tool for the spectral representation of bounded normal, especially unitary, operators. The latter will have an instrumental role in the next chapter dealing with unbounded selfadjoint operators.

4.1 Classes of Sets and Positive Measures

In this chapter we first collect some basic material on measure and integration needed later. We omit many proofs. They can of course be found in many sources, but we mention especially [1, 2] whose presentations are fairly close to ours.

Throughout this chapter, Ω is a set. We denote by 2^Ω the set of its subsets. A collection of subsets $\mathcal{A} \subset 2^\Omega$ is called a σ -algebra if $\emptyset \in \mathcal{A}$ and $A^c (= \Omega \setminus A)$ and $\bigcup_{n=1}^\infty A_n$ belong to $\mathcal{A}$ whenever $A \in \mathcal{A}$ and $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$. If $\mathcal{F} \subset 2^\Omega$, the intersection of the σ -algebras containing $\mathcal{F}$ is clearly a σ -algebra; it is called the σ -algebra generated by $\mathcal{F}$. In particular, if Ω is a topological space (in practice usually $\mathbb{R}$ or $\mathbb{R}^n$), the σ -algebra generated by its topology (i.e. the class of open sets) is called the Borel σ -algebra of Ω and denoted by $\mathcal{B}(\Omega)$. The sets $B \in \mathcal{B}(\Omega)$ are Borel sets. We may use without explicit mention the easily proved fact that if X is a subset of a topological space Ω , then its Borel σ -algebra, i.e. the σ -algebra generated in 2^X by the relative (subspace) topology of X , is the same as the set of the intersections

$B \cap X$ where $B \in \mathcal{B}(\Omega)$. We may also call any set $A \in \mathcal{A}$ $\mathcal{A}$ -measurable. If $\mathcal{A} \subset 2^\Omega$ is a σ -algebra, we call the pair $(\Omega, \mathcal{A})$ a *measurable space*.

We occasionally need the following somewhat more general notions:

Definition 4.1 Let Ω be a set.

- (a) We say that $\mathcal{R} \subset 2^\Omega$ is a *ring* (of sets) if
 - (i) $\emptyset \in \mathcal{R}$;
 - (ii) $E \setminus F \in \mathcal{R}$ whenever $E, F \in \mathcal{R}$;
 - (iii) $E \cup F \in \mathcal{R}$ whenever $E, F \in \mathcal{R}$.
- (b) A ring $\mathcal{R} \subset 2^\Omega$ containing Ω is called an *algebra* (of sets).
- (c) We say that $\mathcal{S} \subset 2^\Omega$ is a *semiring* if
 - (i) $\emptyset \in \mathcal{S}$;
 - (ii) $E \cap F \in \mathcal{S}$ whenever $E, F \in \mathcal{S}$;
 - (iii) whenever $E, F \in \mathcal{S}$, $E \setminus F$ is the union of a finite number of disjoint sets belonging to $\mathcal{S}$.

Since $E \cap F = E \setminus (E \setminus F)$, a ring is closed with respect to the intersection of two (and by induction a finite number of) sets. In particular, a ring is a semiring. It is straightforward to show that if $\mathcal{S} \subset 2^\Omega$ is a semiring, then the set of finite unions of sets belonging to $\mathcal{S}$ is a ring. We denote this ring by $\mathcal{U}(\mathcal{S})$. Clearly $\mathcal{U}(\mathcal{S})$ is the smallest ring containing $\mathcal{S}$. For any subset $\mathcal{R} \subset 2^\Omega$, the smallest ring containing $\mathcal{R}$ exists; it is the intersection of all the rings containing $\mathcal{R}$ and called the *ring generated by $\mathcal{R}$* .

Remark 4.1 It is instructive to compare the structure $(\mathcal{P}(\mathcal{H}), \leq, ^\perp, 0, I)$ of the projection lattice $\mathcal{P}(\mathcal{H})$ of a Hilbert space $\mathcal{H}$ (see Theorem 2.14 and Remark 2.3) with the corresponding structure $(\mathcal{A}, \subset, ^c, \emptyset, \Omega)$ of a σ -algebra $\mathcal{A}$ of subsets of a (nonempty) set Ω . Both are lattices, and the respective mappings $^\perp$ and ^{c} are orthocomplementations (in the sense of Remark 2.3). Let L denote either of these orthocomplemented lattices, with, for instance, $a \wedge b$ and $a \vee b$ denoting the greatest lower bound (infimum) and the least upper bound (supremum) with respect to the relevant order. We say that the elements $a, b, c \in L$ form a *distributive triple* if the equalities

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \quad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

hold, together with the other four equalities obtained by cyclic permutation of a, b, c . It is obvious that any triple $X, Y, Z \in \mathcal{A}$ is distributive. This means that $\mathcal{A}$ is a *distributive lattice*. Distributive orthocomplemented lattices are generally called *Boolean algebras*. On the other hand, it is an easy exercise to show that in the projection lattice $\mathcal{P}(\mathcal{H})$ a triple P, Q, R is distributive if and only if the projections are pairwise commutative. $\mathcal{P}(\mathcal{H})$ is not a Boolean algebra, unless $\dim(\mathcal{H}) = 1$. By Theorem 2.14, $\mathcal{P}(\mathcal{H})$ is *complete* (i.e. every subset has supremum and infimum) whereas $\mathcal{A}$ is usually only σ -complete (i.e. every countable subset has supremum and infimum). $\triangleleft$

4.2 Measurable Functions

In this section we let Ω and Λ be sets and $\mathcal{A} \subset 2^\Omega$, $\mathcal{B} \subset 2^\Lambda$ σ -algebras. The measurability of a function relates to σ -algebras as continuity relates to topologies.

Definition 4.2 (a) We say that a map $f : \Omega \rightarrow \Lambda$ is $(\mathcal{A}, \mathcal{B})$ -measurable if $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$.
 (b) If, in the above, Λ is $\mathbb{R}^n$ and $\mathcal{B}$ is its Borel σ -algebra $\mathcal{B}(\mathbb{R}^n)$ then we call a $(\mathcal{A}, \mathcal{B})$ -measurable map simply $\mathcal{A}$ -measurable or just a *measurable function* if $\mathcal{A}$ is clear from the context. A similar remark applies when we have the set $\overline{\mathbb{R}}$ of the extended real numbers in place of $\mathbb{R}^n$. (We equip $\overline{\mathbb{R}}$ with its natural topology which makes it homeomorphic with the interval $[-1, 1] \subset \mathbb{C}$.)

We list without proof some basic properties of measurable functions. For the $(\mathcal{A}, \mathcal{B})$ -measurability of $f : \Omega \rightarrow \Lambda$ it suffices that for some class $\mathcal{F} \subset 2^\Lambda$ generating $\mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$ whenever $B \in \mathcal{F}$. As a consequence, the following holds.

Proposition 4.1 For a function $f : \Omega \rightarrow \overline{\mathbb{R}}$ the following conditions are equivalent:

- (a) f is $\mathcal{A}$ -measurable;
- (b) $f^{-1}([a, \infty]) \in \mathcal{A}$ for each $a \in \mathbb{R}$;
- (c) $f^{-1}((a, \infty]) \in \mathcal{A}$ for each $a \in \mathbb{R}$;
- (d) $f^{-1}([-\infty, a]) \in \mathcal{A}$ for each $a \in \mathbb{R}$;
- (e) $f^{-1}((-\infty, a)) \in \mathcal{A}$ for each $a \in \mathbb{R}$;
- (f) $f^{-1}(U) \in \mathcal{A}$ for each open set $U \subset \mathbb{R}$, $f^{-1}(\infty) \in \mathcal{A}$ and $f^{-1}(-\infty) \in \mathcal{A}$.

In particular, a function $f : \Omega \rightarrow \mathbb{R}$ is measurable if and only if it is measurable when regarded as a function into the extended real line $\overline{\mathbb{R}}$. The above proposition also easily implies that $\limsup f_k$ and $\liminf f_k$ for a sequence (f_k) of measurable functions are measurable. In particular, the limit of a pointwise convergent sequence of real or extended real valued measurable functions is measurable. The same applies to $\mathbb{R}^n$ -valued functions, since a function $f = (f_1, \dots, f_n) : \Omega \rightarrow \mathbb{R}^n$ is measurable if and only if the real valued functions $f_1, \dots, f_n$ are measurable. The measurable functions $f : \Omega \rightarrow \mathbb{R}^n$ form a vector space with respect to the pointwise operations. If $f : \Omega \rightarrow \mathbb{R}^n$ is a measurable function and g is a continuous $\mathbb{R}^m$ -valued function defined on its range, then the composite function $g \circ f$ is measurable.

4.3 Integration with Respect to a Positive Measure

We now assume that $(\Omega, \mathcal{A}, \mu)$ is a *measure space*, which means that $\mathcal{A} \subset 2^\Omega$ is a σ -algebra and $\mu : \mathcal{A} \rightarrow [0, \infty]$ is a *measure*, i.e. a nonnegative $\overline{\mathbb{R}}$ -valued set function which is σ -additive in the sense that $\mu(\emptyset) = 0$ and $\mu(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$ for any sequence (A_n) of disjoint sets belonging to $\mathcal{A}$. If here $\mu(\Omega) = 1$, μ is a *probability measure* and $(\Omega, \mathcal{A}, \mu)$ is a *probability space*.

We assume that G is a Banach space whose scalar field is $\mathbb{R}$ or $\mathbb{C}$; we use the unified notation $\mathbb{K}$ to stand for either one. We are going to define the *Bochner integral* of a G -valued function. (We use this term since the definition following [1] is equivalent to one originally given by Bochner.) Mostly in the sequel it would suffice to consider the special case where G is the scalar field $\mathbb{R}$ or $\mathbb{C}$ (in which case we obtain the *Lebesgue integral*), but we consider a general Banach space when this generality does not present any extra difficulty.

A *simple function* $f : \Omega \rightarrow G$ is a function which only takes a finite number of values, and the inverse image of each of these values belongs to $\mathcal{A}$. If, moreover, f vanishes outside a set of finite measure, f is an *integrable simple function*, and the *integral* of f is defined as

$$\int_{\Omega} f d\mu = \int_{\Omega} f(\omega) d\mu(\omega) = \sum_{i=1}^n \mu(E_i)x_i,$$

where $f(\Omega) = \{x_1, \dots, x_n\}$ and $E_i = f^{-1}(x_i)$. (We use the convention $\infty \cdot 0 = 0$ even if 0 is the zero element of a Banach space.) The integral of certain more general functions is defined by approximating them in a way described below.

We say that the sequence of functions $f_n : \Omega \rightarrow G$ *converges in (μ) -measure* to the function $f : \Omega \rightarrow G$ (denoted $f_n \rightarrow f(\mu)$) if for every $\varepsilon > 0$ we have

$$\lim_{n \rightarrow \infty} \mu^*(\{\omega \in \Omega \mid \|f_n(\omega) - f(\omega)\| \geq \varepsilon\}) = 0,$$

with the notation $\mu^*(E) = \inf\{\mu(B) \mid B \in \mathcal{A}, E \subset B\}$ for all $E \subset \Omega$.

Definition 4.3 A function $f : \Omega \rightarrow G$ is (μ) -*integrable* if there is a sequence of integrable simple functions $f_n : \Omega \rightarrow G$ such that

- (a) $f_n \rightarrow f(\mu)$ and
- (b) $\lim_{m, n \rightarrow \infty} \int_X |f_m - f_n| d\mu = 0$.

Above, we have denoted the function $\omega \mapsto \|g(\omega)\|$ by $|g|$ whenever g is a G -valued function defined on Ω ; we use this notation throughout this section. The condition (b) can be expressed by saying that (f_n) is an L^1 Cauchy sequence. It can be shown that in the situation of this definition the limit $\lim_{n \rightarrow \infty} \int_{\Omega} f_n$ exists and is independent of the choice of the sequence of functions satisfying these conditions (see [1]). This limit is called the *(Bochner) integral* of the μ -integrable function f (with respect to μ) and denoted by $\int_{\Omega} f d\mu = \int_{\Omega} f(\omega) d\mu(\omega)$. If f is a G -valued function defined on a set containing $E \subset \Omega$, then f is integrable over E if the function f_E satisfying $f_E(\omega) = f(\omega)$ for $\omega \in E$ and $f_E(\omega) = 0$ when $\omega \notin E$ is integrable. We then write

$$\int_E f d\mu = \int_E f(\omega) d\mu(\omega) = \int_{\Omega} f_E d\mu.$$

If $E = \Omega$, it is often omitted in the notation. In these and similar notations we may sometimes write $\mu(d\omega)$ instead of $d\mu(\omega)$.

The set of integrable functions $f : \Omega \rightarrow G$ is a vector space and the integral is a linear operation on it [1]. If f is integrable, then so is $|f|$. Moreover, $\|\int f d\mu\| \leq \int |f| d\mu$ and for any $\varepsilon > 0$

$$\mu^*(\{\omega \in \Omega \mid \|f(\omega)\| \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \int_X |f| d\mu.$$

In the definition of integrability there were no requirements of measurability, and it turns out that even a real-valued integrable function is not necessarily $\mathcal{A}$ -measurable. To elaborate this issue we next introduce some further notions.

We say that a set $E \subset \Omega$ is (μ) -negligible or a (μ) -null set if there is a set $B \in \mathcal{A}$ such that $E \subset B$ and $\mu(B) = 0$. An equivalent requirement is that $\mu^*(E) = 0$. If some condition depending on points $\omega \in \Omega$ is satisfied for all ω in the complement of some (μ) -null set, then it is said to be satisfied (μ) -almost everywhere, abbreviated (μ) -a.e., or for almost every ω .

A measure space $(\Omega, \mathcal{A}, \mu)$ is said to be *complete* if every null set belongs to $\mathcal{A}$. We denote

$$\overline{\mathcal{A}} = \{E \cup N \mid E \in \mathcal{A}, N \text{ is } \mu\text{-negligible}\}.$$

Then $\overline{\mathcal{A}}$ is a σ -algebra containing $\mathcal{A}$, and the formula $\overline{\mu}(A \cup N) = \mu(A)$ for all $A \in \mathcal{A}$ and all μ -null sets $N \subset \Omega$ defines an extension of μ , independently of the representation of $A \cup N$, to a measure defined on the σ -algebra $\overline{\mathcal{A}}$. Now $(\Omega, \overline{\mathcal{A}}, \overline{\mu})$ is a complete measure space. We say that it is the *completion* of the measure space $(\Omega, \mathcal{A}, \mu)$, the measure $\overline{\mu}$ is the *Lebesgue extension* of the measure μ , and $\overline{\mathcal{A}}$ is the *Lebesgue extension* of $\mathcal{A}$ (with respect to μ).

The importance of the completion of a measure space is apparent in the next result. The complex-valued case is contained in it as $\mathbb{C} = \mathbb{R}^2$.

Proposition 4.2 (a) *For a function $f : \Omega \rightarrow \mathbb{R}^n$ the following conditions are equivalent:*

- (i) *f is μ -integrable.*
 - (ii) *f is $\overline{\mathcal{A}}$ -measurable, and $|f|$ is μ -integrable.*
- (b) *If $g : \Omega \rightarrow [0, \infty)$ is μ -integrable and $f : \Omega \rightarrow \mathbb{R}^n$ is a $\mathcal{A}$ -measurable function such that $\|f(\omega)\| \leq g(\omega)$ a.e., then f is μ -integrable.*

Thus even if an $\mathbb{R}$ -valued integrable function were not $\mathcal{A}$ -measurable, it must be $\overline{\mathcal{A}}$ -measurable. If an integrable function is changed in a negligible set, its integrability stays intact, and the value of the integral does not change. For every integrable function $f : \Omega \rightarrow \mathbb{R}^n$ there is an $\mathcal{A}$ -measurable (integrable) function g satisfying $f(\omega) = g(\omega)$ a.e.

The integrability of a function $f : \Omega \rightarrow \overline{\mathbb{R}}$ is defined by requiring that f agrees a.e. with some integrable $\mathbb{R}$ -valued function g , and then we define $\int_{\Omega} f d\mu = \int_{\Omega} g d\mu$. Clearly this definition is independent of the choice of g .

The definition of the integral used here is suitable for Banach space valued functions and is independent of the order structure even when the range space

is $\mathbb{R}$. If a function $f : \Omega \rightarrow \mathbb{R}$ or more generally a function $f : \Omega \rightarrow \overline{\mathbb{R}}$ is $\mathcal{A}$ -measurable we, however, obtain the following result which shows the connection of our present approach to one used in many other presentations. For a function $f : \Omega \rightarrow \overline{\mathbb{R}}$ we define the functions $f^+ : \Omega \rightarrow [0, \infty]$ and $f^- : \Omega \rightarrow [0, \infty]$ by $f^+(\omega) = \max\{f(\omega), 0\}$ and $f^-(\omega) = \max\{-f(\omega), 0\}$. If f is measurable, then so are f^+ and f^- .

Proposition 4.3 (a) *An $\mathcal{A}$ -measurable function $f : \Omega \rightarrow [0, \infty]$ is integrable if and only if there is an increasing sequence of integrable simple functions $f_n : \Omega \rightarrow [0, \infty)$, such that $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega)$ for all $\omega \in \Omega$ and $\lim_{n \rightarrow \infty} \int f_n d\mu < \infty$. Then*

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$$

for every sequence (f_n) satisfying these conditions.

(b) *An $\mathcal{A}$ -measurable function $g : \Omega \rightarrow \overline{\mathbb{R}}$ is integrable if and only if g^+ and g^- are integrable. Then*

$$\int_{\Omega} g d\mu = \int_{\Omega} g^+ d\mu - \int_{\Omega} g^- d\mu.$$

We refrain from listing more results related to integration theory now. When they are needed we may refer directly to some outside literature.

4.4 The Hilbert Space $L^2(\Omega, \mathcal{A}, \mu)$

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. We use the notation L_0^2 for the set of all $\mathcal{A}$ -measurable functions $f : \Omega \rightarrow \mathbb{C}$ for which $|f|^2$ is integrable. We summarise below some basic facts related to this function space. Again, the omitted proof can be found e.g. in [1]. The set L_0^2 is a vector space, and the set

$$N = \{f \in L_0^2 \mid f(\omega) = 0 \text{ a.e. } \omega \in \Omega\}$$

is its vector subspace. We denote $[f] = f + N$. In the quotient space

$$L_0^2/N = \{[f] \mid f \in L_0^2\}$$

there is a well-defined inner product $\langle [g] \mid [f] \rangle$ such that

$$\langle [g] \mid [f] \rangle = \int_{\Omega} f \bar{g} d\mu.$$

With respect to the norm defined by this inner product, L_0^2/N is complete, i.e. a Hilbert space. We denote it by $L^2(\Omega, \mathcal{A}, \mu)$ or briefly by $L^2(\Omega, \mu)$ or by $L^2(\mu)$. If μ is the Lebesgue measure of $\mathbb{R}^n$, we may also write $L^2(\mathbb{R}^n)$ instead of $L^2(\mu)$.

Although strictly speaking the elements of $L^2(\mu)$ are equivalence classes of functions, in practice they are usually regarded as functions, and two functions are identified if they agree almost everywhere.

4.5 Complex Measures and Integration

In this section Ω is a set and $\mathcal{A} \subset 2^\Omega$ is a σ -algebra. A set function $\mu : \mathcal{A} \rightarrow \mathbb{C}$ is a *complex measure* if μ is σ -additive, i.e. $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

whenever (A_n) is a sequence of disjoint members of $\mathcal{A}$. We let $\text{ca}(\Omega, \mathcal{A})$ denote the set all complex measures defined on $\mathcal{A}$. If $\mu \in \text{ca}(\Omega, \mathcal{A})$, we use for each $A \in \mathcal{A}$ the notation $|\mu|(A)$ for the supremum of the sums $\sum_{k=1}^n |\mu(A_k)|$ over all finite $\mathcal{A}$ -partitions of A . (The sets $A_1, \dots, A_n \in \mathcal{A}$ form an $\mathcal{A}$ -*partition* of A if they are disjoint and their union is A .) The set function $A \mapsto |\mu|(A) \in [0, \infty]$ on $\mathcal{A}$ is called the *total variation* of μ .

Proposition 4.4 *Let $\mu \in \text{ca}(\Omega, \mathcal{A})$. Then*

- (a) $\sup_{A \in \mathcal{A}} |\mu(A)| < \infty$;
- (b) $|\mu(A)| \leq |\mu|(A) \leq 4 \sup \{|\mu(B)| \mid B \in \mathcal{A}, B \subset A\}$ whenever $A \in \mathcal{A}$;
- (c) *The set function $|\mu|$ is a finite positive measure.*

Proof See [1].

Remark 4.2 Clearly the set $\text{ca}(\Omega, \mathcal{A})$ is a vector space with the natural argumentwise operations. From the preceding proposition it follows that it is a normed space with respect to the supremum norm, and a standard argument shows that it is complete, i.e. a Banach space. However, unless otherwise stated, we consider it equipped with the (by the preceding proposition) equivalent Banach space norm $\mu \mapsto |\mu|(\Omega)$. We leave the details as an exercise. $\triangleleft$

Remark 4.3 Proposition 4.4 enables us to express any $\mu \in \text{ca}(\Omega, \mathcal{A})$ as a linear combination of four finite positive measures. We write $\mu_1 = \text{Re } \mu$, $\mu_2 = \text{Im } \mu$, and $\nu_1 = \frac{1}{2}(|\mu_1| + \mu_1)$, $\nu_2 = \frac{1}{2}(|\mu_1| - \mu_1)$, $\nu_3 = \frac{1}{2}(|\mu_2| + \mu_2)$, $\nu_4 = \frac{1}{2}(|\mu_2| - \mu_2)$. Then $\nu_1, \dots, \nu_4$ are positive measures and $\mu = \nu_1 - \nu_2 + i(\nu_3 - \nu_4)$. $\triangleleft$

For the rest of this section, let $\mu : \mathcal{A} \rightarrow \mathbb{C}$ be a complex measure and G a (complex) Banach space. We say that a function $f : \Omega \rightarrow G$ is *integrable* with respect to μ (or μ -integrable) if f is integrable with respect to the total variation measure $|\mu|$. If f is μ -integrable, we will see that it also has a well-defined integral which can be expressed in terms of positive measures. If f is an $\mathcal{A}$ -simple function, its μ -integral is defined in the same obvious way as for positive measures: $\int f d\mu = \sum_{i=1}^n \mu(E_i)x_i$ where $f(\Omega) = \{x_1, \dots, x_n\}$ and $E_i = f^{-1}(x_i)$. More generally, assume that $f : \Omega \rightarrow G$ is $|\mu|$ -integrable. Choose a sequence (f_n) of (automatically $|\mu|$ -integrable as $|\mu|(\Omega) < \infty$) simple functions such that $\lim_{m,n \rightarrow \infty} \int |f_m - f_n| d|\mu| = 0$ and $f_n \rightarrow f$ (μ). The same conditions remain valid if $|\mu|$ is replaced by any of the measures $\nu_1, \dots, \nu_4$ discussed in Remark 4.3, since clearly $\nu_j \leq |\mu|$. It follows that the limit

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \lim_{n \rightarrow \infty} \left[\int f_n d\nu_1 - \int f_n d\nu_2 + i \int f_n d\nu_3 - i \int f_n d\nu_4 \right]$$

exists and is independent of the choice of the sequence (f_n) . Using this notation we give the following definition.

Definition 4.4 The limit $\lim_{n \rightarrow \infty} \int f_n d\mu$ is called the *integral* of the μ -integrable function f with respect to μ . We denote

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int_{\Omega} f(\omega) d\mu(\omega) = \int_{\Omega} f d\mu = \int f d\mu.$$

In the same way as in the case of positive measures one defines the integrability and integral of a function over a subset of Ω .

Proposition 4.5 Let $f, g : \Omega \rightarrow G$ be μ -integrable functions. Then using the notation of Remark 4.3

- (a) $\int f d\mu = \int f d\nu_1 - \int f d\nu_2 + i \int f d\nu_3 - i \int f d\nu_4$;
- (b) for all $\alpha, \beta \in \mathbb{C}$, the function $\alpha f + \beta g$ is μ -integrable and

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu;$$

- (c) the function $|f|$ (i.e. the function $\omega \mapsto \|f(\omega)\|$) is $|\mu|$ -integrable and

$$\left\| \int f d\mu \right\| \leq \int |f| d|\mu|,$$

Proof Since (a) is clearly true for simple functions, the preceding discussion implies it in general (in particular, f is ν_j -integrable for all $j = 1, \dots, 4$). The integrability claims in (b) and (c) are clear as they hold for positive measures. The equation in (b) follows from (a), and the inequality in (c) via a limit procedure in the corresponding elementary inequality for simple functions. $\square$

As integration with respect to a complex measure is by definition reduced to the study of positive measures, many well-known results in the latter context are easily transferred to the complex case. As an example we mention the *dominated convergence theorem*:

Theorem 4.1 *Let $f_n : \Omega \rightarrow G$ be a μ -integrable function for each $n \in \mathbb{N}$, and let $g : \Omega \rightarrow [0, \infty)$ be a μ -integrable function such that for each $n \in \mathbb{N}$ $\|f_n(\omega)\| \leq g(\omega)$ for $|\mu|$ -a.e. $\omega \in \Omega$. If $f : \Omega \rightarrow G$ is a function such that $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega)$ $|\mu|$ -a.e., then f is μ -integrable and*

$$\lim_{n \rightarrow \infty} \int_{\Omega} \|f_n(\omega) - f(\omega)\| d|\mu| = 0, \quad \lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} f d\mu.$$

4.6 Positive Operator Measures

We begin the study of certain operator valued measures which will play a key role in the sequel. Throughout this section, Ω will be a set and $\mathcal{A}$ a σ -algebra of subsets of Ω . In most of our applications $(\Omega, \mathcal{A})$ will be $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ or more generally $(\Omega, \mathcal{B}(\Omega))$ where Ω is a closed or open subset of $\mathbb{R}^n$.

- Definition 4.5** (a) A map $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is *additive* if $E(X \cup Y) = E(X) + E(Y)$ for any two disjoint sets $X, Y \in \mathcal{A}$.
- (b) An additive map $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is said to be a *positive operator measure* (PO measure or POM for short) if $E(X) \geq 0$ for all $X \in \mathcal{A}$, and for any decreasing sequence of sets X_n in $\mathcal{A}$ such that $\bigcap_{n=1}^{\infty} X_n = \emptyset$, the zero operator is the greatest lower bound of the operators $E(X_n)$.
- (c) A POM $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is said to be *normalised* if $E(\Omega) = I$ (the identity operator on $\mathcal{H}$). A normalised POM is also called a *semispectral measure*.
- (d) If the range of a POM $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is contained in the projection lattice $\mathcal{P}(\mathcal{H})$, then E is called a *projection valued measure* (PV measure or PVM for short).
- (e) A normalised PV measure is called a *spectral measure*.

In the literature a POM is often called a *positive operator valued measure* (POVM).

Remark 4.4 (a) Since $E(\emptyset) = E(\emptyset) + E(\emptyset)$, any additive $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ satisfies $E(\emptyset) = 0$.

- (b) Suppose that $E(X) \geq 0$ for all $X \in \mathcal{A}$. For a decreasing sequence of positive operators $E(X_n)$ the greatest lower bound always exists, and it is the limit of the sequence $(E(X_n))$ with respect to both the weak and the strong operator topologies, i.e. in the sense of both weak and strong convergence. (See Theorem 2.7.) It follows by adapting the usual simple argument from scalar measure theory and applying the polarisation identity that E is a POM if and only if it is weakly, or equivalently strongly, σ -additive in the sense that $E(\emptyset) = 0$ and

$$\mathbf{E}(\cup_{n=1}^{\infty} X_n) = \sum_{n=1}^{\infty} \mathbf{E}(X_n)$$

for any sequence of disjoint sets $X_n \in \mathcal{A}$, the series converging weakly or, equivalently, strongly.

Another equivalent formulation is to require that for any $\varphi, \psi \in \mathcal{H}$ the set function

$$X \mapsto \mathbf{E}_{\psi, \varphi}(X) = \langle \psi | \mathbf{E}(X) \varphi \rangle$$

on $\mathcal{A}$ be a complex measure. Using again the polarisation identity, we may write

$$\mathbf{E}_{\psi, \varphi}(X) = \frac{1}{4} \sum_{n=0}^3 i^n \mathbf{E}_{\varphi + i^n \psi, \varphi + i^n \psi}(X)$$

which shows that it is enough to consider the cases where $\varphi = \psi$. If $\varphi = \psi$ and $\|\varphi\| = 1$, we also use the notation $p_{\varphi}^{\mathbf{E}}$ for the probability measure $\mathbf{E}_{\varphi, \varphi}$. Clearly, the family $\{p_{\varphi}^{\mathbf{E}} | \varphi \in \mathcal{H}, \|\varphi\| = 1\}$ also determines $\mathbf{E}$ uniquely.

- (c) We adopt the convention of using the above notation $\mathbf{E}_{\psi, \varphi}(X) = \langle \psi | \mathbf{E}(X) \varphi \rangle$ in connection with any set function $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$. $\triangleleft$

Proposition 4.6 *Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be additive.*

- (a) *Let $X, Y \in \mathcal{A}$ and $X \subset Y$. Then $\mathbf{E}(Y \setminus X) = \mathbf{E}(Y) - \mathbf{E}(X)$. If, in addition, $\mathbf{E}(Z) \geq 0$ for all $Z \in \mathcal{A}$, then $\mathbf{E}(X) \leq \mathbf{E}(Y)$.*
 (b) *For any $X, Y \in \mathcal{A}$*

$$\mathbf{E}(X \cup Y) + \mathbf{E}(X \cap Y) = \mathbf{E}(X) + \mathbf{E}(Y).$$

- (c) *If $\mathbf{E}(X) = \mathbf{E}(X)^*$ or $\|\mathbf{E}(X)\| \leq 1$ for all $X \in \mathcal{A}$, then the following conditions are equivalent:*

- (i) $\mathbf{E}(X \cap Y) = \mathbf{E}(X)\mathbf{E}(Y)$ for all $X, Y \in \mathcal{A}$;
 (ii) $\mathbf{E}(X) \in \mathcal{P}(\mathcal{H})$ for all $X \in \mathcal{A}$.

Proof (a) By additivity, $\mathbf{E}(Y) = \mathbf{E}(X) + \mathbf{E}(Y \setminus X)$.

(b) By additivity, we have

$$\mathbf{E}(X \cup Y) = \mathbf{E}(X \setminus Y) + \mathbf{E}(X \cap Y) + \mathbf{E}(Y \setminus X),$$

and adding $\mathbf{E}(X \cap Y)$ to both sides of this equation and again using additivity on the right hand side we obtain the claim.

- (c) If (i) holds, then $\mathbf{E}(X)^2 = \mathbf{E}(X \cap X) = \mathbf{E}(X)$. If $\mathbf{E}(X) = \mathbf{E}(X)^*$ or $\|\mathbf{E}(X)\| \leq 1$, then, by Theorem 2.10, $\mathbf{E}(X)$ is a projection. Conversely, if (ii) holds, then

$$\mathbf{E}(X \cap Y) \leq \mathbf{E}(X) \leq \mathbf{E}(X \cup Y)$$

by (a), and so $E(X)E(X \cap Y) = E(X \cap Y)$ and $E(X)E(X \cup Y) = E(X)$. Multiplying the equation of (b) from the left by $E(X)$ then yields $E(X) + E(X \cap Y) = E(X) + E(X)E(Y)$, implying (i). $\square$

Remark 4.5 Let $E : \mathcal{A} \rightarrow \mathcal{P}(\mathcal{H})$ be a spectral measure. By Proposition 4.6 one sees that E preserves the order ($X \subset Y \Rightarrow E(X) \leq E(Y)$), the orthocomplement ($E(Y^c) = E(Y)^\perp$), as well as the greatest lower bounds ($E(X \cap Y) = E(X) \wedge E(Y)$) and the least upper bounds ($E(X \cup Y) = E(X) \vee E(Y)$). This means that E is a *homomorphism* from the Boolean algebra $(\mathcal{A}, \subset, ^c, \emptyset, \Omega)$ to the complete orthocomplemented lattice $(\mathcal{P}(\mathcal{H}), \leq, ^\perp, 0, I)$. In addition, E is σ -additive, and thus a σ -homomorphism. The range $E(\mathcal{A})$ is clearly closed under the operations $^\perp$, $\wedge$, and $\vee$, and each triple $E(X), E(Y), E(Z) \in E(\mathcal{A})$ is distributive. This means that $E(\mathcal{A})$ is a Boolean subalgebra of $\mathcal{P}(\mathcal{H})$. Again, the σ -additivity of E implies that for any sequence $(E(X_i))$ of pairwise orthogonal elements in $E(\mathcal{A})$, their least upper bound $\vee E(X_i) = E(\cup X_i) \in \mathcal{A}$, that is, $E(\mathcal{A})$ is a Boolean sub- σ -algebra of $\mathcal{P}(\mathcal{H})$. $\triangleleft$

We prepare the definition of the support of a positive operator measure (in a special topological setting) with a lemma. A *Lindelöf space* is a topological space such that its each open cover has a countable subcover. A space is *strongly Lindelöf* if its every open subset is a Lindelöf space.

Lemma 4.1 *Let Ω be a strongly Lindelöf space and $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ a positive operator measure. There exists the smallest closed set $F \subset \Omega$ such that $E(\Omega \setminus F) = 0$.*

Proof The intersection of all closed subsets C of Ω satisfying $E(\Omega \setminus C) = 0$ is the required set F . To see this, note that if $(U_i)_{i \in \mathcal{I}}$ is any family of open subsets of Ω with $E(U_i) = 0$, then there is a countable subset $\{i_k \mid k = 1, 2, \dots\}$ of $\mathcal{I}$ such that $\cup_{i \in \mathcal{I}} U_i$ is already the union of the sets U_{i_k} . Thus by weak σ -(sub)additivity we have $E(\cup_{i \in \mathcal{I}} U_i) = 0$. $\square$

Definition 4.6 In the situation of the preceding lemma, the closed set F is called the *support* of the positive operator measure E and denoted $\text{supp}(E)$.

Remark 4.6 We take for granted that every separable metric space, in particular every subset of $\mathbb{R}^n$, is strongly Lindelöf. Thus if $\Omega \subset \mathbb{R}^n$, and $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ is a positive operator measure, then $\text{supp}(E)$ is defined as a relatively closed subset of Ω . If Ω is a closed subset of $\mathbb{R}^n$, then $\text{supp}(E)$ is closed even in $\mathbb{R}^n$. $\triangleleft$

We use the well-known Carathéodory–Hahn extension theorem in the following form: If $\mathcal{R} \subset 2^\Omega$ is a ring such that Ω is the union of some countable collection of sets in $\mathcal{R}$, then any σ -additive function $\mu : \mathcal{R} \rightarrow [0, \infty)$ has a unique extension to a (not necessarily finite) measure defined on the σ -algebra generated by $\mathcal{R}$. Here σ -additivity on $\mathcal{R}$ means the usual condition provided that the union of the sets also belongs to $\mathcal{R}$. The following lemma is often used to guarantee this hypothesis.

Lemma 4.2 *Let Ω be a Hausdorff topological space, $\mathcal{R} \subset 2^\Omega$ a ring and $\mu : \mathcal{R} \rightarrow [0, \infty)$ an additive set function. Suppose that for each $A \in \mathcal{R}$ and $\epsilon > 0$ there is a set $B \in \mathcal{R}$ with a compact set K such that $B \subset K \subset A$ and $\mu(A) - \mu(B) < \epsilon$. Then μ is σ -additive.*

Proof It is enough to show that $\lim_{n \rightarrow \infty} \mu(A_n) = 0$ whenever (A_n) is a decreasing sequence in $\mathcal{R}$ with $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Let $\epsilon > 0$ and choose for each $n \in \mathbb{N}$ $B_n \in \mathcal{R}$ and a compact set K_n such that $B_n \subset K_n \subset A_n$ and $\mu(A_n) - \mu(B_n) < \epsilon 2^{-n}$. We have $\bigcap_{n=1}^{\infty} K_n = \emptyset$, and since each $K_n \cap K_1$ is a closed subset of the compact set K_1 , there is some $p \in \mathbb{N}$ with $\bigcap_{n=1}^p K_n = \emptyset$. A straightforward argument (resembling one in the proof of Proposition 4.9(a)) now shows (exercise) that $\mu(A_n) \leq \epsilon \sum_{k=1}^n 2^{-k} < \epsilon$ for any $n \geq p$. $\square$

Let Ω be a set. A collection $\mathcal{R} \subset 2^\Omega$ is said to be *monotone class* if it is closed with respect to the unions of increasing sequences and the intersections of decreasing sequences. A σ -ring of subsets of Ω is a ring which is closed with respect to countable unions. We use the well-known *monotone class theorem* [3, p. 27] in the σ -ring version: The monotone class generated by a ring $\mathcal{R} \subset 2^\Omega$ is the same as the σ -ring generated by $\mathcal{R}$. (As usual, e.g. the monotone class generated by $\mathcal{R}$ is the intersection of all monotone classes containing $\mathcal{R}$.)

Lemma 4.3 *Let $\mathcal{R}$ be a ring of subsets of a set Ω such that Ω is the union of a countable collection of sets in $\mathcal{R}$. Let $\mathcal{A}$ be the σ -algebra generated by $\mathcal{R}$. Suppose that $E_0 : \mathcal{R} \rightarrow \mathcal{L}(\mathcal{H})$ is such that for every $\varphi \in \mathcal{H}$ the mapping $X \mapsto \langle \varphi | E_0(X) \varphi \rangle$ has a unique σ -additive extension $\lambda_\varphi : \mathcal{A} \rightarrow [0, \infty)$. Assume that there is a constant $M \in [0, \infty)$ such that $\lambda_\varphi(\Omega) \leq M \|\varphi\|^2$ for all $\varphi \in \mathcal{H}$. Then E_0 has a unique extension to a positive operator measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$. If $E_0(X)$ is a projection for every $X \in \mathcal{R}$, then so is $E(X)$ for every $X \in \mathcal{A}$.*

Proof Clearly there can be at most one such extension. For any $\xi, \eta \in \mathcal{H}$ and $X \in \mathcal{A}$ define

$$B_X(\xi, \eta) = \frac{1}{4} \sum_{n=0}^3 i^n \lambda_{\xi + i^n \eta}(X).$$

For any $X \in \mathcal{R}$ $B_X(\xi, \eta) = \langle \xi | E_0(X) \eta \rangle$ by the polarisation identity. Let $\mathcal{F}$ denote the set of those $X \in \mathcal{A}$ for which B_X is a sesquilinear form. Since each $\lambda_{\xi + i^n \eta}$ is a measure, $\mathcal{F}$ is a monotone class, and since it contains $\mathcal{R}$, it coincides with $\mathcal{A}$ (which in this case is also the σ -ring generated by $\mathcal{R}$). Fix $X \in \mathcal{A}$. For any $\xi, \eta \in \mathcal{H}$ with $\|\xi\| \leq 1, \|\eta\| \leq 1$, we have $|B_X(\xi, \eta)| \leq \frac{1}{4} \sum_{n=0}^3 \lambda_{\xi + i^n \eta}(X) \leq \frac{1}{4} \sum_{n=0}^3 M \|\xi + i^n \eta\|^2 \leq 4M$. Thus B_X is a bounded sesquilinear form, and so by Proposition 2.1 there is a unique operator $T_X \in \mathcal{L}(\mathcal{H})$ satisfying $B_X(\xi, \eta) = \langle \xi | T_X \eta \rangle$. Defining $E(X) = T_X$ for each $X \in \mathcal{A}$ we get the desired positive operator measure E . The last statement follows from the fact that the set of those $X \in \mathcal{A}$ for which $E(X)$ is a projection is a monotone class by Theorem 2.8(b). $\square$

Remark 4.7 The boundedness condition in the above result holds for example if $\Omega \in \mathcal{R}$, for then $\lambda_\varphi(X) \leq \lambda_\varphi(\Omega) = \langle \varphi | E(\Omega)\varphi \rangle \leq \|E(\Omega)\| \|\varphi\|^2$. $\triangleleft$

We now prove an extension result for positive operator measures.

Proposition 4.7 *Let Ω be a Hausdorff topological space and $\mathcal{R}$ a ring of subsets of Ω such that Ω is the union of some countable collection of sets in $\mathcal{R}$. Let $E_0 : \mathcal{R} \rightarrow \mathcal{L}(\mathcal{H})$ be an additive positive operator valued set function such that it is dominated by some positive operator D (i.e. $E_0(X) \leq D$ for all $X \in \mathcal{R}$) and for any $A \in \mathcal{R}$, $\varphi \in \mathcal{H}$ and $\epsilon > 0$ there is a set $B \in \mathcal{R}$ with a compact set K such that $B \subset K \subset A$ and*

$$\langle \varphi | E_0(A)\varphi \rangle - \langle \varphi | E_0(B)\varphi \rangle < \epsilon.$$

Then E_0 has a unique weakly σ -additive extension to a positive operator measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ where $\mathcal{A}$ is the σ -algebra generated by $\mathcal{R}$. If E_0 is projection valued, then so is E .

Proof For each $\varphi \in \mathcal{H}$ the set function $X \mapsto \langle \varphi | E_0(X)\varphi \rangle$ is σ -additive by Lemma 4.2, and so by the Carathéodory–Hahn theorem it has a unique extension to a measure $\lambda_\varphi : \mathcal{A} \rightarrow [0, \infty]$. Taking an increasing sequence (X_n) in $\mathcal{R}$ with union Ω , we see that $\lambda_\varphi(\Omega) = \lim_{n \rightarrow \infty} \langle \varphi | E_0(X_n)\varphi \rangle \leq \langle \varphi | D\varphi \rangle \leq \|D\| \|\varphi\|^2$. In particular, λ_φ is a finite measure. The claim now follows from Lemma 4.3. $\square$

4.7 Positive Operator Bimeasures

The setup of this section draws its historical motivation from the study of product spectral measures used e.g. in connection with the spectral representation of a normal operator, and this is an important application also for us. The techniques are, however, developed in much greater generality dictated by physical applications in later chapters. A convenient general framework is afforded by positive operator bimeasures, arising for example in the study of mutually commuting pairs of positive operator measures.

Suppose $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ are measurable spaces. The *product σ -algebra* of $\mathcal{A}_1$ and $\mathcal{A}_2$ is defined as the σ -algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ generated in $2^{\Omega_1 \times \Omega_2}$ by the set $\{X \times Y | X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}$.

Classically, already for spectral measures an extension problem arises. If $E_i : \mathcal{A}_i \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure for $i = 1, 2$, such that always $E_1(X)E_2(Y) = E_2(Y)E_1(X)$, then there is a unique additive projection valued set function E_0 on the algebra generated by the set $\{X \times Y | X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}$ such that $E_0(X \times Y) = E_1(X)E_2(Y)$ for all $X \in \mathcal{A}_1, Y \in \mathcal{A}_2$. It is, however, known that E_0 does not necessarily extend to a weakly σ -additive set function on the whole of $\mathcal{A}_1 \otimes \mathcal{A}_2$. We refer to [4, 5] for discussions with references. With suitable, typically topological, assumptions this obstacle can be circumvented. We give below one such approach which yields positive results also for commuting positive operator measures.

Lemma 4.4 *Let $\mathcal{R}_i$ be a collection of subsets of a set Ω_i , such that Ω_i is the union of a countable collection of members of $\mathcal{R}_i$ for $i = 1, 2$, and let $\mathcal{A}_i$ be the σ -algebra generated by $\mathcal{R}_i$. Let $\mathcal{A} \subset 2^{\Omega_1 \times \Omega_2}$ be the σ -algebra generated by $\{X \times Y \mid X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}$, and let $\mathcal{B} \subset 2^{\Omega_1 \times \Omega_2}$ be the σ -algebra generated by $\{X \times Y \mid X \in \mathcal{R}_1, Y \in \mathcal{R}_2\}$. Then $\mathcal{A} = \mathcal{B}$.*

Proof First fix $X \in \mathcal{R}_1$. The set $F_X = \{Y \in \mathcal{A}_2 \mid X \times Y \in \mathcal{B}\}$ is a σ -algebra containing $\mathcal{R}_2$. (For example, if $Y \in F_X$, then taking $\Omega_2 = \bigcup_{n=1}^{\infty} F_n$ with $F_n \in \mathcal{R}_2$ we have $X \times (\Omega_2 \setminus Y) = (X \times (\bigcup_{n=1}^{\infty} F_n)) \setminus (X \times Y) = (\bigcup_{n=1}^{\infty} X \times F_n) \setminus (X \times Y) \in \mathcal{B}$.) Therefore $\mathcal{A}_2 \subset F_X$. Thus $X \times Y \in \mathcal{B}$ for all $X \in \mathcal{R}_1$ and $Y \in \mathcal{A}_2$. A similar argument now shows that $X \times Y \in \mathcal{B}$ for all $X \in \mathcal{A}_1$ and $Y \in \mathcal{A}_2$. It follows that $\mathcal{B} = \mathcal{A}$, since the inclusion $\mathcal{B} \subset \mathcal{A}$ is obvious. $\square$

For convenience we use the following terminology.

Definition 4.7 (a) We say that the pair $(\Omega, \mathcal{R})$ has the *property (C)* if Ω is a Hausdorff topological space, $\mathcal{R}$ is a ring of subsets of Ω , and every σ -additive set function $\mu : \mathcal{R} \rightarrow [0, \infty)$ is such that for each $A \in \mathcal{R}$ and $\epsilon > 0$ there is a set $B \in \mathcal{R}$ with a compact set K such that $B \subset K \subset A$ and $\mu(A) - \mu(B) < \epsilon$.
 (b) Let $(\Omega, \mathcal{A})$ be a measurable space. We say that it has the *property (D)* if Ω is a Hausdorff space and there is a ring $\mathcal{R}$ generating the σ -algebra $\mathcal{A}$ such that Ω is the union of a countable collection of members of $\mathcal{R}$ and the pair $(\Omega, \mathcal{R})$ has the property (C).

We use the notion of a (scalar) *bimeasure* $\beta : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow [0, \infty)$. This simply means that β is *separately σ -additive*: the functions $\beta(X, \cdot)$ and $\beta(\cdot, Y)$ are σ -additive for all $X \in \mathcal{A}_1, Y \in \mathcal{A}_2$.

Lemma 4.5 *Suppose that $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces having the property (D). Let $\beta : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow [0, \infty)$ be a bimeasure. Then there is a unique measure $\mu : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow [0, \infty)$ such that $\mu(X \times Y) = \beta(X, Y)$ for all $X \in \mathcal{A}_1, Y \in \mathcal{A}_2$.*

Proof For $i = 1, 2$, we choose a ring $\mathcal{R}_i \subset \mathcal{A}_i$ having the properties required in Definition 4.7. Let $\mathcal{R}$ be the ring generated by the set $\{X \times Y \mid X \in \mathcal{R}_1, Y \in \mathcal{R}_2\}$. Each element of $\mathcal{R}$ is a finite union of sets of the type $X \times Y$ where $X \in \mathcal{R}_1$ and $Y \in \mathcal{R}_2$, and a routine argument shows that there is a unique additive set function $\mu_0 : \mathcal{R} \rightarrow [0, \infty)$ such that $\mu_0(X \times Y) = \beta(X, Y)$ for all $X \in \mathcal{R}_1, Y \in \mathcal{R}_2$. To see that μ_0 satisfies the kind of approximation condition assumed in Lemma 4.2, it is enough to consider one set of the type $X \times Y$ where $X \in \mathcal{R}_1, Y \in \mathcal{R}_2$. Let $\epsilon > 0$. Since the function $\beta(X, \cdot) : \mathcal{R}_2 \rightarrow [0, \infty)$ is σ -additive, by our property (C) assumption there is a set $W \in \mathcal{R}_2$ with a compact set L such that $W \subset L \subset Y$ and $\beta(X, W) \geq \beta(X, Y) - \frac{\epsilon}{2}$. Similarly, there is a set $Z \in \mathcal{R}_1$ with a compact set K such that $Z \subset K \subset X$ and $\beta(Z, W) \geq \beta(X, W) - \frac{\epsilon}{2}$. Thus $\beta(Z, W) \geq \beta(X, Y) - \epsilon$, where $Z \times W \subset K \times L \subset X \times Y$ and $K \times L$ is compact. By Lemma 4.2 μ_0 is σ -additive, and so by the Carathéodory–Hahn theorem it has a unique σ -additive extension μ to the σ -algebra generated by $\mathcal{R}$. But by Lemma 4.4 this σ -algebra equals $\mathcal{A}_1 \otimes \mathcal{A}_2$. We still need to show that μ is finite and $\mu(X \times Y) = \beta(X, Y)$ for

all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. Fixing first $X \in \mathcal{R}_1$ we can apply the Carathéodory–Hahn theorem (to the function $Y \mapsto \beta(X, Y)$ on the ring $\mathcal{R}_2$) and then repeat the argument for the first variable X with a fixed $Y \in \mathcal{A}$, thus getting the claim (in particular, $\mu(\Omega_1 \times \Omega_2) = \beta(\Omega_1, \Omega_2) < \infty$). $\square$

Remark 4.8 The above proof can be used to show that if $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces having the property (D), then $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ also has the property (D). Of course on $\Omega_1 \times \Omega_2$ we take the product topology, and $\mathcal{R}$ is as in the above proof. We leave the simple proof of the property (C) of the pair $(\Omega_1 \times \Omega_2, \mathcal{R})$ as an exercise. $\triangleleft$

Definition 4.8 Let $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, be measurable spaces. We say that a mapping $B : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ is a *positive operator bimeasure* if for every $X \in \mathcal{A}_1$ the set function $Y \mapsto B(X, Y)$ on $\mathcal{A}_2$ is a positive operator measure and similarly for every $Y \in \mathcal{A}_2$ the set function $X \mapsto B(X, Y)$ on $\mathcal{A}_1$ is a positive operator measure.

Theorem 4.2 Suppose that $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces having the property (D). If $B : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ is a positive operator bimeasure, then there is a unique positive operator measure $\mathbf{E} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $B(X, Y) = \mathbf{E}(X \times Y)$ for all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. If B is projection valued, then so is $\mathbf{E}$.

Proof For every $\varphi \in \mathcal{H}$ we define $\beta_\varphi(X, Y) = \langle \varphi | B(X, Y) \varphi \rangle$ whenever $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. Each β_φ is a bimeasure, and so by Lemma 4.5 there is a unique measure $\lambda_\varphi : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow [0, \infty)$ satisfying $\lambda_\varphi(X \times Y) = \beta_\varphi(X, Y)$ for all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. Let $\mathcal{R}$ be the ring generated by $\{X \times Y | X \in \mathcal{A}_1, Y \in \mathcal{A}_2\}$. There is a unique additive set function $\mathbf{E}_0 : \mathcal{R} \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{E}_0(X \times Y) = B(X, Y)$ for $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. As $\lambda_\varphi(\Omega_1 \times \Omega_2) = \langle \varphi | B(\Omega_1, \Omega_2) \varphi \rangle \leq \|B(\Omega_1, \Omega_2)\| \|\varphi\|^2$, Lemma 4.3 yields our claims. $\square$

Example 4.1 Assume that $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces having the property (D). Suppose that $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{E}_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ are commuting positive operator measures, i.e. $\mathbf{E}_1(X)\mathbf{E}_2(Y) = \mathbf{E}_2(Y)\mathbf{E}_1(X)$ whenever $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. Then also $\mathbf{E}_2(Y)^{\frac{1}{2}}$ commutes with $\mathbf{E}_1(X)$ (exercise), and so $\mathbf{E}_1(X)\mathbf{E}_2(Y) \geq 0$ since $\langle \varphi | \mathbf{E}_1(X)\mathbf{E}_2(Y) \varphi \rangle = \langle \mathbf{E}_2(Y)^{\frac{1}{2}} \varphi | \mathbf{E}_1(X)\mathbf{E}_2(Y)^{\frac{1}{2}} \varphi \rangle \geq 0$. Thus by Theorem 2.8(a) the map $(X, Y) \mapsto \mathbf{E}_1(X)\mathbf{E}_2(Y)$ is a positive operator bimeasure, and by Theorem 4.2 it determines a unique positive operator measure $\mathbf{E} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $\mathbf{E}(X \times Y) = \mathbf{E}_1(X)\mathbf{E}_2(Y)$ for all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. In the important special case where $\mathbf{E}_1$ and $\mathbf{E}_2$ are projection valued or, more specifically, commuting spectral measures (with the obvious definition), then by Theorem 4.2 so is $\mathbf{E}$. $\triangleleft$

In our physical applications a somewhat related result, Proposition 4.8 below, will also be useful. We prepare its proof with a lemma.

Lemma 4.6 (a) Suppose $P \in \mathcal{L}(\mathcal{H})$ is a projection and $A \in \mathcal{L}_s(\mathcal{H})$ satisfies $0 \leq A \leq I$. If $A \leq P$, then $A = PA = AP$. If $PA = AP$, then $PA \leq P$ and $PA \leq A$. If any positive operator B satisfies $B \leq P$ and $B \leq A$, then $B \leq PA$.

- (b) If $(\Omega, \mathcal{A})$ is a measurable space and $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a semispectral measure, then any projection in the range of E commutes with every operator in the range of E .

Proof (a) The first claim follows Proposition 2.5: $PA = A = A^* = AP$. The second claim: Proposition 2.5 shows that $PA \leq P$, and $PA \leq A$ since $\langle \xi | PA\xi \rangle = \|PA^{\frac{1}{2}}\xi\|^2 \leq \|A^{\frac{1}{2}}\xi\|^2 = \langle \xi | A\xi \rangle$. Finally suppose that $0 \leq B \leq P$ and $B \leq A$. Using the first claim we get $\langle \xi | B\xi \rangle = \langle \xi | PB\xi \rangle = \langle P\xi | BP\xi \rangle \leq \langle P\xi | AP\xi \rangle = \langle \xi | PA\xi \rangle$.

- (b) Assume $P = E(X)$ is a projection and $Y \in \mathcal{A}$. Then $E(Y) = E(Y \setminus X) + E(X \cap Y)$. Here $E(Y \setminus X) \leq E(\Omega \setminus X) = I - P$ and $E(X \cap Y) \leq E(X) = P$. It follows from the first claim in (a) that each of the operators $E(Y \setminus X)$ and $E(X \cap Y)$ commutes with P , and hence so does their sum $E(Y)$. $\square$

Proposition 4.8 Suppose that $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, are measurable spaces. Let $E : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure. Define $E_1(X) = E(X \times \Omega_2)$ and $E_2(Y) = E(\Omega_1 \times Y)$ for all $X \in \mathcal{A}_1$ and $Y \in \mathcal{A}_2$, i.e. E_1 and E_2 are the marginal semispectral measures of E . If E_1 or E_2 is projection valued, then $E_1(X)E_2(Y) = E_2(Y)E_1(X) = E(X \times Y)$ for all $X \in \mathcal{A}_1$ and $Y \in \mathcal{A}_2$.

Proof Assume that e.g. E_1 is projection valued. Then every $E_1(X) = E(X \times \Omega_2)$ commutes with every $E_2(Y) = E(\Omega_1 \times Y)$ by Lemma 4.6(b). Since $E(X \times Y) \leq E(X \times \Omega_2) = E_1(X)$ and similarly $E(X \times Y) \leq E_2(Y)$, Lemma 4.6(a) shows that $E(X \times Y) \leq E_1(X)E_2(Y)$. We apply this principle four times and get

$$\begin{aligned} I &= E(\Omega_1 \times \Omega_2) = E(X \times Y) + E(X \times Y^c) + E(X^c \times Y) + E(X^c \times Y^c) \\ &\leq E_1(X)E_2(Y) + E_1(X)E_2(Y^c) + E_1(X^c)E_2(Y) + E_1(X^c)E_2(Y^c) \\ &= (E_1(X) + E_1(X^c))(E_2(Y) + E_2(Y^c)) = E_1(\Omega_1)E_2(\Omega_2) = I. \end{aligned}$$

For each of the four terms we have the inequality in the same direction which implies that there must be an equality everywhere. In particular, $E(X \times Y) = E_1(X)E_2(Y)$. $\square$

We now consider Borel measures on certain spaces which in our physical applications will be of the greatest (even sole) interest.

Proposition 4.9 Suppose that Ω is a closed or open subset of $\mathbb{R}^n$.

- (a) Any measure $\mu : \mathcal{B}(\Omega) \rightarrow [0, \infty)$ is regular in the sense that, for any $X \in \mathcal{B}(\Omega)$, $\mu(X) = \sup\{\mu(K) \mid K \subset X \text{ compact}\} = \inf\{\mu(G \cap \Omega) \mid X \subset G \cap \Omega, G \subset \mathbb{R}^n \text{ open}\}$.
- (b) The measurable space $(\Omega, \mathcal{B}(\Omega))$ has the property (D).

Proof (a) Let $\mathcal{R}$ be the ring generated by the intersections of bounded intervals with Ω . Any interval in $\mathbb{R}^n$ is the union of an increasing sequence of compact intervals. It follows immediately that if Ω is a closed subset of $\mathbb{R}^n$, then

- any member of $\mathcal{R}$ is the union of an increasing sequence of compact members of $\mathbb{R}$. The same is true if Ω is open as can easily be seen by first expressing Ω as the union of an increasing sequence of finite unions of compact intervals. Now consider a measure $\mu : \mathcal{B}(\Omega) \rightarrow [0, \infty)$ and denote by $\mathcal{A}$ the class of those sets $X \in \mathcal{B}(\Omega)$ for which $\mu(X) = \sup\{\mu(K) \mid K \subset X \text{ compact}\}$. We have seen that $\mathcal{R} \subset \mathcal{A}$ (since μ commutes with monotone limits). We show that $\mathcal{A}$ is a monotone class. If (A_k) is an increasing sequence in $\mathcal{A}$ with union A , then for any $\epsilon > 0$ we can find some $k \in \mathbb{N}$ such that $\mu(A) - \mu(A_k) < \epsilon$, and then choosing a compact set $K \subset A_k$ satisfying $\mu(A_k) - \mu(K) < \epsilon - \mu(A) + \mu(A_k)$ we have $\mu(A) - \mu(K) < \epsilon$. Thus $A \in \mathcal{A}$. Now let (B_k) be a decreasing sequence in $\mathcal{A}$ with intersection B . Let $\epsilon > 0$. For each $k \in \mathbb{N}$ we choose a compact set $K_k \subset B_k$ with $\mu(B_k) - \mu(K_k) < \epsilon 2^{-k}$. The set $K = \bigcap_{k=1}^{\infty} K_k$ is compact and $\mu(B) - \mu(K) = \mu(\bigcap_{k=1}^{\infty} B_k \setminus \bigcap_{k=1}^{\infty} K_k) = \mu((\bigcap_{k=1}^{\infty} B_k) \cap (\bigcup_{k=1}^{\infty} (\Omega \setminus K_k))) \leq \mu(\bigcup_{k=1}^{\infty} B_k \setminus K_k) \leq \sum_{k=1}^{\infty} \mu(B_k \setminus K_k) < \epsilon$. Thus $B \in \mathcal{A}$. By the monotone class theorem the σ -ring (actually, σ -algebra as Ω is a countable union of members of $\mathcal{R}$) generated by $\mathcal{R}$ is contained in $\mathcal{A}$, and so $\mathcal{B}(\Omega) = \mathcal{A}$. Using this result on inner compact approximation to the complement $\Omega \setminus X$ of a set $X \in \mathcal{B}(\Omega)$ we see that $\mu(X) = \inf\{\mu(U) \mid X \subset U \subset \Omega, U \text{ relatively open}\}$.
- (b) This is an immediate consequence of (a): we may take $\mathcal{B}(\Omega)$ itself for the ring $\mathcal{R}$ required in Definition 4.7(b). (The full force of (a) could be dispensed with by taking $\mathcal{R}$ as in the above proof.) $\square$

The preceding result could be proved in much more general situations, but the present generality, allowing for a simple proof, is sufficient for our needs. We conclude this section with a result which will be of interest later in our applications of some of the above material.

Proposition 4.10 *Suppose the topology in each of the topological spaces Ω_1 and Ω_2 has a countable base. For their Borel σ -algebras we have the formula*

$$\mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2) = \mathcal{B}(\Omega_1 \times \Omega_2).$$

Proof Let $\mathcal{B}_i$, $i = 1, 2$, be a countable base for the topology of Ω_i . Then $\{A_1 \times A_2 \mid A_i \in \mathcal{B}_i, i = 1, 2\}$ is a countable base for the product topology of $\Omega_1 \times \Omega_2$. Let $U \subset \Omega_1 \times \Omega_2$ be an open set. We may express U as the union of a countable family (B_j) where $B_j = A_j^1 \times A_j^2$ with some $A_j^i \in \mathcal{B}_i$, $i = 1, 2$. It follows that $U = \bigcup_{j=1}^{\infty} A_j^1 \times A_j^2 \in \mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2)$. Thus $\mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2) \supset \mathcal{B}(\Omega_1 \times \Omega_2)$. The reverse inclusion follows from the fact that the projection maps $\text{pr}_i : \Omega_1 \times \Omega_2 \rightarrow \Omega_i$, $i = 1, 2$, are continuous, implying that $A_1 \times A_2 = \text{pr}_1^{-1}(A_1) \cap \text{pr}_2^{-1}(A_2) \in \mathcal{B}(\Omega_1 \times \Omega_2)$ for all $A_i \in \mathcal{B}(\Omega_i)$, $i = 1, 2$. $\square$

4.8 Integration of Bounded Functions with Respect to a Positive Operator Measure

The range of an operator measure $\mathbf{E}$ is a collection of bounded operators $\mathbf{E}(X)$, $X \in \mathcal{A}$. The theory of integration of measurable complex functions f with respect to a positive operator measure $\mathbf{E}$ allows one to construct a wider class of operators from $\mathbf{E}$. In this section we develop this theory for bounded measurable functions.

Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be an additive set function and $f = \sum_{j=1}^n \alpha_j \chi_{X_j}$ an $\mathcal{A}$ -simple function, i.e. $X_j \in \mathcal{A}$, $\alpha_j \in \mathbb{C}$ for each $j = 1, \dots, n$, and χ_{X_j} is the characteristic function (or indicator) of X_j . Then the *integral* of f with respect to $\mathbf{E}$ is defined by the formula

$$\int_{\Omega} f \, d\mathbf{E} = \sum_{j=1}^n \alpha_j \mathbf{E}(X_j).$$

A routine check shows that the result is independent of the way f is represented as a linear combination of characteristic functions. Moreover, the integral is readily seen to define a linear map in the vector space of $\mathcal{A}$ -simple functions. If $\mathcal{A}$ is clear from the context, we usually suppress it and talk just of simple functions.

We let $\mathcal{F}_{\mathcal{A}}(\Omega, \mathbb{C})$ or simply $\mathcal{F}_{\mathcal{A}}$ denote the space of all bounded $\mathcal{A}$ -measurable functions $f : \Omega \rightarrow \mathbb{C}$. We equip $\mathcal{F}_{\mathcal{A}}$ with the supremum norm $\|f\| = \sup_{\omega \in \Omega} |f(\omega)|$. Then $\mathcal{F}_{\mathcal{A}}$ is a Banach space, being a closed subspace of the Banach space of all bounded complex functions of Ω with respect to the supremum norm.

Lemma 4.7 *If $f \in \mathcal{F}_{\mathcal{A}}$, there exists a sequence of simple functions f_n converging uniformly to f on Ω and such that $|f_n(\omega)| \leq |f(\omega)|$ for all $\omega \in \Omega$. In particular, the space of simple functions is dense in $\mathcal{F}_{\mathcal{A}}$.*

Proof For each $n \in \mathbb{N}$, we express $\{z \in \mathbb{C} \mid |z| \leq \|f\|\}$ as the union of a finite number of disjoint Borel sets $B_1, \dots, B_k$ having diameter at most $\frac{1}{n}$. If $1 \leq j \leq k$, using the compactness of the closure $\overline{B_j}$ we choose $z_j \in \overline{B_j}$ such that $|z_j| \leq |z|$ whenever $z \in B_j$. Then $f_n = \sum_{j=1}^k z_j \chi_{f^{-1}(B_j)}$ is a simple function satisfying $\|f_n - f\| \leq \frac{1}{n}$ and $|f_n(\omega)| \leq |f(\omega)|$ for all $\omega \in \Omega$. $\square$

Proposition 4.11 *Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a positive operator measure. For any simple function $f : \Omega \rightarrow \mathbb{C}$,*

$$\left\| \int f \, d\mathbf{E} \right\| \leq 2 \|\mathbf{E}(\Omega)\| \sup_{\omega \in \Omega} |f(\omega)|.$$

The mapping $f \mapsto \int f \, d\mathbf{E}$ on the space of simple functions on Ω can be uniquely extended to a bounded linear map $L : \mathcal{F}_{\mathcal{A}} \rightarrow \mathcal{L}(\mathcal{H})$.

Proof Suppose first that f is a real valued simple function. If $\varphi \in \mathcal{H}$ and $\|\varphi\| \leq 1$, we have

$$\left| \left\langle \varphi \left| \left(\int f d\mathbf{E} \right) \varphi \right\rangle \right| = \left| \int f d\mathbf{E}_{\varphi, \varphi} \right| \leq \|f\| \mathbf{E}_{\varphi, \varphi}(\Omega) = \|f\| \langle \varphi | \mathbf{E}(\Omega) \varphi \rangle,$$

and since $\int f d\mathbf{E} \in \mathcal{L}_s(\mathcal{H})$, in view of Proposition 2.2 this implies

$\|\int f d\mathbf{E}\| \leq \|f\| \|\mathbf{E}(\Omega)\|$. In the general case, $f = f_1 + if_2$ with real valued simple functions f_1, f_2 , and we get

$$\begin{aligned} \left\| \int f d\mathbf{E} \right\| &\leq \left\| \int f_1 d\mathbf{E} \right\| + \left\| \int f_2 d\mathbf{E} \right\| \\ &\leq (\|f_1\| + \|f_2\|) \|\mathbf{E}(\Omega)\| \leq 2 \|f\| \|\mathbf{E}(\Omega)\|. \end{aligned}$$

The latter claim follows from Lemma 4.7 and the completeness of $\mathcal{L}(\mathcal{H})$. $\square$

Definition 4.9 In the situation of the preceding proposition the operator $L(f) \in \mathcal{L}(\mathcal{H})$, or more precisely $L(f, \mathbf{E})$, is called the *integral* of the function $f \in \mathcal{F}_A$ with respect to the positive operator measure $\mathbf{E}$, and we denote

$$L(f, \mathbf{E}) = \int f d\mathbf{E} = \int_{\Omega} f d\mathbf{E} = \int f(\omega) d\mathbf{E}(\omega) = \int_{\Omega} f(\omega) d\mathbf{E}(\omega) = \mathbf{E}[f].$$

When obvious from the context, we may suppress $\mathbf{E}$ from the notation $L(f, \mathbf{E})$.

Recall that we may use the notation $\mathbf{E}_{\psi, \varphi}(X) = \langle \psi | \mathbf{E}(X) \varphi \rangle$ for any set function $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$.

Proposition 4.12 Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a positive operator measure and $f \in \mathcal{F}_A$.

- (a) $\int f d\mathbf{E}_{\psi, \varphi} = \langle \psi | \left(\int f d\mathbf{E} \right) \varphi \rangle$ whenever $\varphi, \psi \in \mathcal{H}$, so that in particular $\int f d\mathbf{E} \geq 0$ if $f \geq 0$.
- (b) $\int \bar{f} d\mathbf{E} = \left(\int f d\mathbf{E} \right)^*$.
- (c) $\left\| \int f d\mathbf{E} \right\| \leq 2 \|\mathbf{E}(\Omega)\| \sup_{\omega \in \Omega} |f(\omega)|$.
- (d) If (f_n) is a sequence in $\mathcal{F}_A$ converging uniformly to a function f , then $\lim_{n \rightarrow \infty} \left\| \int f_n d\mathbf{E} - \int f d\mathbf{E} \right\| = 0$.
- (e) If (f_n) is a sequence in $\mathcal{F}_A$ converging pointwise to a function f and satisfying $\sup_{n \in \mathbb{N}} \|f_n\| < \infty$, then $w\text{-}\lim_{n \rightarrow \infty} \int f_n d\mathbf{E} = \int f d\mathbf{E}$.

Proof (a) The equality clearly holds for simple functions f , and since both sides depend continuously on the integrand, the claim follows from Proposition 4.11.

(b) Again the equality is seen by approximating f uniformly by simple functions.

(c) The inequality follows from Definition 4.9 and Proposition 4.11.

(d) This is an immediate consequence of (c) and the (obvious) linearity of the integral.

- (e) Each $\mathbf{E}_{\varphi, \varphi}$ is a finite positive measure, and so by (a) and the dominated convergence theorem we get

$$\left\langle \varphi \left| \left(\int f_n d\mathbf{E} \right) \varphi \right\rangle = \int f_n d\mathbf{E}_{\varphi, \varphi} \rightarrow \int f d\mathbf{E}_{\varphi, \varphi} = \left\langle \varphi \left| \left(\int f d\mathbf{E} \right) \varphi \right\rangle. \quad \square$$

Proposition 4.13 *Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a positive measure.*

- (a) *For an operator $T \in \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:*
- (i) $T\mathbf{E}(X) = \mathbf{E}(X)T$ for all $X \in \mathcal{A}$;
 - (ii) $TL(f, \mathbf{E}) = L(f, \mathbf{E})T$ for all $f \in \mathcal{F}_{\mathcal{A}}$.
- (b) *The following conditions are equivalent:*
- (i) $L(fg, \mathbf{E}) = L(f, \mathbf{E})L(g, \mathbf{E})$ for all $f, g \in \mathcal{F}_{\mathcal{A}}$;
 - (ii) *For every $X \in \mathcal{A}$ the operator $\mathbf{E}(X)$ is a projection.*
- (c) *If $\mathbf{E}(X)$ is a projection for all $X \in \mathcal{A}$, then for any $\psi, \varphi \in \mathcal{H}$ the total variation of the complex measure $\mathbf{E}_{\psi, \varphi}$ satisfies $|\mathbf{E}_{\psi, \varphi}|(\Omega) \leq \|\psi\| \|\varphi\|$, and $\|L(f, \mathbf{E})\| \leq \sup_{x \in \Omega} |f(x)|$ for all $f \in \mathcal{F}_{\mathcal{A}}$.*

Proof (a) Clearly (ii) implies (i) since $\chi_X \in \mathcal{F}_{\mathcal{A}}$. Assume now (i). Then (ii) evidently holds for simple functions, and the general case follows from Proposition 4.12(d), since both sides depend continuously (with respect to the operator norm) on f .

(b) If (i) holds, then in particular $\mathbf{E}(X) = L(\chi_X^2, \mathbf{E}) = L(\chi_X, \mathbf{E})L(\chi_X, \mathbf{E}) = \mathbf{E}(X)^2$, and since $\mathbf{E}(X)$ is selfadjoint, $\mathbf{E}(X) \in \mathcal{P}(\mathcal{H})$. Now assume (ii). If $f, g \in \mathcal{F}_{\mathcal{A}}$, by Lemma 4.7 we can choose two sequences (f_n) and (g_n) of simple functions such that $f_n \rightarrow f$ and $g_n \rightarrow g$ uniformly. Then $L(f)L(g) = \lim_{n \rightarrow \infty} L(f_n)L(g_n) = \lim_{n \rightarrow \infty} L(f_n g_n) = L(fg)$ by Proposition 4.12(d), since it easily follows from (ii) in view of Proposition 4.6(c) that (i) holds for simple functions.

(c) If the sets $A_1, \dots, A_n \in \mathcal{A}$ are disjoint and $\cup_{k=1}^n A_k = \Omega$, we have

$$\begin{aligned} \sum_{k=1}^n |\mathbf{E}_{\psi, \varphi}(A_k)| &= \sum_{k=1}^n |\langle \psi | \mathbf{E}(A_k) \varphi \rangle| = \sum_{k=1}^n |\langle \mathbf{E}(A_k) \psi | \mathbf{E}(A_k) \varphi \rangle| \\ &\leq \sum_{k=1}^n \|\mathbf{E}(A_k) \psi\| \|\mathbf{E}(A_k) \varphi\| \\ &\leq \sqrt{\sum_{k=1}^n \|\mathbf{E}(A_k) \psi\|^2 \sum_{k=1}^n \|\mathbf{E}(A_k) \varphi\|^2} \leq \|\psi\| \|\varphi\| \end{aligned}$$

by the Cauchy–Schwarz inequality and the Pythagorean theorem, since the projections $\mathbf{E}(A_k)$ are by Proposition 4.6(c) mutually orthogonal and their sum is a projection. Thus $|\mathbf{E}_{\psi, \varphi}|(\Omega) \leq \|\psi\| \|\varphi\|$. Using Proposition 4.5 we thus find that

$$\left| \langle \psi | \left(\int f d\mathbf{E} \right) \varphi \rangle \right| = \left| \int f d\mathbf{E}_{\psi, \varphi} \right| \leq \|f\| |\mathbf{E}_{\psi, \varphi}|(\Omega) \leq \|f\|$$

whenever $\|\psi\| \leq 1$, $\|\varphi\| \leq 1$, implying the latter claim. $\square$

We next prove an elementary transformation result for operator integrals.

Proposition 4.14 *Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a positive operator measure, Λ a set, and $\mathcal{B} \subset 2^\Lambda$ a σ -algebra. Let $\alpha : \Omega \rightarrow \Lambda$ be a measurable function, that is, $\alpha^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$. Then the map $\mathbf{E}^\alpha : \mathcal{B} \rightarrow \mathcal{L}(\mathcal{H})$ defined by $\mathbf{E}^\alpha(B) = \mathbf{E}(\alpha^{-1}(B))$ is a positive operator measure, and*

$$\int_\Lambda f d\mathbf{E}^\alpha = \int_\Omega (f \circ \alpha) d\mathbf{E},$$

for all $f \in \mathcal{F}_\mathcal{B}$. If $\mathbf{E}$ is a spectral measure, then $\mathbf{E}^\alpha$ is a spectral measure, too.

Proof It is immediately seen that $\mathbf{E}^\alpha$ is a PO measure, and the claimed equality holds for the characteristic functions of the sets in $\mathcal{B}$ and then by linearity for simple functions. The general case follows from Lemma 4.7 and Proposition 4.12. The last claim is obvious. $\square$

4.9 The Connection Between (Semi)Spectral Measures and (Semi)Spectral Functions

The construction of spectral measures on the real line traditionally often involves so-called spectral functions. The present section is devoted to this question (also in a somewhat more general setting).

Definition 4.10 We say that a function $\mathbf{F} : \mathbb{R} \rightarrow \mathcal{L}_s(\mathcal{H})^+$ is a *semispectral function* if $\mathbf{F}(s) \leq \mathbf{F}(t) \leq I$ whenever $s \leq t$, $\mathbf{F}$ is weakly continuous from the right (i.e. $\text{w-lim}_{t \rightarrow s+} \mathbf{F}(t) = \mathbf{F}(s)$ for all $s \in \mathbb{R}$), $\text{w-lim}_{t \rightarrow -\infty} \mathbf{F}(t) = 0$ and $\text{w-lim}_{t \rightarrow \infty} \mathbf{F}(t) = I$. A semispectral function $\mathbf{F}$ is called a *spectral function* if $\mathbf{F}(t)$ is a projection for all $t \in \mathbb{R}$.

From Theorem 2.7 it follows that since $\mathbf{F}$ is increasing, in the above definition $\mathbf{F}$ is even strongly continuous, i.e. we may replace each w-lim with s-lim. From the same theorem it follows that $\mathbf{F}(s) = \inf \{ \mathbf{F}(t) \mid t > s \}$ for all $s \in \mathbb{R}$.

We take for granted the well-known bijective correspondence $g \mapsto \mu_g$ between the set of the increasing right continuous functions $g : \mathbb{R} \rightarrow [0, 1]$ with $\lim_{t \rightarrow -\infty} g(t) = 0$, $\lim_{t \rightarrow \infty} g(t) = 1$, and the set of the probability measures $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$, characterised by $g(t) = \mu_g((-\infty, t])$ for $t \in \mathbb{R}$. We say that μ_g is the *Lebesgue–Stieltjes measure* defined by g . (This name is also used without normalisation assumptions.) The following theorem generalises this correspondence.

Theorem 4.3 *If $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}_s(\mathcal{H})^+$ is a semispectral measure, then the function $F : \mathbb{R} \rightarrow \mathcal{L}_s(\mathcal{H})^+$ defined by $F(t) = E((-\infty, t])$ for $t \in \mathbb{R}$ is a semispectral function. Conversely, if $F : \mathbb{R} \rightarrow \mathcal{L}_s(\mathcal{H})^+$ is a semispectral function, there is a uniquely determined semispectral measure $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}_s(\mathcal{H})^+$ such that $E((-\infty, t]) = F(t)$ for all $t \in \mathbb{R}$. Here E is a spectral measure if and only if F is a spectral function.*

Proof The function F defined in terms of E is increasing in view of Proposition 4.6, and using Theorem 2.7 we obtain the limit conditions required by the definition of a semispectral function. For example,

$$\begin{aligned} \text{w-lim}_{t \rightarrow s+} F(t) &= \inf_{t > s} F(t) = \inf \{F(s + 1/n) \mid n \in \mathbb{N}\} \\ &= \text{w-lim}_{n \rightarrow \infty} F(s + 1/n) = \text{w-lim}_{n \rightarrow \infty} E((-\infty, s + 1/n)) \\ &= E(\cap_{n=1}^{\infty} (-\infty, s + 1/n)) = F(s). \end{aligned}$$

If E is a spectral measure then of course F is a spectral function.

Suppose next, to prove the converse, that $F : \mathbb{R} \rightarrow \mathcal{L}_s(\mathcal{H})^+$ is a semispectral function. Let $\mathcal{R}$ denote the ring consisting of $\emptyset$ and the finite disjoint unions of half-open intervals $(a, b]$. There is a unique additive set function $E_0 : \mathcal{R} \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $E_0((a, b]) = F(b) - F(a)$ whenever $a < b$. Since F is weakly right continuous, E_0 satisfies the assumption in Proposition 4.7, and so we get a uniquely determined weakly σ -additive extension $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ of E_0 . Then $E((-\infty, t]) = \lim_{n \rightarrow \infty} E((-n, t])$ for all $t \in \mathbb{R}$, and clearly this condition determines E on $\mathcal{R}$ and hence on $\mathcal{B}(\mathbb{R})$. The last statement also follows from Proposition 4.7. $\square$

We state explicitly the following useful consequence of the above result.

Corollary 4.1 *Let $E_1, E_2 : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}_s(\mathcal{H})^+$ be semispectral measures. The following conditions are equivalent:*

- (i) $E_1 = E_2$;
- (ii) $\int_{\mathbb{R}} f(t) dE_1(t) = \int_{\mathbb{R}} f(t) dE_2(t)$ for any continuous function $f : \mathbb{R} \rightarrow [0, 1]$ vanishing outside a compact set;
- (iii) $E_1((a, b]) = E_2((a, b])$ whenever $a < b$;
- (iv) $E_1((-\infty, t]) = E_2((-\infty, t])$ for every $t \in \mathbb{R}$.

Proof Obviously (i) $\implies$ (ii). Assume now (ii) and take $a < b$. For each great enough $n \in \mathbb{N}$ define $f_n(t) = 1$ if $a + \frac{1}{n} \leq t \leq b$, $f_n(t) = 0$ if $t \geq b + \frac{1}{n}$ or $t \leq a$, and require f_n to be affine and continuous on the closed intervals $[a, a + \frac{1}{n}]$ and $[b, b + \frac{1}{n}]$. The sequence (f_n) converges pointwise to $\chi_{(a,b]}$, and so the equality in (iii) follows from the dominated convergence theorem applied to the positive measures $X \mapsto \langle \varphi \mid E_j(X)\varphi \rangle, j = 1, 2$. The implication (iii) $\implies$ (iv) is clear since $E_j((-\infty, t]) = \text{w-lim}_{n \rightarrow \infty} E_j((t - n, t])$. The implication (iv) $\implies$ (i) follows from the uniqueness statement in the preceding theorem. $\square$

4.10 A Riesz–Markov–Kakutani Type Representation Theorem for Positive Operator Measures

We find it convenient to insert here a short discussion of a connection between positive operator measures and operator valued linear maps. Strictly speaking we need this information only in a much later chapter, but even at this stage it gives an alternative approach to a measure theoretic part in the proof of the spectral representation theorem of a bounded selfadjoint operator to be presented in the next section via an elementary though slightly tedious argument. From now on we take for granted the basics of general topology. The notions we need can be found in any introductory topology text.

Let Ω be a locally compact Hausdorff space and $C_0(\Omega)$ the Banach space of continuous complex functions on Ω vanishing at infinity, equipped with the supremum norm. A complex measure $\mu : \mathcal{B}(\Omega) \rightarrow \mathbb{C}$ is *regular* if its total variation $|\mu|$ is regular in the sense that for any $X \in \mathcal{B}(\Omega)$ $|\mu|(X) = \sup\{|\mu|(K) \mid K \subset X \text{ compact}\} = \inf\{|\mu|(G) \mid G \supset X \text{ open}\}$. We take for granted the *Riesz–Markov–Kakutani representation theorem* which says that for any bounded linear functional $F : C_0(\Omega) \rightarrow \mathbb{C}$ there is a unique regular complex measure $\mu_F : \mathcal{B}(\Omega) \rightarrow \mathbb{C}$ such that $F(f) = \int_{\Omega} f d\mu_F$ whenever $f \in C_0(\Omega)$. Moreover, the correspondence $F \mapsto \mu_F$ satisfies $\|F\| = |\mu_F|(\Omega)$, and μ_F is a positive measure if and only if F is a positive linear functional.

We let $\text{supp}(f)$ denote the *support* of a function $f : \Omega \rightarrow \mathbb{C}$, i.e. the closure of the complement of the set where f vanishes. A positive operator measure $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ is said to be *regular* if the complex measure $\langle \xi \mid E(\cdot)\eta \rangle$ is regular for all $\xi, \eta \in \mathcal{H}$. It is actually enough to require this, when $\xi = \eta$ (exercise).

Theorem 4.4 *Let $\Phi : C_0(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ be a linear map which is positive, i.e. $\Phi(f) \geq 0$ whenever $f \geq 0$.*

- (a) *The linear map Φ is bounded.*
- (b) *There is a unique regular positive operator measure $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ such that*

$$\Phi(f) = \int_{\Omega} f dE$$

for all $f \in C_0(\Omega)$.

- (c) *For any open set $G \subset \Omega$ we have $E(G) = \sup\{\Phi(f) \mid f \in C_0(\Omega), 0 \leq f \leq 1, \text{supp}(f) \subset G\}$. In particular, the POM E is normalised, i.e. a semispectral measure, if and only if $\sup\{\Phi(f) \mid f \in C_0(\Omega), 0 \leq f \leq 1\} = I$.*
- (d) *We have $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in C_0(\Omega)$ if and only if E is projection valued.*

Proof (a) We first show that there is an upper bound for the norms $\|\Phi(f)\|$ with $0 \leq f \leq 1$. If not, there is a sequence (f_n) in $C_0(\Omega)$ with $0 \leq f_n \leq 1$ and $\|\Phi(f_n)\| \geq 4^n$. Then the series $\sum_{n=1}^{\infty} 2^{-n} f_n$ converges uniformly to some $f \in C_0(\Omega)$ since $C_0(\Omega)$ is complete. Then $\Phi(f) \geq 2^{-n} \Phi(f_n)$ for all $n \in \mathbb{N}$, implying $\|\Phi(f)\| \geq$

$2^{-n} \|\Phi(f_n)\| \geq 2^{-n} 4^n = 2^n$ for all $n \in \mathbb{N}$, which is impossible. Any $g \in C_0(\Omega)$ with $\|f\| \leq 1$ may be written as $g = g_1 - g_2 + ig_3 - ig_4$ with each $g_k \in C_0(\Omega)$ satisfying $0 \leq g_k \leq 1$, the general boundedness claim follows.

- (b) For any $\xi, \eta \in \mathcal{H}$ define the bounded linear functional $\Phi_{\xi, \eta} : C_0(\Omega) \rightarrow \mathbb{C}$ by the formula $\Phi_{\xi, \eta}(f) = \langle \xi | \Phi(f)\eta \rangle$. There is a unique regular complex measure $\mu_{\xi, \eta} : \mathcal{B}(\Omega) \rightarrow \mathbb{C}$ such that $\Phi_{\xi, \eta}(f) = \int_{\Omega} f d\mu_{\xi, \eta}$ for all $f \in C_0(\Omega)$. For $X \in \mathcal{B}(\Omega)$, $\xi, \eta \in \mathcal{H}$, we define $S_X(\xi, \eta) = \mu_{\xi, \eta}(X)$. The map $(\xi, \eta) \mapsto S_X(\xi, \eta)$ is easily seen to be sesquilinear since the map $(\xi, \eta) \mapsto \langle \xi | \Phi(f)\eta \rangle$ is sesquilinear for all $f \in C_0(\Omega)$ and the correspondence which maps $\Phi_{\xi, \eta}$ to $\mu_{\xi, \eta}$ is injective. Since $|S_X(\xi, \eta)| \leq |\mu_{\xi, \eta}|(\Omega) = \|\Phi_{\xi, \eta}\| \leq \|\Phi\| \|\xi\| \|\eta\|$, S_X is a bounded sesquilinear form, and so using Proposition 2.1 we get a unique operator $E(X) \in \mathcal{L}(\mathcal{H})$ such that $\langle \xi | E(X)\eta \rangle = \mu_{\xi, \eta}(X)$ for all $\xi, \eta \in \mathcal{H}$. Since each $\mu_{\xi, \xi}$ is a positive measure, E is a positive operator measure, and the claimed formula holds by Proposition 4.12(a). By construction, E is regular, and its uniqueness follows from the uniqueness part of the Riesz–Markov–Kakutani theorem.
- (c) Let $G \subset \Omega$ be an open set and F the set of the functions $f \in C_0(G)$ satisfying $0 \leq f \leq 1$ and having their supports contained in G . By regularity and Urysohn’s lemma we easily see (exercise) the claimed equality. In fact F is a directed set with pointwise order and $E(G)$ is the weak limit of the net $(\Phi(f))_{f \in F}$.
- (d) Assume first the multiplicativity condition $\Phi(fg) = \Phi(f)\Phi(g)$. In the notation of the proof of (b), using Theorem 2.8(b) we find that $E(G)^2 = [\text{w-lim } \Phi(f)]^2 = \text{w-lim } \Phi(f^2) = E(G)$, since the functions f^2 with $f \in F$ are just the elements of F . Thus $E(G)^2 = E(G)$ for any open set G . By regularity, for any $X \in \mathcal{B}(\Omega)$ $\Phi(X)$ is the weak limit of the net $(\Phi(G))$ defined on the set of the open sets containing X , directed by requiring that $G_1 \leq G_2$ if $G_2 \subset G_1$. Using Theorem 2.8(b) again we get $E(X)^2 = [\text{w-lim } E(G)]^2 = \text{w-lim } E(G)^2 = \text{w-lim } E(G) = E(X)$. Thus E is projection valued. The converse follows from Proposition 4.13. $\square$

In the situation of the above theorem, we call the positive linear map Φ *normalised* if the corresponding regular POM E is normalised, i.e. $E(\Omega) = I$.

4.11 The Spectral Representation of Bounded Selfadjoint Operators

We assume throughout this section that $A \in \mathcal{L}(\mathcal{H})$ is a selfadjoint operator. We fix two real numbers a and b such that $a < b$ and $aI \leq A \leq bI$; this is possible since $\sup_{\|\varphi\| \leq 1} |\langle \varphi | A\varphi \rangle| < \infty$. As an intermediate step towards proving the spectral representation of a bounded selfadjoint operator, we construct a suitable bounded linear map from the space $C[a, b]$ of continuous functions (with the supremum norm) on the interval $[a, b]$ into the operator space $\mathcal{L}(\mathcal{H})$. We begin with an auxiliary result.

Lemma 4.8 *Let p be a polynomial with real coefficients such that $p(t) \geq 0$ for all $t \in [a, b]$. The polynomial p can be expressed in the form*

$$p(t) = \sum_{i=1}^m q_i(t)^2 + \sum_{j=1}^n (t-a)r_j(t)^2 + \sum_{k=1}^{\ell} (b-t)s_k(t)^2$$

where q_i , r_j , and s_k are polynomials with real coefficients.

Proof We may assume that $p \neq 0$ is not constant and write it as a constant α times a product of first degree polynomials and monic second degree polynomials without real roots as follows:

$$p(t) = \alpha \prod_{i=1}^{n_1} (t - c_i)^{\lambda_i} \prod_{j=1}^{n_2} (t - d_j)^{\mu_j} \prod_{k=1}^{n_3} (e_k - t)^{\nu_k} \prod_{r=1}^s Q_r(t).$$

Here we assume that $c_i \leq a$, $a < d_j < b$, and $e_k \geq b$. Each μ_j must be even, for otherwise p could not preserve its sign on the interval (a, b) . Clearly $\alpha > 0$. For each $r = 1, \dots, s$ there are $\alpha_r \in \mathbb{R}$ and $\beta_r > 0$ such that $Q_r(t) = (t - \alpha_r)^2 + \beta_r$. Writing $t - c_i = (t - a) + (a - c_i)$ and $e_k - t = (e_k - b) + (b - t)$ (where $a - c_i \geq 0$ and $e_k - b \geq 0$) and performing the multiplications we find that $p(t)$ can be expressed as stated in the lemma, except that in addition there may be terms of the form $(t - a)(b - t)q(t)^2$ where q is a real polynomial. But the formula

$$(t - a)(b - t) = (b - a)^{-1}[(t - a)^2(b - t) + (t - a)(b - t)^2]$$

shows that these terms can also be included in a sum of the same form. $\square$

The next results contains in an algebraic form the essence of the spectral representation.

Theorem 4.5 *There is a uniquely determined continuous linear map $\Phi : C[a, b] \rightarrow \mathcal{L}(\mathcal{H})$ satisfying the following conditions:*

- (i) $\Phi(1) = I$ where 1 is the function with the constant value 1 ;
- (ii) $\Phi(\text{id}) = A$ where $\text{id}(t) = t$ for all $t \in [a, b]$;
- (iii) $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in C[a, b]$.

Proof If $p(t) = a_0 + a_1t + \dots + a_nt^n$ for $t \in [a, b]$, the coefficients $a_k \in \mathbb{C}$ are uniquely determined, and so we may define

$$\Phi_0(p) = a_0I + \sum_{k=1}^n a_k A^k.$$

Denoting by $P[a, b]$ the set of the restrictions of the polynomials with complex coefficients to $[a, b]$, we may in this way define a linear map $\Phi_0 : P[a, b] \rightarrow \mathcal{L}(\mathcal{H})$ satisfying the above conditions (i)–(iii). If $q \in P[a, b]$ is the restriction of a polynomial with real coefficients, then $\|q\| \pm q(t) \geq 0$. In view of the preceding lemma we may thus write

$$\|q\| I + \Phi_0(q) = \sum_{i=1}^m \Phi_0(q_i)^2 + \sum_{j=1}^n (A - aI) \Phi_0(r_j)^2 + \sum_{k=1}^p (bI - A) \Phi_0(s_k)^2$$

for certain polynomials q_i , r_j and s_k with real coefficients (restricted to $[a, b]$). It follows that $\|q\| I + \Phi_0(q) \geq 0$ since e.g. $A - aI \geq 0$ and the operators $A - aI$ and $\Phi_0(r_j)$ commute. Similarly, $\|q\| I - \Phi_0(q) \geq 0$, and it follows from Proposition 2.2 that $\|\Phi_0(q)\| \leq \|q\|$. If q is the restriction to $[a, b]$ of an arbitrary polynomial, we find

$$\|\Phi_0(q)\| \leq \|\Phi_0(\operatorname{Re} q)\| + \|\Phi_0(\operatorname{Im} q)\| \leq \|\operatorname{Re} q\| + \|\operatorname{Im} q\| \leq 2\|q\|,$$

and so Φ_0 is a continuous linear map. The Stone–Weierstrass theorem implies that $P[a, b]$ is dense in $C[a, b]$, and so Φ_0 can be extended to a continuous linear map $\Phi : C[a, b] \rightarrow \mathcal{L}(\mathcal{H})$. The map Φ also satisfies the conditions (i)–(iii); for example $\Phi(fg) = \lim_{n \rightarrow \infty} \Phi_0(f_n g_n) = \lim_{n \rightarrow \infty} \Phi_0(f_n) \Phi_0(g_n) = \Phi(f) \Phi(g)$, whenever $f_n, g_n \in P[a, b]$ and $f_n \rightarrow f, g_n \rightarrow g$ uniformly. The uniqueness of the map Φ follows from continuity and the fact that it is by linearity and the conditions (i)–(iii) completely determined in the dense subspace $P[a, b]$ of $C[a, b]$. $\square$

Remark 4.9 If Φ is as in the preceding theorem, then $\Phi(\bar{f}) = \Phi(f)^*$ for all $f \in C[a, b]$ or, equivalently, $\Phi(f)$ is selfadjoint for every real valued function $f \in C[a, b]$. Since A^k is selfadjoint for any $k \in \mathbb{N}$, this follows from the fact that the restrictions of the polynomials with real coefficients are dense in the space of real continuous functions on $[a, b]$. Moreover, if $f \in C[a, b]$ and $f \geq 0$, then $\Phi(f) = \Phi(f^{1/2})^2 \geq 0$, since $\langle \varphi | \Phi(f^{1/2})^2 \varphi \rangle = \langle \Phi(f^{1/2}) \varphi | \Phi(f^{1/2}) \varphi \rangle \geq 0$. $\triangleleft$

The map Φ obtained in Theorem 4.5 can be represented by integration with respect to a suitable spectral measure. Before proving this, we consider a somewhat more general situation.

Theorem 4.6 *Assume that $\Phi : C[a, b] \rightarrow \mathcal{L}(\mathcal{H})$ is a linear map satisfying $\Phi(1) = I$ and $\Phi(f) \geq 0$ whenever $f \geq 0$. There is a uniquely determined semispectral measure $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}_s(\mathcal{H})^+$ such that $\Phi(f) = \int_{[a, b]} f d\mathbf{E}$ for all $f \in C[a, b]$. If $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in C[a, b]$, then $\mathbf{E}$ is a spectral measure.*

Proof (First proof) Since every complex measure on $[a, b]$ is regular by Proposition 4.9, this is a special case of Theorem 4.4. $\square$

Since this proof depends on the Riesz–Markov–Kakutani representation theorem which we have used without proof, we also give a more elementary approach based on the use of a semispectral function. We retain the assumptions of Theorem 4.6 and define a function $\mathbf{F} : \mathbb{R} \rightarrow \mathcal{L}_s(\mathcal{H})^+$ in the following way: $\mathbf{F}(\lambda) = 0$ for $\lambda < a$, $\mathbf{F}(\lambda) = I$ for $\lambda \geq b$, and for any number $\lambda \in [a, b]$, $\mathbf{F}(\lambda)$ is the infimum of the set of the operators $\Phi(f)$ where $f \in C[a, b]$, $f \geq 0$ and $f(t) = 1$ if $t \in [a, \lambda]$. (According to Theorem 2.7 this infimum exists.)

In item (a) below, the estimate $\|\Phi\| \leq 2$ could be replaced by $\|\Phi\| = 1$, but the proof would be more complicated and we do not need this result (which of course also follows from Theorem 4.6).

Lemma 4.9 (a) *The linear map Φ is continuous and $\|\Phi\| \leq 2$.*

(b) *The function F is a semispectral function.*

(c) *If $\lambda \in [a, b]$, $\varphi \in \mathcal{H}$ and $\epsilon > 0$, then there is a $\delta > 0$ such that*

$$|\langle \varphi | \Phi(f)\varphi \rangle| < \epsilon$$

whenever $f \in C[a, b]$ is such that $\|f\| \leq 1$ and $f(t) = 0$ outside the interval $(\lambda, \lambda + \delta]$.

Proof (a) Since $\|f\| I \pm \Phi(f) \geq 0$ when $f \in C[a, b]$ is real valued, the corresponding argument in the proof of Theorem 4.5 works here.

(b) We show that F is weakly right continuous; the other conditions in Definition 4.10 are obvious. It is enough to consider a point $\lambda \in [a, b]$. Suppose that $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$, and $\epsilon > 0$. We choose a function $f \in C[a, b]$, $f \geq 0$, such that $f(t) = 1$ for all $t \in [a, \lambda]$ and $\langle \varphi | \Phi(f)\varphi \rangle < \langle \varphi | F(\lambda)\varphi \rangle + \frac{\epsilon}{2}$ (see Theorem 2.7). For any number $\delta \in (0, b - \lambda)$ we define $f_\delta(t) = \tilde{f}(t - \delta)$ for all $t \in [a, b]$, where $\tilde{f}(t) = f(t)$ when $t \in [a, b]$, and $\tilde{f}(t) = 1$, when $t < a$. Since f is uniformly continuous, we have $\|f - f_\delta\| < \frac{\epsilon}{4}$, when δ is small enough, so that by (a) we get

$$\langle \varphi | F(\lambda + \delta)\varphi \rangle \leq \langle \varphi | \Phi(f_\delta)\varphi \rangle \leq \langle \varphi | \Phi(f)\varphi \rangle + \frac{\epsilon}{2} < \langle \varphi | F(\lambda)\varphi \rangle + \epsilon.$$

This implies the claim.

(c) Choose $\delta > 0$ such that $\langle \varphi | F(\lambda + \delta)\varphi \rangle < \langle \varphi | F(\lambda)\varphi \rangle + \epsilon$ (see (b)). If f is the kind of function mentioned in the claim (c), we can construct a function $g \in C[a, b]$, $g \geq 0$, such that $g(t) = 1$ when $t \in [a, \lambda]$, $g(t) + |f(t)| \leq 1$ when $t \in [\lambda, \lambda + \delta]$, and $g(t) = 0$ when $t \in [\lambda + \delta, b]$. (Put e.g. $g(t) = (1 - |f(t)|)(1 - \delta^{-1}(t - \lambda))$ when $t \in [\lambda, \lambda + \delta]$.) If $\alpha \in \mathbb{C}$ is such that $|\alpha| = 1$ and $|\langle \varphi | \Phi(f)\varphi \rangle| = \alpha \langle \varphi | \Phi(f)\varphi \rangle$, then we get

$$\begin{aligned} |\langle \varphi | \Phi(f)\varphi \rangle| &= \operatorname{Re} \langle \varphi | \Phi(\alpha f)\varphi \rangle = \langle \varphi | \Phi(\operatorname{Re} \alpha f)\varphi \rangle \leq \langle \varphi | \Phi(|\alpha f|)\varphi \rangle \\ &= \langle \varphi | \Phi(|f|)\varphi \rangle = \langle \varphi | \Phi(g + |f|)\varphi \rangle - \langle \varphi | \Phi(g)\varphi \rangle \\ &\leq \langle \varphi | F(\lambda + \delta)\varphi \rangle - \langle \varphi | F(\lambda)\varphi \rangle < \epsilon, \end{aligned}$$

for $g + |f| \leq h$ whenever $h \in C[a, b]$, $h \geq 0$, and $h(t) = 1$ when $a \leq t \leq \lambda + \delta$. $\square$

Proof (Second proof of Theorem 4.6) Consider the semispectral function F studied in the above lemma and let $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}_s(\mathcal{H})^+$ be the semispectral measure corresponding to it in accordance with Theorem 4.3, i.e. $E((-\infty, t]) = F(t)$ for all $t \in \mathbb{R}$. Let $f \in C[a, b]$. We are going to show that

$$\Phi(f) = \int_{[a,b]} f d\mathbf{E}. \quad (4.1)$$

Suppose that $\varphi \in \mathcal{H}$ and $\epsilon > 0$. Since f is uniformly continuous, we can find some points $a = t_0 < t_1 < \dots < t_p = b$ and numbers $\alpha_1, \dots, \alpha_p \in \mathbb{C}$ such that $|f(t) - \alpha_i| < \epsilon$ whenever $t_{i-1} \leq t \leq t_i$. By Lemma 4.9(b) there is a $\delta > 0$ such that

$$\langle \varphi | \mathbf{F}(t_i + \delta) \varphi \rangle - \langle \varphi | \mathbf{F}(t_i) \varphi \rangle < \frac{\epsilon}{p} \quad (4.2)$$

for $i = 1, \dots, p$, and that the claim of Lemma 4.9(c) is true at all points $\lambda = t_0, \dots, t_{p-1}$ for the number $\frac{\epsilon}{p}$ in place of ϵ . We may assume that $\delta < t_{i+1} - t_i$. We construct a function $g \in C[a, b]$ as follows: $g(t) = \alpha_1$ for $t \in [a, t_1]$, $g(t) = \alpha_{i+1}$ for $t \in [t_i + \delta, t_{i+1}]$, $i = 1, \dots, p-1$, and $g(t) = \alpha_i + \delta^{-1}(\alpha_{i+1} - \alpha_i)(t - t_i)$ for $t \in [t_i, t_i + \delta]$, $i = 1, \dots, p-1$. Since $|\alpha_{i+1} - \alpha_i| \leq |\alpha_{i+1} - f(t_i)| + |f(t_i) - \alpha_i| < 2\epsilon$, we have $|g(t) - \alpha_i| < 2\epsilon$ when $t \in [t_{i-1}, t_i]$, implying $\|f - g\| < 3\epsilon$. Using the definition of $\mathbf{F}$ we may define for each $i = 1, \dots, p-1$ a continuous function $h_i : [a, b] \rightarrow [0, 1]$ such that $h_i(t) = 1$ when $t \in [a, t_i]$, $h_i(t) = 0$ when $t \in [t_i + \delta, b]$, and

$$\langle \varphi | \mathbf{F}(t_i) \varphi \rangle \leq \langle \varphi | \Phi(h_i) \varphi \rangle < \langle \varphi | \mathbf{F}(t_i) \varphi \rangle + \frac{\epsilon}{p}. \quad (4.3)$$

By (4.2)

$$\langle \varphi | \mathbf{F}(t_i) \varphi \rangle \leq \int h_i d\mathbf{E}_{\varphi, \varphi} \leq \langle \varphi | \mathbf{F}(t_i + \delta) \varphi \rangle < \langle \varphi | \mathbf{F}(t_i) \varphi \rangle + \frac{\epsilon}{p}, \quad (4.4)$$

and so

$$\left| \langle \varphi | \Phi(h_i) \varphi \rangle - \left\langle \varphi \left| \left(\int h_i d\mathbf{E} \right) \varphi \right\rangle \right| < \frac{\epsilon}{p}. \quad (4.5)$$

We denote $h = \sum_{i=1}^p \alpha_i (h_i - h_{i-1})$ where identically $h_0 = 0$, $h_p = 1$. Then using (4.5) we see that

$$\left| \left\langle \varphi \left| \left(\int h d\mathbf{E} \right) \varphi \right\rangle - \langle \varphi | \Phi(h) \varphi \rangle \right| \leq \max_{1 \leq i \leq p} |\alpha_i| 2\epsilon \leq (\|f\| + \epsilon) 2\epsilon. \quad (4.6)$$

On the other hand, by Lemma 4.9,

$$\begin{aligned} |\langle \varphi | \Phi(f) \varphi \rangle - \langle \varphi | \Phi(h) \varphi \rangle| &\leq 2 \|f - g\| \|\varphi\|^2 + |\langle \varphi | \Phi(g) \varphi \rangle - \langle \varphi | \Phi(h) \varphi \rangle| \\ &\leq 6\epsilon \|\varphi\|^2 + \sum_{i=1}^{p-1} |\langle \varphi | \Phi(g_i) \varphi \rangle| \\ &\leq 6\epsilon \|\varphi\|^2 + \epsilon(2 \|f\| + 4\epsilon), \end{aligned} \quad (4.7)$$

for we can write $h - g = \sum_{i=1}^{p-1} g_i$ with functions $g_i \in C[a, b]$ satisfying $\|g_i\| \leq 2\|f\| + 4\epsilon$ and such that each g_i vanishes outside the interval $[t_i, t_i + \delta)$ (for $|\alpha_i(h_i(t) - h_{i-1}(t))| = |\alpha_i| \leq f + \epsilon$ when $t \in [t_{i-1}, t_i]$, and $\|g\| \leq \|f\| + 3\epsilon$). Moreover,

$$\begin{aligned} & \left| \left\langle \varphi \left| \left(\int h d\mathbf{E} \right) \varphi \right\rangle - \left\langle \varphi \left| \left(\int f d\mathbf{E} \right) \varphi \right\rangle \right| \\ & \leq \|f - g\| \|\varphi\|^2 + \sum_{i=1}^{p-1} 3(\|f\| + \epsilon) \langle \varphi | (\mathbf{F}(t_i + \delta) - \mathbf{F}(t_i)) \varphi \rangle \\ & \leq 3\epsilon \|\varphi\|^2 + 3\epsilon(\|f\| + \epsilon). \end{aligned} \quad (4.8)$$

by (4.2). As $\epsilon > 0$ is arbitrary, by (4.6)–(4.8) we get

$$\langle \varphi | \Phi(f) \varphi \rangle = \left\langle \varphi \left| \left(\int f d\mathbf{E} \right) \varphi \right\rangle,$$

implying (4.1) by the polarisation identity.

To prove the uniqueness claim, suppose that two semispectral measures $\mathbf{E}_1, \mathbf{E}_2$ satisfy $\Phi(f) = \int_{[a,b]} f d\mathbf{E}_j$ for all $f \in C[a, b]$, $j = 1, 2$. Since $\mathbf{E}_j([a, b]) = \Phi(1) = I$, we must have $\mathbf{E}_j(\mathbb{R} \setminus [a, b]) = 0$. Thus if $f : \mathbb{R} \rightarrow [0, 1]$ is any continuous function vanishing outside a compact interval, then

$$\int_{\mathbb{R}} f d\mathbf{E}_1 = \int_{[a,b]} f d\mathbf{E}_1 = \Phi(f|[a, b]) = \int_{[a,b]} f d\mathbf{E}_2 = \int_{\mathbb{R}} f d\mathbf{E}_2.$$

Therefore $\mathbf{E}_1 = \mathbf{E}_2$ by Corollary 4.1.

Assume now that $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in C[a, b]$. In order to see that $\mathbf{E}$ is a spectral measure, it is by Theorem 4.3 enough to show that $\mathbf{F}(\lambda)$ is a projection whenever $\lambda \in [a, b)$. Choose a decreasing sequence (f_n) of continuous functions $f_n : [a, b] \rightarrow [0, 1]$ such that $f_n(t) = 1$ when $t \in [a, \lambda]$, and $f_n(t) \rightarrow 0$ when $t \in (\lambda, b]$. Then $\mathbf{F}(\lambda) = \int \chi_{[a,\lambda]} d\mathbf{E} = \text{w-lim}_{n \rightarrow \infty} \int f_n d\mathbf{E}$ according to Proposition 4.13(e). Similarly $\mathbf{F}(\lambda) = \text{w-lim}_{n \rightarrow \infty} \int f_n^2 d\mathbf{E}$. But $\text{w-lim} \int f_n^2 d\mathbf{E} = \text{w-lim} (\int f_n d\mathbf{E})^2 = \mathbf{F}(\lambda)^2$, so that $\mathbf{F}(\lambda)^2 = \mathbf{F}(\lambda)$. Conversely, by Proposition 4.13(b) Φ is multiplicative if $\mathbf{E}$ is a spectral measure. $\square$

We are now in a position to prove the spectral representation of a bounded self-adjoint operator.

Theorem 4.7 *If $A \in \mathcal{L}_s(\mathcal{H})$ and $aI \leq A \leq bI$, there is a uniquely determined spectral measure $\mathbf{E} : \mathcal{B}([a, b]) \rightarrow \mathcal{L}(\mathcal{H})$ such that*

$$A = \int_{[a,b]} \lambda d\mathbf{E}(\lambda).$$

Proof Let Φ be as in Theorem 4.5 (so that $\Phi(f) \geq 0$ whenever $f \geq 0$, since $\langle \varphi | \Phi(f) \varphi \rangle = \langle \Phi(\sqrt{f}) \varphi | \Phi(\sqrt{f}) \varphi \rangle \geq 0$), and let $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be the

associated spectral measure of Theorem 4.6. Since $E(\mathbb{R} \setminus [a, b]) = 0$, by restricting E to the σ -algebra $\mathcal{B}([a, b])$ we obtain the spectral measure E of the claim. Suppose that E' is also a spectral measure on $\mathcal{B}([a, b])$ such that $A = \int_{[a, b]} \lambda dE'(\lambda)$. The mappings $f \mapsto \int_{[a, b]} f dE$ and $f \mapsto \int_{[a, b]} f dE'$ satisfy the assumptions made on Φ in Theorem 4.5 (see Proposition 4.13(b)), and so by its uniqueness part we see that they agree on $C[a, b]$. The proof of the uniqueness part of Theorem 4.6 then shows that $E = E'$. $\square$

As an application of the spectral representation theorem, we give another proof for the fact that every positive operator has a unique positive square root.

Proposition 4.15 *If $A \in \mathcal{L}_s(\mathcal{H})^+$, then there is a uniquely determined $B \in \mathcal{L}_s(\mathcal{H})^+$ such that $B^2 = A$.*

Proof Since $0 \leq A \leq \|A\| I$, by Theorem 4.7 there is a unique spectral measure $E : \mathcal{B}([0, \|A\|]) \rightarrow \mathcal{L}(\mathcal{H})$ such that $A = \int_{[0, \|A\|]} \lambda dE(\lambda)$. We define $B = \int_{[0, \|A\|]} \sqrt{\lambda} dE(\lambda)$. From Proposition 4.12 it follows that $B \in \mathcal{L}_s(\mathcal{H})^+$, and $B^2 = A$ by Proposition 4.13. On the other hand, let $B_0 \in \mathcal{L}_s(\mathcal{H})^+$ be such that $B_0^2 = A$. For B_0 there is a spectral representation $B_0 = \int_{[0, \|B_0\|]} \lambda dE_0(\lambda)$, and if $\alpha : [0, \|B_0\|] \rightarrow [0, \|A\|]$ is defined by $\alpha(t) = t^2$ (note that $\|B_0\|^2 = \|A\|$ by Theorem 2.6 (e)), we may apply Proposition 4.14 to the spectral measure $\tilde{E}_0 : \mathcal{B}([0, \|A\|]) \rightarrow \mathcal{L}(\mathcal{H})$ defined by $\tilde{E}_0(B) = E_0(\alpha^{-1}(B))$ and find that

$$\int_{[0, \|A\|]} \lambda d\tilde{E}_0(\lambda) = \int_{[0, \|B_0\|]} t^2 dE_0(t) = B_0^2 = A$$

(see Proposition 4.13(b)). Therefore by the uniqueness part of Theorem 4.7 we have $\tilde{E}_0 = E$, and so Proposition 4.14 shows that

$$B_0 = \int_{[0, \|A\|]} \sqrt{\lambda} d\tilde{E}_0(\lambda) = \int_{[0, \|A\|]} \sqrt{\lambda} dE(\lambda) = B.$$

$\square$

We conclude this section with a commutation result.

Proposition 4.16 *In the situation of Theorem 4.7 for an operator $T \in \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:*

- (i) $TA = AT$;
- (ii) $TE(X) = E(X)T$ for all $X \in \mathcal{B}([a, b])$.

Proof By Proposition 4.13(a), (ii) implies (i). Assume now (i). We retrace the main steps in the construction of the spectral measure E . Clearly, T commutes with every polynomial in A , and it follows from the proof of Theorem 4.5 that $T\Phi(f) = \Phi(f)T$ where Φ is as in Theorem 4.5. The construction of the spectral function F before Lemma 4.9 involves weakly convergent nets of operators, i.e. infima of certain sets

of positive operators, directed with respect to the order opposite to the usual operator order (compare Theorem 2.7) and since weak convergence commutes with multiplication from the left or right by T (see Theorem 2.8(a)), it follows that $F(t)T = TF(t)$ for all $t \in \mathbb{R}$. Finally, the construction of the spectral measure E from the spectral function F in the proof of Theorem 4.3 only involves weakly convergent nets of operators, and so again the commutation condition is inherited by the next generation. Thus (ii) holds. $\square$

4.12 The Spectrum of a Bounded Operator

In this section we assume that $T \in \mathcal{L}(\mathcal{H})$. In the discussion preceding Theorem 3.4, the spectrum of an operator was defined, as well as the subdivision of the spectrum into various parts. It was shown there that the following definition is equivalent to the definition of the spectrum we used previously.

Definition 4.11 The *spectrum* $\sigma(T)$ of T is the set of those $\lambda \in \mathbb{C}$ for which the operator $T_\lambda = \lambda I - T$ has no inverse belonging to $\mathcal{L}(\mathcal{H})$.

We use the following lemma to prove that the spectrum of any bounded linear operator is a compact subset of the complex plane. We let $\text{Inv}(\mathcal{L}(\mathcal{H}))$ denote the set of the operators in $\mathcal{L}(\mathcal{H})$ having an inverse in $\mathcal{L}(\mathcal{H})$. (The spectrum is known to be always nonempty, but we do not need or prove this general fact. However, in the case of a normal operator this will follow from the spectral representation Theorem 4.10.)

Lemma 4.10 (a) If $T \in \mathcal{L}(\mathcal{H})$ is such that $\|T\| < 1$, then $I - T \in \text{Inv}(\mathcal{L}(\mathcal{H}))$.
 (b) The set $\text{Inv}(\mathcal{L}(\mathcal{H}))$ is open in $\mathcal{L}(\mathcal{H})$.

Proof In (a) the inverse of $I - T$ is obtained as the sum of the absolutely convergent series $I + \sum_{n=1}^{\infty} T^n$, and (b) easily follows from (a) (exercise). $\square$

Theorem 4.8 (a) If $\lambda \in \sigma(T)$, then $|\lambda| \leq \|T\|$.
 (b) The spectrum $\sigma(T)$ of T is closed.

Proof Part (a) follows at once from Lemma 4.10(a), and (b) is a consequence of Lemma 4.10(b) and the continuity of the map $\lambda \mapsto \lambda I - T$. $\square$

Corollary 4.2 If $U \in \mathcal{L}(\mathcal{H})$ is unitary and $\lambda \in \sigma(U)$, then $|\lambda| = 1$.

Proof Since $\|U\| = 1$, $|\lambda| \leq 1$ by Theorem 4.8. On the other hand, if $|\lambda| < 1$, then $|\lambda^{-1}| > 1$, and since $\|U^*\| = 1$, the same result shows that $\lambda^{-1} - U^* \in \text{Inv}(\mathcal{L}(\mathcal{H}))$ implying that $\lambda - U = (-\lambda U)(\lambda^{-1} - U^*)$ is also in $\text{Inv}(\mathcal{L}(\mathcal{H}))$. $\square$

4.13 The Spectral Representations of Unitary and Other Normal Operators

In this section we assume that $T \in \mathcal{L}(\mathcal{H})$ is a normal operator. We write $T = A_1 + iA_2$ with commuting selfadjoint operators A_1 and A_2 . We fix real numbers a_j, b_j such that $a_j I \leq A_j \leq b_j I, j = 1, 2$. By Theorem 4.7 there are unique spectral measures $E_j : \mathcal{B}([a_j, b_j]) \rightarrow \mathcal{L}(\mathcal{H})$ such that

$$A_j = \int_{[a_j, b_j]} \lambda dE_j(\lambda),$$

$j = 1, 2$. These will be used to construct a spectral measure for T . When convenient, we may use the complex notation $(x, y) = x + iy$ for elements of $\mathbb{R}^2 (= \mathbb{C})$.

In the proof of the following theorem we make essential use of the observation made on commuting spectral measures in Example 4.1, which is applicable by Proposition 4.9(b). In this situation $\mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2) = \mathcal{B}(\Omega_1 \times \Omega_2)$ (Proposition 4.10).

Theorem 4.9 *There is a unique spectral measure $E : \mathcal{B}([a_1, b_1] \times [a_2, b_2]) \rightarrow \mathcal{L}(\mathcal{H})$ such that*

$$T = \int_{[a_1, b_1] \times [a_2, b_2]} \lambda dE(\lambda).$$

Proof To prove the existence of a spectral measure E satisfying the formula in the claim, we note that since the selfadjoint operators A_1 and A_2 commute, there are commuting spectral measures $E_j : \mathcal{B}([a_j, b_j]) \rightarrow \mathcal{L}(\mathcal{H})$ such that

$$A_j = \int_{[a_j, b_j]} \lambda dE_j(\lambda),$$

$j = 1, 2$ (see Theorem 4.7 and Proposition 4.16). By Example 4.1 there is a spectral measure $E : \mathcal{B}([a_1, b_1] \times [a_2, b_2]) \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $E(X_1 \times X_2) = E_1(X_1)E_2(X_2)$ for all $X_j \in \mathcal{B}([a_j, b_j]), j = 1, 2$. If $h : [a_1, b_1] \times [a_2, b_2] \rightarrow [a_1, b_1]$ is the map defined by $h(x, y) = x$, then $E(h^{-1}(X)) = E_1(X)$ for all $X \in \mathcal{B}([a_1, b_1])$, and so

$$\int_{[a_1, b_1] \times [a_2, b_2]} \operatorname{Re} \lambda dE(\lambda) = \int_{[a_1, b_1]} x dE_1(x) = A_1.$$

Similarly,

$$\int_{[a_1, b_1] \times [a_2, b_2]} \operatorname{Im} \lambda dE(\lambda) = A_2.$$

It follows that

$$\int_{[a_1, b_2] \times [a_2, b_2]} \lambda dE(\lambda) = A_1 + iA_2 = T.$$

We now prove the uniqueness statement. Assume that $\mathbf{E}$ is as in the claim. Define $\tilde{\mathbf{E}}_1(X) = \mathbf{E}(X \times [a_2, b_2])$ for all $X \in \mathcal{B}([a_1, b_1])$. Then $\tilde{\mathbf{E}}_1$ is a spectral measure and by Proposition 4.14

$$\int_{[a_1, b_1] \times [a_2, b_2]} \operatorname{Re} \lambda d\mathbf{E}(\lambda) = \int_{[a_1, b_1]} x d\tilde{\mathbf{E}}_1(x).$$

Since

$$A_1 = \frac{1}{2}(T + T^*) = \int_{[a_1, b_1] \times [a_2, b_2]} \frac{1}{2}(\lambda + \bar{\lambda}) d\mathbf{E}(\lambda) = \int_{[a_1, b_1]} x d\tilde{\mathbf{E}}_1(x)$$

it follows from Theorem 4.7 that $\tilde{\mathbf{E}}_1 = \mathbf{E}_1$. A similar argument shows that $\mathbf{E}_2(Y) = \mathbf{E}([a_1, b_2] \times Y)$ for all $Y \in \mathcal{B}([a_2, b_2])$. Hence

$$\begin{aligned} \mathbf{E}(X \times Y) &= \mathbf{E}((X \times [a_2, b_2]) \cap ([a_1, b_1] \times Y)) = \mathbf{E}((X \times [a_2, b_2])\mathbf{E}([a_1, b_1] \times Y)) \\ &= \mathbf{E}_1(X)\mathbf{E}_2(Y) \end{aligned}$$

for all $X \in \mathcal{B}([a_1, b_1])$, $Y \in \mathcal{B}([a_2, b_2])$. But by Example 4.1 there is only one spectral measure on $\mathcal{B}([a_1, b_1] \times [a_2, b_2])$ with this property. $\square$

We next prove the spectral representation of a normal operator in a form involving the spectrum.

Theorem 4.10 *There is a unique spectral measure $\mathbf{E} : \mathcal{B}(\sigma(T)) \rightarrow \mathcal{L}(\mathcal{H})$ such that $T = \int_{\sigma(T)} \lambda d\mathbf{E}(\lambda)$. Moreover, $\sigma(T)$ is the support of $\mathbf{E}$.*

Proof Since $\sigma(T)$ is compact, in the situation described before Theorem 4.9 we may assume that $\sigma(T) \subset [a_1, b_1] \times [a_2, b_2]$. Let $\mathbf{E} : \mathcal{B}([a_1, b_1] \times [a_2, b_2]) \rightarrow \mathcal{L}(\mathcal{H})$ be a spectral measure such that $T = \int_{[a_1, b_1] \times [a_2, b_2]} \lambda d\mathbf{E}(\lambda)$. We show that $\mathbf{E}([a_1, b_1] \times [a_2, b_2] \setminus \sigma(T)) = 0$. Since $[a_1, b_1] \times [a_2, b_2] \setminus \sigma(T)$ is the union of a countable number of compact sets, using the σ -subadditivity of the measures $\mathbf{E}_{\varphi, \varphi}$ we see that it suffices to show that $\mathbf{E}(K) = 0$ for any compact set $K \subset [a_1, b_1] \times [a_2, b_2] \setminus \sigma(T)$. By compactness and subadditivity, it is enough to show that for each $z \in K$ there is a $\delta > 0$ such that $\mathbf{E}(B(z, \delta)) = 0$ where $B(z, \delta) = \{w \in [a_1, b_1] \times [a_2, b_2] \mid \|w - z\| < \delta\}$. Suppose that for some $z \in [a_1, b_1] \times [a_2, b_2] \setminus \sigma(T)$ we have $\mathbf{E}(B(z, \frac{1}{n})) \neq 0$ whenever $n \in \mathbb{N}$. Choose $\varphi_n \in \mathbf{E}(B(z, \frac{1}{n}))(\mathcal{H})$ with $\|\varphi_n\| = 1$ for each $n \in \mathbb{N}$. Then by Proposition 4.13

$$\begin{aligned}
1 &= \|\varphi_n\| = \|(zI - T)^{-1}(zI - T)\mathbf{E}(B(z, 1/n))\varphi_n\| \\
&= \left\| (zI - T)^{-1} \int_{B(z, 1/n)} (z - \lambda) d\mathbf{E}(\lambda) \varphi_n \right\| \\
&\leq \|(zI - T)^{-1}\| \left\| \int_{B(z, 1/n)} (z - \lambda) d\mathbf{E}(\lambda) \right\| \\
&\leq \|(T - zI)^{-1}\| \frac{1}{n} \rightarrow 0,
\end{aligned}$$

when $n \rightarrow \infty$. This contradiction proves our claim, and thus the existence of $\mathbf{E}$ as a consequence of Theorem 4.9. The uniqueness part also follows from Theorem 4.9, since any spectral measure on $\mathcal{B}(\sigma(T))$ may be regarded as a spectral measure on $\mathcal{B}([a_1, b_1] \times [a_2, b_2])$ vanishing on $[a_1, b_1] \times [a_2, b_2] \setminus \sigma(T)$. To prove that $\sigma(T) = \text{supp}(\mathbf{E})$ we now only need to show that $\sigma(T) \subset \text{supp}(\mathbf{E})$. Suppose that $z \in \mathbb{C} \setminus \text{supp}(\mathbf{E})$. Then the function $\lambda \mapsto (z - \lambda)^{-1}$ is defined, continuous and bounded on the closed set $\text{supp}(\mathbf{E})$. Define

$$B = \int_{\text{supp}(\mathbf{E})} (z - \lambda)^{-1} d\mathbf{E}(\lambda) \quad (\in \mathcal{L}(\mathcal{H})).$$

From Proposition 4.13(b) it follows that $(zI - T)B = B(zI - T) = I$, and so $z \notin \sigma(T)$. $\square$

The spectral measure $\mathbf{E} : \mathcal{B}(\sigma(T)) \rightarrow \mathcal{L}(\mathcal{H})$ in the above theorem is called the spectral measure of the normal operator T . For convenience we may use the same expression for the spectral measure obtained by extending $\mathbf{E}$, as zero outside $\sigma(T)$, to the Borel σ -algebra of a suitable set containing $\sigma(T)$. Thus in the following corollary dealing with the spectral representation of a unitary operator, $\mathbf{E}$ could be called the spectral measure of U .

Since the spectrum of a unitary operator $U \in \mathcal{L}(\mathcal{H})$ is contained in $\mathbb{T} = \{z \in \mathbb{C} \mid |z| \leq 1\}$ by Corollary 4.2, the above result yields the following corollary. (The uniqueness part is seen by modifying an argument at the end of the proof of Theorem 4.10.)

Corollary 4.3 *If $U \in \mathcal{L}(\mathcal{H})$ is a unitary operator, there is a unique spectral measure $\mathbf{E} : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$ such that $U = \int_{\mathbb{T}} \lambda d\mathbf{E}(\lambda)$.*

4.14 Exercises

Unless otherwise stated, $\mathcal{H}$ is an arbitrary Hilbert space.

1. Let Ω be a set and $\mathcal{F}$ the set of the singletons $\{x\}$, $x \in \Omega$. Describe
 - (a) the ring generated by $\mathcal{F}$,
 - (b) the σ -algebra generated by $\mathcal{F}$.

2. Let Ω be a set. The *counting measure* $\mu : 2^\Omega \rightarrow [0, \infty]$ is defined as follows: $\mu(A)$ is the number of the elements in the set $A \subset \Omega$, if A is a finite set, and $\mu(A) = \infty$ otherwise. Show that μ is a measure.
3. In the situation of the preceding exercise, show that a function $f : \Omega \rightarrow \mathbb{C}$ is μ -integrable if and only if $(f(x))_{x \in \Omega}$ is a summable family, and if this is the case, then

$$\int_{\Omega} f d\mu = \sum_{x \in \Omega} f(x).$$

4. In the situation of exercise 2, let $f : \Omega \rightarrow \mathbb{C}$ be a μ -integrable function. For each $X \subset \Omega$, define $\nu(X) = \int_X f(x) d\mu(x) (= \sum_{x \in X} f(x))$. Show that $\nu : 2^\Omega \rightarrow \mathbb{C}$ is a complex measure.
5. Describe the total variation measure $|\nu|$ of the complex measure ν defined in the previous exercise.
6. Let Ω be a set and $\mathcal{A} \subset 2^\Omega$ an *algebra*, i.e. a ring having Ω as one of its elements. Let $\text{bfa}(\Omega, \mathcal{A})$ denote the set of the bounded (finitely) additive set functions $\mu : \mathcal{A} \rightarrow \mathbb{C}$. (A function $\mu : \mathcal{A} \rightarrow \mathbb{C}$ is (finitely) additive if $\mu(X \cup Y) = \mu(X) + \mu(Y)$ for any two disjoint sets $X, Y \in \mathcal{A}$.) Define the total variation $|\mu| : \mathcal{A} \rightarrow [0, \infty]$ as in the case of a complex measure. Show that $\text{bfa}(\Omega, \mathcal{A})$ is a Banach space (with the obvious vector space structure) with respect to the norm $\mu \mapsto \|\mu\| = |\mu|(\Omega)$. (In particular, $\|\mu\| < \infty$.)
7. Let Ω be a set and $\mathcal{A} \subset 2^\Omega$ a σ -algebra. Show that $\text{ca}(\Omega, \mathcal{A})$ is a closed subspace of the Banach space $\text{bfa}(\Omega, \mathcal{A})$ defined in the previous exercise.
8. Let Ω be a set and $\mathcal{A} \subset 2^\Omega$ a σ -algebra. Show that the total variation $|\mu|$ of μ is the smallest among the measures $\nu : \mathcal{A} \rightarrow [0, \infty]$ satisfying $|\mu(A)| \leq \nu(A)$ for all $A \in \mathcal{A}$.
9. Let $g : \mathbb{N} \rightarrow [0, \infty)$ be a bounded function. For each $X \subset \mathbb{N}$ and $f \in \ell^2$, denote $E_g(X)f = g\chi_X f$. Show that the map $X \mapsto E_g(X)$ on $2^\mathbb{N}$ is a positive operator measure.
10. Give a necessary and sufficient condition on the function g in the previous exercise for E_g to be
 - (a) projection valued;
 - (b) a semispectral measure;
 - (c) a spectral measure.
11. Give a necessary and sufficient condition on the function g in Exercise 9 for E_g to be countably additive with respect to the norm topology of $\mathcal{L}(\mathcal{H})$. (First state the obvious definition of being countably additive with respect to the norm topology of $\mathcal{L}(\mathcal{H})$.)
12. Is it possible for the positive operator measure in exercise 9 to be both countably additive with respect to the norm topology of $\mathcal{L}(\mathcal{H})$ and
 - (a) projection-valued;
 - (b) a semispectral measure;
 - (c) a spectral measure?

13. Explain in what sense the spectral representation of a compact selfadjoint operator on $\mathcal{H}$ may be expressed in the form

$$A = \int_{\Omega} \lambda d\mathbf{E}(\lambda)$$

where $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure for a suitable set Ω and a σ -algebra $\mathcal{A} \subset 2^{\Omega}$.

14. Prove the following statement in the proof of Proposition 4.13: “It easily follows from (ii) in view of Proposition 4.6(c) that (i) holds for simple functions”.
15. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, so that μ is a not necessarily a finite (positive) measure. Suppose that $h : \Omega \rightarrow [-\infty, \infty] = \overline{\mathbb{R}}$ is an $\mathcal{A}$ -measurable function. Denote

$$B = \{\alpha \in \mathbb{R} \mid \mu(h^{-1}((\alpha, \infty))) = 0\}.$$

Show that $\inf B \in \overline{\mathbb{R}}$ is the smallest $\lambda \in \overline{\mathbb{R}}$ having the property that $h(x) \leq \lambda$ for μ -almost every $x \in \Omega$. We call $\inf B$ the *essential supremum* of h and denote it by $\text{ess sup } h$ or $\text{ess sup}_{x \in \Omega} h(x)$.

16. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space where $\mu(\Omega) > 0$ and consider the Hilbert space $L^2(\mu) = L^2(\Omega, \mathcal{A}, \mu)$. Let $g : \Omega \rightarrow \mathbb{C}$ be an $\mathcal{A}$ -measurable function and assume that g is *essentially bounded*, i.e. $\text{ess sup}_{x \in \Omega} |g(x)| < \infty$. We then denote $\text{ess sup}_{x \in \Omega} |g(x)| = \|g\|_{\infty}$. Show that via the formula $T_g f = gf$ we get a well-defined bounded linear map $T_g : L^2(\mu) \rightarrow L^2(\mu)$ such that $\|T_g\| \leq \|g\|_{\infty}$.
17. (a) Show by an example that in the previous exercise the strict inequality $\|T_g\| < \|g\|_{\infty}$ is possible.
- (b) Show that the equality $\|T_g\| = \|g\|_{\infty}$ holds in the case of the Lebesgue measure μ on $\Omega = \mathbb{R}$.
- (c) Can you abstract some general feature from (b) to guarantee the equality $\|T_g\| = \|g\|_{\infty}$?
18. Let $\mathcal{A} \subset 2^{\Omega}$ be a σ -algebra and $\mu \in \text{ca}(\Omega, \mathcal{A})$. Let $f : \Omega \rightarrow \mathbb{C}$ be a μ -integrable function, and denote $\nu(A) = \int_A f d\mu$ for all $A \in \mathcal{A}$. Show that ν is a complex measure and prove the formula $|\nu|(A) = \int_A |f| d|\mu|$ for all $A \in \mathcal{A}$.
19. Let $\mathcal{A} \subset 2^{\Omega}$ be a σ -algebra and $\mu \in \text{ca}(\Omega, \mathcal{A})$. Show that there is an $\mathcal{A}$ -measurable $|\mu|$ -integrable function $g : \Omega \rightarrow \mathbb{C}$ such that $|g(x)| = 1$ for all $x \in \Omega$ and $\mu(A) = \int_A g(x) d|\mu|(x)$ for all $x \in \Omega$. (Hint: Consider the Hilbert space $L^2(\Omega, \mathcal{A}, |\mu|)$ and the bounded (why?) linear functional $\varphi \mapsto \int_{\Omega} \varphi(x) d\mu(x)$ on it.)
20. Let $\mathcal{A} \subset 2^{\Omega}$ be a σ -algebra and $\mu : \mathcal{A} \rightarrow [0, \infty)$ a (finite positive) measure. Show that there is a Hilbert space $\mathcal{H}$ with a spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ and a vector $\varphi \in \mathcal{H}$ such that $\mu(A) = \langle \varphi | \mathbf{E}(A)\varphi \rangle$ for all $A \in \mathcal{A}$.
21. Let $\mathcal{A} \subset 2^{\Omega}$ be a σ -algebra and $\mu : \mathcal{A} \rightarrow \mathbb{C}$ a complex measure. Show that there is a Hilbert space $\mathcal{H}$ with a spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ and two vectors $\varphi, \psi \in \mathcal{H}$ such that $\mu(A) = \langle \psi | \mathbf{E}(A)\varphi \rangle$ for all $A \in \mathcal{A}$.
22. Complete the details in the proof of Lemma 4.2.
23. Complete the details in Remark 4.8.

24. Prove the commutation claim left as an exercise in Example 4.1.
25. Let Ω be a closed or open subset of $\mathbb{R}^n$ and $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ a positive operator measure. Show that for any $X \in \mathcal{B}(\Omega)$, $E(X)$ is the infimum of the set of the operators $E(U)$ where each U is a relatively open subset of Ω .
26. Let Ω be a closed or open subset of $\mathbb{R}^n$ and $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ a positive operator measure. Show that for any $X \in \mathcal{B}(\Omega)$, $E(X)$ is the supremum of the set of the operators $E(K)$ where each K is a compact subset of Ω .
27. Let Ω be a locally compact Hausdorff space and $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ a positive operator measure. Show that the complex measure $\langle \xi | E(\cdot) \eta \rangle$ is regular for all $\xi, \eta \in \mathcal{H}$ if the measure $\langle \xi | E(\cdot) \xi \rangle$ is regular for all $\xi \in \mathcal{H}$.
28. Complete the details in the proof of Theorem 4.4(c).
29. Complete the details in the proof of Lemma 4.10.
30. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be a bounded function and $T_g : \ell^2 \rightarrow \ell^2$ the bounded linear operator defined by the formula $T_g f = gf$. Show that the spectrum of T_g is the closure of the set $\{g(n) \mid n \in \mathbb{N}\}$.
31. The operator T_g in the previous exercise is clearly normal, since its adjoint is $T_{\bar{g}}$. Describe the spectral measure of T_g .
32. Give an example of an isometric operator whose spectrum is not contained in the set $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$.
33. (a) Let K be a compact subset of $\mathbb{C}$. Show that there is a normal operator T on ℓ^2 such that $\sigma(T) = K$.
 (b) Let $K \subset \mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ be a compact set. Show that there is a unitary operator U on ℓ^2 such that $\sigma(U) = K$.
 (c) State and prove the analogue of (b) for (bounded) selfadjoint operators.
34. (a) Let $U \in \mathcal{L}(\mathcal{H})$ be a normal operator whose spectrum is a subset of $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$. Show that U is unitary.
 (b) State and prove the analogue of (a) for (bounded) selfadjoint operators.
35. Let $T \in \mathcal{L}(\mathcal{H})$ be a normal operator and $T = \int_{\sigma(T)} \lambda dE(\lambda)$ its spectral representation. Show that the series $I + \sum_{n=1}^{\infty} \frac{1}{n!} T^n$ converges in norm, and that if $\exp(T)$ denotes its sum, then $\exp(T) = \int_{\sigma(T)} \exp(\lambda) dE(\lambda)$.
36. Let $T \in \mathcal{L}(\mathcal{H})$ be a normal operator. Suppose that there is some $w \in \mathbb{C}$, $w \neq 0$, such that $\{\lambda w \mid \lambda \geq 0\} \cap \sigma(T) = \emptyset$. Show that there is a normal operator $S \in \mathcal{L}(\mathcal{H})$ such that $T = \exp(S)$.
37. Show that Theorem 3.5 remains valid if the selfadjoint compact operator A is replaced by a normal compact operator. (Then the eigenvalues λ_n are of course not necessarily real.)

References

1. Dunford, N., Schwartz, J.T.: Linear Operators. Part I. Wiley Classics Library. Wiley, New York (1988). Reprint of the 1958 original, A Wiley-Interscience Publication
2. Friedman, A.: Foundations of Modern Analysis. Holt Rinehart and Winston Inc, New York (1970)

3. Halmos, P.R.: Measure Theory, 4th edn. Springer, Heidelberg (1988). Reprint of the 1950 edition
4. Berg, C., Christensen, J.P.R., Ressel, P.: Harmonic Analysis on Semigroups. Graduate Texts in Mathematics. Springer, New York (1984)
5. Ylinen, K.: Positive operator bimeasures and a noncommutative generalization. *Studia Math.* **118**(2), 157–168 (1996)

Chapter 5

Operator Integrals and Spectral Representations: The Unbounded Case

In the traditional Hilbert space quantum mechanics, (generally unbounded) selfadjoint operators are taken as physical observables. In the present work a more comprehensive approach is used: observables are represented as normalised positive operator measures. Selfadjoint operators, however, have independent interest, and their spectral representation in terms of spectral measures shows the connection to the approach to quantum mechanics with positive operator measures. The spectral representation theory of selfadjoint operators utilises the integration of unbounded functions with respect to spectral measures. We develop this theory in the more general setting of positive operator measures, paving the way for important physical applications in the sequel. As the final highlight of the chapter we prove the Stone representation theorem for strongly continuous one-parameter groups of unitary operators.

5.1 Elementary Notes on Unbounded Operators

In many important applications of Hilbert space theory, e.g. in quantum mechanics, one is naturally led to study linear maps that are neither continuous nor defined on the whole of $\mathcal{H}$. In this section we make some basic observations related to such situations.

From now on, we use the term *operator* (in $\mathcal{H}$) for any linear map T whose domain $\mathcal{D}(T)$ is a linear subspace of $\mathcal{H}$ and which satisfies $\mathcal{R}(T) = \{T\varphi \mid \varphi \in \mathcal{D}(T)\} \subset \mathcal{H}$. The *graph* of an operator T in $\mathcal{H}$ is the subset $\mathcal{G}(T) = \{(\varphi, T\varphi) \mid \varphi \in \mathcal{D}(T)\}$ of the Cartesian product $\mathcal{H} \times \mathcal{H}$. For operators S and T , we write $S \subset T$ if $\mathcal{G}(S) \subset \mathcal{G}(T)$. (From a set theoretical point of view, of course T is the same as its graph; the introduction of the notion of graph is a concession to historical usage.) An operator T is said to be *densely defined* if its domain is dense in $\mathcal{H}$, i.e. $\overline{\mathcal{D}(T)} = \mathcal{H}$. We now generalise the concept of the adjoint of an element of $\mathcal{L}(\mathcal{H})$.

Definition 5.1 Let T be a densely defined operator in $\mathcal{H}$. The *adjoint* of T is the operator T^* defined as follows:

- (i) The domain $\mathcal{D}(T^*)$ is the set of those $\psi \in \mathcal{H}$ for which the linear functional $\varphi \mapsto \langle \psi | T\varphi \rangle$ is continuous on $\mathcal{D}(T)$, and
- (ii) whenever $\psi \in \mathcal{D}(T^*)$, $T^*\psi$ is the vector in $\mathcal{H}$ satisfying $\langle T^*\psi | \varphi \rangle = \langle \psi | T\varphi \rangle$ for all $\varphi \in \mathcal{D}(T)$.

The vector $T^*\psi$ mentioned in the definition exists (and is unique) by the Fréchet–Riesz theorem, for if $\psi \in \mathcal{D}(T)$, the map $\varphi \mapsto \langle \psi | T\varphi \rangle$ can be uniquely extended to a continuous linear functional on $\mathcal{H}$.

It is immediately seen that T^* is an operator, i.e. $\mathcal{D}(T^*)$ is a vector subspace of $\mathcal{H}$ and T^* is linear. However, in general T^* need not be densely defined.

Definition 5.2 If S and T are operators in $\mathcal{H}$, then the operators $S + T$ and ST are defined as follows: $\mathcal{D}(S + T) = \mathcal{D}(S) \cap \mathcal{D}(T)$ and $(S + T)\varphi = S\varphi + T\varphi$ for all $\varphi \in \mathcal{D}(S + T)$; $\mathcal{D}(ST) = \{\varphi \in \mathcal{D}(T) \mid T\varphi \in \mathcal{D}(S)\}$ and $(ST)\varphi = S(T\varphi)$ for all $\varphi \in \mathcal{D}(ST)$.

Proposition 5.1 If S , T , and ST are densely defined operators in $\mathcal{H}$, then $T^*S^* \subset (ST)^*$.

Proof Suppose $\psi \in \mathcal{D}(T^*S^*)$. Then $S^*\psi \in \mathcal{D}(T^*)$. Thus the map

$$\varphi \mapsto \langle \psi | ST\varphi \rangle = \langle S^*\psi | T\varphi \rangle$$

on $\mathcal{D}(ST)$ is continuous, so that $\psi \in \mathcal{D}((ST)^*)$. Moreover,

$$\langle T^*S^*\psi | \varphi \rangle = \langle S^*\psi | T\varphi \rangle = \langle \psi | ST\varphi \rangle = \langle (ST)^*\psi | \varphi \rangle$$

for all φ in the dense subspace $\mathcal{D}(ST)$ of $\mathcal{H}$, so that $T^*S^*\psi = (ST)^*\psi$. $\square$

Definition 5.3 Let T be an operator in $\mathcal{H}$. We say that T is *symmetric* if $\langle \psi | T\varphi \rangle = \langle T\psi | \varphi \rangle$ for all $\varphi, \psi \in \mathcal{D}(T)$. If T is densely defined and $T = T^*$, then T is said to be *selfadjoint*.

Thus in our usage a symmetric operator need not be densely defined. It is clear that a densely defined operator T is symmetric if and only if $T \subset T^*$.

Next we study an operator and its adjoint in terms of their graphs. An operator T is said to be a *closed operator* if its graph $\mathcal{G}(T)$ is a closed subspace of the Cartesian product $\mathcal{H} \times \mathcal{H}$ (regarded as the direct sum Hilbert space $\mathcal{H} \oplus \mathcal{H}$ in the usual way). Clearly T is closed if and only if $\varphi \in \mathcal{D}(T)$ and $\psi = T\varphi$, whenever (φ_n) and (ψ_n) are convergent sequences in $\mathcal{H}$ such that $\varphi_n \in \mathcal{D}(T)$ and $\psi_n = T\varphi_n$ for all $n \in \mathbb{N}$ and $\varphi = \lim_{n \rightarrow \infty} \varphi_n$, $\psi = \lim_{n \rightarrow \infty} \psi_n$.

If the closure (in $\mathcal{H} \times \mathcal{H}$) of the graph of an operator T has the property that it can contain the vector pairs (φ, ξ) and (φ, η) only if $\xi = \eta$, then an easy argument shows that this closure is the graph of a linear map defined on a vector subspace of

$\mathcal{H}$. We then denote this linear map by $\overline{T}$ and call it the *closure* of the operator T . In this situation T is said to be *closable*. Clearly T is closable if and only if there is some closed operator S such that $T \subset S$.

Proposition 5.2 *Let T be a densely defined operator in $\mathcal{H}$.*

- (a) *The graph of T^* is the orthogonal complement (in the Hilbert space $\mathcal{H} \times \mathcal{H}$) of the set $\{(-T\xi, \xi) \in \mathcal{H} \times \mathcal{H} \mid \xi \in \mathcal{D}(T)\}$.*
- (b) *The adjoint T^* is a closed operator.*
- (c) *The operator T is closable if and only if $\mathcal{D}(T^*)$ is dense in which case $\overline{T} = T^{**}$.*
- (d) *If T is symmetric, then it is closable and thus $\overline{T} = T^{**}$.*
- (e) *If T is selfadjoint, then T is closed.*

Proof The following statements are equivalent for a pair (φ, ψ) of vectors in $\mathcal{H}$: $\langle (\varphi, \psi) \mid (-T\xi, \xi) \rangle = 0$ for all $\xi \in \mathcal{D}(T)$; $\langle \varphi \mid T\xi \rangle = \langle \psi \mid \xi \rangle$ for all $\xi \in \mathcal{D}(T)$; $\varphi \in \mathcal{D}(T^*)$ and $\psi = T^*\varphi$. Thus (a) holds. Since any orthogonal complement is closed, (b) and (e) follow from (a).

To prove (c), suppose first that $\mathcal{D}(T^*)$ is dense. We denote by U the unitary operator on $\mathcal{H} \times \mathcal{H}$ defined by $U(\varphi, \psi) = (-\psi, \varphi)$. In the proof of (a) we saw that $\mathcal{G}(T^*) = (U(\mathcal{G}(T)))^\perp$, and so by using this also for T^* in place of T and noting that $U^{-1} = -U$, we get $\overline{\mathcal{G}(T)} = (-U\mathcal{G}(T^*))^\perp = (U\mathcal{G}(T^*))^\perp = \mathcal{G}(T^{**})$. Conversely, assume that $\mathcal{D}(T^*)$ is not dense and choose $\psi \in \mathcal{D}(T^*)^\perp \setminus \{0\}$. Then $(\psi, 0) \in \mathcal{G}(T^*)^\perp$ which shows that $(U(\mathcal{G}(T^*)))^\perp$ contains $(0, \psi)$ and hence is not the graph of an operator. Since $\overline{\mathcal{G}(T)} = (U(\mathcal{G}(T^*)))^\perp$, we see that T is not closable.

Part (d) follows from (c) as $T \subset T^*$. □

We show next that a selfadjoint operator is maximally symmetric.

Proposition 5.3 *If S is a selfadjoint operator and T a symmetric operator such that $S \subset T$, then $S = T$.*

Proof Since also T is densely defined, T^* exists. If $\varphi \in \mathcal{D}(S)$ and $\psi \in \mathcal{D}(T^*)$, then $\langle T^*\psi \mid \varphi \rangle = \langle \psi \mid T\varphi \rangle = \langle \psi \mid S\varphi \rangle$, so that also $\psi \in \mathcal{D}(S^*)$ and $S^*\psi = T^*\psi$. We thus have $T \subset T^* \subset S^* = S$, and so $S = T$. □

A densely defined symmetric operator T (which is closable by Proposition 5.2 (d)) is called *essentially selfadjoint* if its closure $\overline{T} = T^{**}$ is selfadjoint. If T is essentially selfadjoint, then by Proposition 5.3 it has only one selfadjoint extension, since any selfadjoint extension is also an extension of $\overline{T}$. In typical applications such a T appears as the sum of two selfadjoint operators (or the restriction of the sum to a dense subspace).

5.2 Integration of Unbounded Functions with Respect to Positive Operator Measures

In this section we extend the definition of the operator integral, which we previously only studied for bounded functions, to the case where the integrand is not necessarily bounded. Then the integral $L(f, E)$ is not necessarily a bounded operator.

Throughout this section, unless specified otherwise, we assume that $\mathcal{A} \subset 2^\Omega$ is a σ -algebra and $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a positive operator measure. Let $f : \Omega \rightarrow \mathbb{C}$ be an $\mathcal{A}$ -measurable function. We let $\mathcal{D}(f, E)$, or just $\mathcal{D}(f)$ if there is no ambiguity about E , denote the set of those $\varphi \in \mathcal{H}$ for which f is integrable with respect to the complex measure $E_{\psi, \varphi}$ whenever $\psi \in \mathcal{H}$.

Lemma 5.1 *Suppose that $\varphi \in \mathcal{D}(f, E)$. Then there is a unique vector $L(f, E)\varphi \in \mathcal{H}$ such that*

$$\langle \psi | L(f, E)\varphi \rangle = \int f dE_{\psi, \varphi}$$

for all $\psi \in \mathcal{H}$.

Proof Choose a sequence (f_n) of simple functions converging pointwise to f , with $|f_n| \leq |f|$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we let α_n denote the integral (with the obvious definition) of f_n with respect to the additive set function $X \mapsto E(X)\varphi$. Consider the bounded linear functionals $h_n : \mathcal{H} \rightarrow \mathbb{C}$ defined by $h_n(\psi) = \langle \psi | \alpha_n \rangle$. It follows from the dominated convergence theorem that the sequence $(h_n(\psi))$ converges to the complex conjugate of $\int f dE_{\psi, \varphi}$ for every $\psi \in \mathcal{H}$. The principle of uniform boundedness can then be used to show that the function $\psi \mapsto \overline{\int_\Omega f dE_{\psi, \varphi}}$ is a bounded linear functional on $\mathcal{H}$, and so the claim follows from the Fréchet–Riesz theorem. $\square$

Definition 5.4 In the situation considered above, we call $L(f, E)$ (i.e. the map $\varphi \mapsto L(f, E)\varphi$) the *integral* of f with respect to E . We may also denote $L(f, E)$ by $E[f]$, or by $L(f)$ if E is clear from the context.

Proposition 5.4 *The set $\mathcal{D}(f, E)$ is a vector subspace of $\mathcal{H}$, and the map $L(f, E) : \mathcal{D}(f, E) \rightarrow \mathcal{H}$ defined in Lemma 5.1 is linear.*

Proof Suppose now that $\xi, \eta \in \mathcal{D}(f, E)$ and $\psi \in \mathcal{H}$. From the definition of the total variation it follows that $|E_{\psi, \xi+\eta}|(X) \leq |E_{\psi, \xi}|(X) + |E_{\psi, \eta}|(X)$ for any $X \in \mathcal{A}$. Thus from the monotone convergence theorem (applied to a suitable increasing sequence of nonnegative simple functions) it follows that a nonnegative measurable function that is integrable with respect to $|E_{\psi, \xi}|$ and $|E_{\psi, \eta}|$ is also $|E_{\psi, \xi+\eta}|$ -integrable. In particular $|f|$, and hence f , is $|E_{\psi, \xi+\eta}|$ -integrable. The equality $\int_\Omega f dE_{\psi, \xi+\eta} = \int_\Omega f dE_{\psi, \xi} + \int_\Omega f dE_{\psi, \eta}$ then follows from the dominated convergence theorem as is obvious for f_n in place of f . The scalar multiplication part is obvious. $\square$

We now consider a different, in general more restricted, approach to operator integrals. For projection valued measures, however, the approaches will in the next section be shown to coincide. Let $\tilde{\mathcal{D}}(f, \mathbf{E})$, or just $\tilde{\mathcal{D}}(f)$, denote the set of those $\varphi \in \mathcal{H}$ for which $|f|^2$ is $\mathbf{E}_{\varphi, \varphi}$ -integrable.

Proposition 5.5 *The set $\tilde{\mathcal{D}}(f, \mathbf{E})$ is a vector subspace of $\mathcal{H}$ contained in $\mathcal{D}(f, \mathbf{E})$. Moreover, if $\varphi \in \tilde{\mathcal{D}}(f, \mathbf{E})$, then*

$$\int |f| d\mathbf{E}_{\psi, \varphi} \leq \sqrt{\|\mathbf{E}(\Omega)\|} \|\psi\| \sqrt{\int |f|^2 d\mathbf{E}_{\varphi, \varphi}},$$

for all $\psi \in \mathcal{H}$, and the map $\psi \mapsto \int f d\mathbf{E}_{\psi, \varphi}$ on $\mathcal{H}$ is a bounded linear functional whose norm is at most $\sqrt{\|\mathbf{E}(\Omega)\|} \sqrt{\int |f|^2 d\mathbf{E}_{\varphi, \varphi}}$.

Proof Choose a sequence (f_n) of simple functions converging pointwise to f , with $|f_n| \leq |f|$ for all $n \in \mathbb{N}$, and let $f_n = \sum_{k=1}^p c_k \chi_{X_k}$, with $X_1, \dots, X_p \in \mathcal{A}$ constituting a partition of Ω . For each $k = 1, \dots, p$, let $Y_{k1}, \dots, Y_{kr_k} \in \mathcal{A}$ form a partition of X_k . Then for all $\varphi, \psi \in \mathcal{H}$,

$$|\mathbf{E}_{\psi, \varphi}(Y_{jk})| = |\langle \psi | \mathbf{E}(Y_{jk}) \varphi \rangle| \leq \|\mathbf{E}(Y_{jk})^{1/2} \psi\| \|\mathbf{E}(Y_{jk})^{1/2} \varphi\|,$$

which gives

$$\begin{aligned} \sum_{k=1}^p |c_k| \sum_{j_k=1}^{r_k} |\mathbf{E}_{\psi, \varphi}(Y_{jk})| &\leq \sum_{k=1}^p |c_k| \sum_{j_k=1}^{r_k} \|\mathbf{E}(Y_{jk})^{1/2} \psi\| \|\mathbf{E}(Y_{jk})^{1/2} \varphi\| \\ &\leq \sqrt{\sum_{k=1}^p \sum_{j_k=1}^{r_k} \|\mathbf{E}(Y_{jk})^{1/2} \psi\|^2} \sqrt{\sum_{k=1}^p \sum_{j_k=1}^{r_k} |c_k|^2 \|\mathbf{E}(Y_{jk})^{1/2} \varphi\|^2} \\ &= \sqrt{\sum_{k=1}^p \sum_{j_k=1}^{r_k} \langle \psi | \mathbf{E}(Y_{jk}) \psi \rangle} \sqrt{\sum_{k=1}^p \sum_{j_k=1}^{r_k} |c_k|^2 \langle \varphi | \mathbf{E}(Y_{jk}) \varphi \rangle} \\ &= \sqrt{\sum_{k=1}^p \langle \psi | \mathbf{E}(X_k) \psi \rangle} \sqrt{\sum_{k=1}^p |c_k|^2 \langle \varphi | \mathbf{E}(X_k) \varphi \rangle} \\ &= \sqrt{\langle \psi | \mathbf{E}(\Omega) \psi \rangle} \sqrt{\sum_{k=1}^p |c_k|^2 \langle \varphi | \mathbf{E}(X_k) \varphi \rangle} \\ &\leq \sqrt{\|\mathbf{E}(\Omega)\|} \|\psi\| \sqrt{\int |f_n|^2 d\mathbf{E}_{\varphi, \varphi}}. \end{aligned}$$

Taking the suprema over all finite partitions of the sets X_k one gets

$$\int |f_n| d|\mathbf{E}_{\psi, \varphi}| \leq \sqrt{\|\mathbf{E}(\Omega)\|} \|\psi\| \sqrt{\int |f_n|^2 d\mathbf{E}_{\varphi, \varphi}},$$

where $|\mathbf{E}_{\psi, \varphi}|$ denotes the total variation of $\mathbf{E}_{\psi, \varphi}$. If $\varphi \in \widetilde{\mathcal{D}}(f)$, using the dominated convergence theorem one obtains $\lim_{n \rightarrow \infty} \int |f_n|^2 d\mathbf{E}_{\varphi, \varphi} = \int |f|^2 d\mathbf{E}_{\varphi, \varphi}$. Hence, by Fatou's lemma, $|f|$ is $|\mathbf{E}_{\psi, \varphi}|$ -integrable, and thus $\mathbf{E}_{\psi, \varphi}$ -integrable, showing that $\varphi \in \mathcal{D}(f)$. This also proves the inequality stated in the proposition, and the last claim is proved by observing that

$$\begin{aligned} \left| \int f d\mathbf{E}_{\psi, \varphi} \right| &\leq \int |f| d|\mathbf{E}_{\psi, \varphi}| \leq \lim_{n \rightarrow \infty} \int |f_n| d|\mathbf{E}_{\psi, \varphi}| \\ &\leq \sqrt{\|\mathbf{E}(\Omega)\|} \|\psi\| \lim_{n \rightarrow \infty} \sqrt{\int |f_n|^2 d\mathbf{E}_{\varphi, \varphi}} \\ &= \sqrt{\|\mathbf{E}(\Omega)\|} \|\psi\| \sqrt{\int |f|^2 d\mathbf{E}_{\varphi, \varphi}}. \end{aligned}$$

Let now $\varphi, \psi \in \mathcal{H}$, $c, d \in \mathbb{C}$, and $X \in \mathcal{A}$. Denoting $\xi = c\varphi + d\psi$ we have

$$\begin{aligned} \langle \xi | \mathbf{E}(X) \xi \rangle &= \|\mathbf{E}(X)^{1/2} \xi\|^2 \leq (|c| \|\mathbf{E}(X)^{1/2} \varphi\| + |d| \|\mathbf{E}(X)^{1/2} \psi\|)^2 \\ &\leq 2|c|^2 \langle \varphi | \mathbf{E}(X) \varphi \rangle + 2|d|^2 \langle \psi | \mathbf{E}(X) \psi \rangle, \end{aligned}$$

which implies that $\widetilde{\mathcal{D}}(f)$ is a linear subspace. $\square$

Remark 5.1 If $\varphi \in \widetilde{\mathcal{D}}(f)$, then the Fréchet–Riesz theorem in conjunction with the above proposition may be used to give a quick proof of the existence of what we have denoted by $L(f, \mathbf{E})\varphi$, independently of the uniform boundedness principle used in the proof of Lemma 5.1. $\triangleleft$

Remark 5.2 In general, the inclusion $\widetilde{\mathcal{D}}(f) \subset \mathcal{D}(f)$ in Proposition 5.5 may be strict. For example, let $\mu : \mathcal{A} \rightarrow [0, 1]$ be a probability measure, and consider the positive operator measure $X \mapsto \mathbf{E}(X) = \mu(X)I$. Take any measurable f which is μ -integrable and such that $|f|^2$ is not μ -integrable. Then $\{0\} = \widetilde{\mathcal{D}}(f) \neq \mathcal{D}(f) = \mathcal{H}$. Choosing an f which is not μ -integrable one gets the extreme case of $\mathcal{D}(f)$ being the null space. $\triangleleft$

Proposition 5.6 (a) *If f is real valued, then $L(f, \mathbf{E})$ is symmetric, that is, for any $\varphi, \psi \in \mathcal{D}(f, \mathbf{E})$, $\langle \psi | L(f, \mathbf{E})\varphi \rangle = \langle L(f, \mathbf{E})\psi | \varphi \rangle$.*
 (b) *If $\mathcal{D}(f, \mathbf{E})$ is a dense subspace of $\mathcal{H}$, then the adjoint $L(f, \mathbf{E})^*$ is an extension of $L(\overline{f}, \mathbf{E})$.*

Proof (a) Let (f_n) be a sequence of real simple functions converging pointwise to f and satisfying $|f_n| \leq |f|$. By the dominated convergence theorem

$$\begin{aligned}
\langle \psi | L(f)\varphi \rangle &= \lim_{n \rightarrow \infty} \langle \psi | L(f_n)\varphi \rangle = \lim_{n \rightarrow \infty} \langle L(f_n)^*\psi | \varphi \rangle \\
&= \lim_{n \rightarrow \infty} \langle L(f_n)\psi | \varphi \rangle = \langle L(f)\psi | \varphi \rangle,
\end{aligned}$$

where we have used the obvious fact that for a (simple) bounded real function f_n the operator $L(f_n)$ is selfadjoint.

(b) Since f is measurable, $\bar{f}$ is also measurable, and the domains of the operators $L(f)$ and $L(\bar{f})$ are the same, $\mathcal{D}(f) = \mathcal{D}(\bar{f})$. If $\mathcal{D}(f)$ is dense in $\mathcal{H}$, then the adjoint of $L(f)$ is defined. For each $n \in \mathbb{N}$, let $g_n(x) = f(x)$ if $|f(x)| \leq n$, and $g_n(x) = 0$ otherwise. Since g_n is bounded, $L(g_n)^* = L(\bar{g}_n)$. By the dominated convergence theorem we may write for all $\psi, \varphi \in \mathcal{D}(f) = \mathcal{D}(\bar{f})$

$$\begin{aligned}
\langle \psi | L(f)\varphi \rangle &= \lim_{n \rightarrow \infty} \langle \psi | L(g_n)\varphi \rangle \\
&= \lim_{n \rightarrow \infty} \langle L(\bar{g}_n)\psi | \varphi \rangle = \langle L(\bar{f})\psi | \varphi \rangle
\end{aligned}$$

which shows that $\psi \in \mathcal{D}(L(f)^*)$ and $L(\bar{f})\psi = L(f)^*\psi$, that is, $L(\bar{f}) \subset L(f)^*$. $\square$

5.3 Integration of Unbounded Functions with Respect to Projection Valued Measures

Throughout this section, $\mathcal{A} \subset 2^\Omega$ is a σ -algebra and $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is assumed to be a projection valued measure. We first show that in this case the two approaches to integration described in the previous section are actually equivalent.

Proposition 5.7 *In the case of a projection valued measure $\mathbf{E}$, we have $\tilde{\mathcal{D}}(f, \mathbf{E}) = \mathcal{D}(f, \mathbf{E})$ for any measurable function f , and this subspace is dense in $\mathcal{H}$.*

Proof We already know that $\tilde{\mathcal{D}}(f, \mathbf{E}) \subset \mathcal{D}(f, \mathbf{E})$ (see Proposition 5.5). Now suppose that $\varphi \in \mathcal{D}(f)$. Choose a sequence (f_n) of simple functions converging pointwise to f , with $|f_n| \leq |f|$ for all $n \in \mathbb{N}$. For a fixed $n \in \mathbb{N}$ we write $f_n = \sum_{k=1}^p c_k \chi_{X_k}$, where $X_1, \dots, X_p \in \mathcal{A}$ are disjoint sets with union Ω . Then a simple calculation based on the fact that the projections $\mathbf{E}(X_j)$ are pairwise orthogonal (by Proposition 4.6 (c)) shows that

$$\int |f_n|^2 d\mathbf{E}_{\varphi, \varphi} = \left\| \sum_{k=1}^p c_k \mathbf{E}(X_k)\varphi \right\|^2 = \left\| \int f_n d\mathbf{E}(\cdot)\varphi \right\|^2$$

with the obvious definition of the integral $\int f_n d\mathbf{E}(\cdot)\varphi$ of the simple function f_n with respect to the additive vector valued set function $X \mapsto \mathbf{E}(X)\varphi$ on $\mathcal{A}$.

For any ψ the dominated convergence theorem shows that the sequence of the numbers

$$\left\langle \psi \left| \int f_n d\mathbf{E}(\cdot)\varphi \right. \right\rangle = \int f_n d\mathbf{E}_{\psi, \varphi}$$

tends to the limit $\int f d\mathbf{E}_{\psi, \varphi}$ as $n \rightarrow \infty$. It thus follows from the uniform boundedness principle that the sequence of the integrals $\int |f_n|^2 d\mathbf{E}_{\varphi, \varphi}$ is bounded, and so Fatou's lemma implies that $|f|^2$ is integrable with respect to $\mathbf{E}_{\varphi, \varphi}$.

We now show that $\widetilde{\mathcal{D}}(f)$ is dense in $\mathcal{H}$. Take some $\varphi \in \mathcal{H}$. Write $\varphi = \mathbf{E}(\Omega)\varphi + \psi$ where $\psi \perp \mathbf{E}(\Omega)(\mathcal{H})$. Since $\mathbf{E}(\Omega)\psi = 0$, it is clear that $\psi \in \widetilde{\mathcal{D}}(f)$, so that it is enough to show that $\mathbf{E}(\Omega)\varphi$ can be approximated by vectors from $\widetilde{\mathcal{D}}(f)$. For each $n \in \mathbb{N}$ denote $A_n = \{\omega \in \Omega \mid |f(\omega)| \leq n\}$ and $\varphi_n = \mathbf{E}(A_n)\varphi$. As (A_n) is an increasing sequence, $\mathbf{E}(\Omega)\varphi = \lim_{n \rightarrow \infty} \varphi_n$ (see Remark 4.4(b)). But $\varphi_n \in \widetilde{\mathcal{D}}(f)$, since $\mathbf{E}_{\varphi_n, \varphi_n}(\Omega \setminus A_n) = \langle \mathbf{E}(A_n)\varphi \mid (\mathbf{E}(\Omega) - \mathbf{E}(A_n))\mathbf{E}(A_n)\varphi \rangle = 0$, and f is bounded on A_n . $\square$

In the sequel we use the notation $\mathcal{D}_f$ for the space $\widetilde{\mathcal{D}}(f, \mathbf{E}) = \mathcal{D}(f, \mathbf{E})$ for any measurable function $f : \Omega \rightarrow \mathbb{C}$.

Lemma 5.2 *If $f : \Omega \rightarrow \mathbb{C}$ is a measurable function, then*

$$\|L(f)\varphi\|^2 = \int_{\Omega} |f|^2 d\mathbf{E}_{\varphi, \varphi}$$

for all $\varphi \in \mathcal{D}_f$.

Proof If $h = \sum_{i=1}^m \alpha_i \chi_{A_i}$ is a simple function where the sets $A_i \in \mathcal{A}$ are disjoint and their union is Ω , then

$$\begin{aligned} \|L(h)\varphi\|^2 &= \left\langle \sum_{i=1}^m \alpha_i \mathbf{E}(A_i)\varphi \left| \sum_{i=1}^m \alpha_i \mathbf{E}(A_i)\varphi \right. \right\rangle = \sum_{i=1}^m |\alpha_i|^2 \|\mathbf{E}(A_i)\varphi\|^2 \\ &= \sum_{i=1}^m |\alpha_i|^2 \langle \varphi \mid \mathbf{E}(A_i)\varphi \rangle = \int_{\Omega} |h|^2 d\mathbf{E}_{\varphi, \varphi}. \end{aligned}$$

The claim follows from this observation in the case when f is bounded, for we can approximate such an f uniformly by simple functions (see Lemma 4.7) and then apply Proposition 4.11. In the general case we define for all $n \in \mathbb{N}$ $f_n(\omega) = f(\omega)$ if $|f(\omega)| \leq n$, and $f_n(\omega) = 0$ if $|f(\omega)| > n$. Since f_n is bounded, we have $\mathcal{D}_{f-f_n} = \mathcal{D}_f$, and using Proposition 5.5 and the dominated convergence theorem we find that for all $\varphi \in \mathcal{D}_f$

$$\begin{aligned} \|L(f)\varphi - L(f_n)\varphi\| &= \sup_{\|\psi\| \leq 1} |\langle \psi \mid L(f - f_n)\varphi \rangle| \leq \sup_{\|\psi\| \leq 1} \int_{\Omega} |f - f_n| d|\mathbf{E}_{\psi, \varphi}| \\ &\leq \left(\int_{\Omega} |f - f_n|^2 d\mathbf{E}_{\varphi, \varphi} \right)^{\frac{1}{2}} \rightarrow 0 \end{aligned} \tag{5.1}$$

when $n \rightarrow \infty$. Since the claim holds for each f_n , for each $\varphi \in \mathcal{D}_f$ the dominated convergence theorem shows that

$$\|L(f)\varphi\|^2 = \lim_{n \rightarrow \infty} \|L(f_n)\varphi\|^2 = \lim_{n \rightarrow \infty} \int_{\Omega} |f_n|^2 d\mathbf{E}_{\varphi, \varphi} = \int_{\Omega} |f|^2 d\mathbf{E}_{\varphi, \varphi}. \quad \square$$

Proposition 5.8 *If $f, g : \Omega \rightarrow \mathbb{C}$ are measurable functions, then $L(f)L(g) \subset L(fg)$ and $\mathcal{D}(L(f)L(g)) = \mathcal{D}_g \cap \mathcal{D}_{fg}$; in particular, $L(f)L(g) = L(fg)$ if and only if $\mathcal{D}_{fg} \subset \mathcal{D}_g$.*

Proof Assume first that f is bounded. Then $\mathcal{D}_g \subset \mathcal{D}_{fg}$. Define $g_n(\omega) = g(\omega)$ if $|g(\omega)| \leq n$, and $g_n(\omega) = 0$ if $|g(\omega)| > n$, $n \in \mathbb{N}$. Suppose that $\varphi \in \mathcal{D}_g$. Now $L(f)$ is bounded, and since by (5.1) we have $L(g)\varphi = \lim_{n \rightarrow \infty} L(g_n)\varphi$ and $L(fg)\varphi = \lim_{n \rightarrow \infty} L(fg_n)\varphi$, Proposition 4.13(b) implies

$$L(f)L(g)\varphi = \lim_{n \rightarrow \infty} L(f)L(g_n)\varphi = \lim_{n \rightarrow \infty} L(fg_n)\varphi = L(fg)\varphi. \quad (5.2)$$

If $\psi = L(g)\varphi$, we thus get using Lemma 5.2,

$$\int_{\Omega} |f|^2 d\mathbf{E}_{\psi, \psi} = \|L(f)\psi\|^2 = \|L(fg)\varphi\|^2 = \int_{\Omega} |fg|^2 d\mathbf{E}_{\varphi, \varphi}. \quad (5.3)$$

Let now f be arbitrary (not necessarily bounded). Since

$$\mathcal{D}(L(f)L(g)) = \{\varphi \in \mathcal{D}_g \mid L(g)\varphi \in \mathcal{D}(L(f))\}$$

and since by the monotone convergence theorem (5.3) implies, when applied to the functions f_n used in the proof of Lemma 5.2, that for a vector $\varphi \in \mathcal{D}_g$ we have $L(g)\varphi \in \mathcal{D}_f$ if and only if $\varphi \in \mathcal{D}_{fg}$, we see that $\mathcal{D}(L(f)L(g)) = \mathcal{D}_g \cap \mathcal{D}_{fg}$. Suppose now that $\varphi \in \mathcal{D}_g \cap \mathcal{D}_{fg}$, $\psi = L(g)\varphi$, and f_n is as in the proof of Lemma 5.2. It follows from the dominated convergence theorem that $f_n \rightarrow f$ in $L^2(\mathbf{E}_{\psi, \psi})$ (for $|f - f_n|^2 \leq |f|^2$) and $f_n g \rightarrow fg$ in $L^2(\mathbf{E}_{\varphi, \varphi})$. Hence because by (5.2) $L(f_n)L(g)\varphi = L(f_n g)\varphi$, we find using (5.1) that $L(f)L(g)\varphi = L(f)\psi = \lim_{n \rightarrow \infty} L(f_n)\psi = L(f_n g)\varphi = L(fg)\varphi$. $\square$

Proposition 5.9 *For any measurable function $f : \Omega \rightarrow \mathbb{C}$ we have $L(f)^* = L(\bar{f})$ and $L(f)L(f)^* = L(|f|^2) = L(f)^*L(f)$.*

Proof Let $\varphi \in \mathcal{D}_f$ and $\psi \in \mathcal{D}_{\bar{f}} = \mathcal{D}_f$. If f_n is as in the proof of Lemma 5.2, using (5.1) and Proposition 4.12(b) we get $\langle \psi \mid L(f)\varphi \rangle = \lim_{n \rightarrow \infty} \langle \psi \mid L(f_n)\varphi \rangle = \lim_{n \rightarrow \infty} \langle L(\bar{f}_n)\psi \mid \varphi \rangle = \langle L(\bar{f})\psi \mid \varphi \rangle$. It follows that $\psi \in \mathcal{D}(L(f)^*)$ and $L(\bar{f})\psi = L(f)^*\psi$. To prove the equality $L(\bar{f}) = L(f)^*$ it now suffices to show that $\mathcal{D}(L(f)^*) \subset \mathcal{D}_{\bar{f}}$. Let $\xi \in \mathcal{D}(L(f)^*)$ and $\eta = L(f)^*\xi$. As $f_n = fh_n$ where h_n is the characteristic function of the set $\{\omega \in \Omega \mid |f(\omega)| \leq n\}$, by Proposition 5.8 we get $L(f_n) = L(f)L(h_n)$, since $\mathcal{D}_{h_n} = \mathcal{H}$. Since $L(h_n)$ is a projection, using

Propositions 5.1 and 4.12(b) we find that $L(h_n)L(f)^* \subset [L(f)L(h_n)]^* = L(f_n)^* = L(\overline{f_n})$, and so

$$L(h_n)\eta = L(\overline{f_n})\xi \quad \text{for all } n \in \mathbb{N}. \quad (5.4)$$

Since $|h_n| \leq 1$, (5.4) and Lemma 5.2 together imply

$$\begin{aligned} \int_{\Omega} |f_n|^2 d\mathbf{E}_{\xi, \xi} &= \int_{\Omega} |\overline{f_n}|^2 d\mathbf{E}_{\xi, \xi} = \|L(\overline{f_n})\xi\|^2 = \|L(h_n)\eta\|^2 \\ &= \int_{\Omega} h_n d\mathbf{E}_{\eta, \eta} \leq \mathbf{E}_{\eta, \eta}(\Omega), \end{aligned}$$

and so by the monotone convergence theorem we have $\xi \in \mathcal{D}_f$. The latter equations in the claim follow from the above by Proposition 5.8, since $\mathcal{D}_{f\overline{f}} \subset \mathcal{D}_f$ in view of the fact that

$$\int_{\Omega} |f|^2 d\mathbf{E}_{\varphi, \varphi} \leq \left(\int_{\Omega} 1^2 d\mathbf{E}_{\varphi, \varphi} \right)^{\frac{1}{2}} \left(\int_{\Omega} |f|^4 d\mathbf{E}_{\varphi, \varphi} \right)^{\frac{1}{2}} < \infty$$

if $\varphi \in \mathcal{D}_{f\overline{f}}$. □

We still mention some immediate consequences of the preceding discussion.

Corollary 5.1 *Let $f, g : \Omega \rightarrow \mathbb{C}$ be $\mathcal{A}$ -measurable functions.*

- (a) *The operator $L(f)$ is closed.*
- (b) *If g is bounded, then $L(g)(\mathcal{D}_f) \subset \mathcal{D}_f$.*
- (c) *If g is bounded and vanishes outside the set $\{\omega \in \Omega \mid |f(\omega)| \leq M\}$ for some $M \in [0, \infty)$, then $L(g)(\mathcal{H}) \subset \mathcal{D}_f$.*

Proof (a) This follows from Proposition 5.9 (applied to $\overline{f}$) since any adjoint is closed by Proposition 5.2.

(b) Since $\mathcal{D}_{fg} \supset \mathcal{D}_f$ and $\mathcal{D}_g = \mathcal{H}$, Proposition 5.8 shows that $\mathcal{D}(L(f)L(g)) \supset \mathcal{D}_f$, implying the claim.

(c) Now fg is bounded, so that $\mathcal{D}(L(f)L(g)) = \mathcal{H}$. □

5.4 The Cayley Transform

If $z \in \mathbb{C} \setminus \{-i\}$, we define $f(z) = \frac{z-i}{z+i}$. In this way we get a homeomorphism $f : \mathbb{C} \setminus \{-i\} \rightarrow \mathbb{C} \setminus \{1\}$ whose inverse is given by the formula $z = f^{-1}(w) = i \frac{1+w}{1-w}$. Since

$$\operatorname{Im} z = \frac{1}{2i}(z - \overline{z}) = \frac{1}{2} \left(\frac{1+w}{1-w} + \frac{1+\overline{w}}{1-\overline{w}} \right) = \frac{1 - |w|^2}{|1-w|^2},$$

f defines a homeomorphism between the real axis and the set $\mathbb{T} \setminus \{1\}$ where $\mathbb{T} = \{w \in \mathbb{C} \mid |w| = 1\}$ is the *unit circle*. This simple observation provides a blueprint for

the definition of the so-called Cayley transform, which will turn out to be an efficient tool in the study of a (generally unbounded) selfadjoint operator: such an operator can be transformed into a unitary operator, for which we already have a spectral representation. Before defining the Cayley transform we make some preparations. Recall that an operator T in $\mathcal{H}$ is *isometric* if $\|T\varphi\| = \|\varphi\|$ for all $\varphi \in \mathcal{D}(T)$.

Proposition 5.10 *Let S be a symmetric operator in $\mathcal{H}$.*

- (a) *For all $\varphi \in \mathcal{D}(S)$ we have $\|(S \pm iI)\varphi\|^2 = \|\varphi\|^2 + \|S\varphi\|^2$.*
- (b) *The map $S + iI : \mathcal{D}(S) \rightarrow \mathcal{R}(S + iI)$ is a bijection and its inverse $(S + iI)^{-1} : \mathcal{R}(S + iI) \rightarrow \mathcal{D}(S)$ is continuous. Moreover, $\mathcal{R}(S + iI)$ is closed if S is a closed operator.*
- (c) *The map $V = (S - iI)(S + iI)^{-1}$ is isometric, $\mathcal{D}(V) = \mathcal{R}(S + iI)$ and $\mathcal{R}(V) = \mathcal{R}(S - iI)$. Moreover, $\mathcal{R}(V)$ is closed if S is a closed operator.*
- (d) *The map $I - V$ is injective, $\mathcal{D}(S) = \mathcal{R}(I - V)$, and $S = i(I + V)(I - V)^{-1}$.*

Proof (a) Since S is symmetric we have $\|(S \pm iI)\varphi\|^2 = \langle S\varphi | S\varphi \rangle \pm \langle S\varphi | i\varphi \rangle \pm \langle i\varphi | S\varphi \rangle + \langle \varphi | \varphi \rangle = \|S\varphi\|^2 + \|\varphi\|^2$ for all $\varphi \in \mathcal{D}(S)$.

(b) If $(S + iI)\varphi = 0$, it follows from (a) that $\varphi = 0$. Thus $S + iI$ is an injection. Moreover,

$$\|\varphi\| \leq \|(S + iI)\varphi\|, \quad (5.5)$$

and so $(S + iI)^{-1}$ is continuous. If $S\varphi_n + i\varphi_n \rightarrow \psi$ where $\varphi_n \in \mathcal{D}(S)$ for all $n \in \mathbb{N}$, by (5.5) (φ_n) is a Cauchy sequence, so that it converges to some $\varphi \in \mathcal{H}$. If S is closed, it is easily seen that $S + iI$ is also closed, and so $\varphi \in \mathcal{D}(S + iI)$ and $\psi = S\varphi + i\varphi$. Hence $\psi \in \mathcal{R}(S + iI)$, so that $\mathcal{R}(S + iI)$ is closed.

(c) Clearly V is defined in $\mathcal{R}(S + iI)$ (since $\mathcal{D}(S - iI) = \mathcal{D}(S) = \mathcal{D}(S + iI)$) and $\mathcal{R}(V) = \mathcal{R}(S - iI)$. Moreover, by (a) we have $\|(S - iI)\varphi\| = \|(S + iI)\varphi\|$ for all $\varphi \in \mathcal{D}(S) = \mathcal{R}((S + iI)^{-1})$ implying that V is isometric. As in (b), we see that $\mathcal{R}(V) = \mathcal{R}(S - iI)$ is closed if S is a closed operator.

(d) Between the sets $\mathcal{R}((S + iI)^{-1}) = \mathcal{D}(S + iI) = \mathcal{D}(S)$ and $\mathcal{D}(V) = \mathcal{R}(S + iI)$ there is in view of (b) and (c) a bijective correspondence making $\varphi \in \mathcal{D}(S)$ correspond to $\psi \in \mathcal{D}(V)$ where

$$\psi = (S + iI)\varphi, \quad V\psi = (S - iI)\varphi \quad (5.6)$$

or, equivalently,

$$(I - V)\psi = 2i\varphi, \quad (I + V)\psi = 2S\varphi. \quad (5.7)$$

If $(I - V)\psi = 0$, then denoting $\varphi = (S + iI)^{-1}\psi$ we have $(I - V)\psi = 2i\varphi$, so $\varphi = 0$ and hence $\psi = 0$. This shows that $I - V$ is an injection and $\mathcal{R}(I - V) = \mathcal{D}(S)$. In addition $S\varphi = \frac{1}{2}(I + V)\psi = i(I + V)(I - V)^{-1}\varphi$ for all $\varphi \in \mathcal{D}(S) = \mathcal{D}(S + iI)$ by (5.7). $\square$

Definition 5.5 The operator $V = (S - iI)(S + iI)^{-1}$ defined in the above proposition is called the *Cayley transform* of the operator S .

A symmetric operator S can be studied with the help of its Cayley transform. It will turn out that a densely defined symmetric operator is selfadjoint if and only if its Cayley transform is unitary. We prove this with the help of the following result.

Proposition 5.11 *For a symmetric densely defined operator S , the following three conditions are equivalent:*

- (i) S is selfadjoint;
- (ii) S is closed, and zero is the only vector $\psi \in \mathcal{D}(S^*)$ such that $(S^* + iI)\psi = 0$ or $(S^* - iI)\psi = 0$;
- (iii) $\mathcal{R}(S + iI) = \mathcal{R}(S - iI) = \mathcal{H}$.

Proof Assume first (i). According to Proposition 5.2 S is closed. Let now $\psi \in \mathcal{D}(S^*)$ be such that $S^*\psi + i\psi = 0$. Then $S\psi + i\psi = 0$ since S is selfadjoint, and Proposition 5.10 (a) shows that $\psi = 0$. Similarly, $\psi = 0$ if $(S^*\psi - iI)\psi = 0$. Thus (ii) holds.

Next assume (ii). Let $\psi \in \mathcal{R}(S - iI)^\perp$. Then

$$\langle \psi | S\varphi \rangle = \langle \psi | (S - iI)\varphi \rangle + \langle \psi | i\varphi \rangle = \langle -i\psi | \varphi \rangle$$

for all $\varphi \in \mathcal{D}(S)$, so $\psi \in \mathcal{D}(S^*)$ and $S^*\psi = -i\psi$, which implies by (ii) that $\psi = 0$. Therefore by the projection theorem, Theorem 2.4, $\mathcal{R}(S - iI)$ is dense in $\mathcal{H}$. But according to Proposition 5.10 (c) $\mathcal{R}(S - iI)$ is closed, so $\mathcal{R}(S - iI) = \mathcal{H}$. It is similarly seen (using Proposition 5.10 (b)) that $\mathcal{R}(S + iI) = \mathcal{H}$.

Finally assume (iii). Let $\varphi \in \mathcal{D}(S^*)$. Since $\mathcal{R}(S - iI) = \mathcal{H}$, there is a vector $\psi \in \mathcal{D}(S)$ such that $(S - iI)\psi = (S^* - iI)\varphi$. As $S \subset S^*$, we have $\varphi - \psi \in \mathcal{D}(S^*)$ and $(S^* - iI)(\varphi - \psi) = 0$. For any $\xi \in \mathcal{D}(S)$ we have $\langle \varphi - \psi | (S + iI)\xi \rangle = \langle S^*(\varphi - \psi) | \xi \rangle + \langle -i(\varphi - \psi) | \xi \rangle = \langle (S^* - iI)(\varphi - \psi) | \xi \rangle = 0$, so $\varphi = \psi \in \mathcal{D}(S)$, since $\mathcal{R}(S + iI) = \mathcal{H}$. It follows that $\mathcal{D}(S^*) = \mathcal{D}(S)$, and so S is selfadjoint. $\square$

Theorem 5.1 *A symmetric densely defined operator is selfadjoint if and only if its Cayley transform is unitary.*

Proof The Cayley transform is unitary if and only if it is an isometric surjection defined on all of $\mathcal{H}$. Therefore the claim follows from Propositions 5.10 (c) and 5.11. $\square$

5.5 The Spectral Representation of an Unbounded Selfadjoint Operator

We are now in a position to prove the spectral representation theorem for a general, not necessarily bounded, selfadjoint operator. Now the spectral measure will be defined on $\mathcal{B}(\mathbb{R})$, and in general there is no compact set whose complement would be mapped to the null operator. The proof depends on the Cayley transform and the spectral representation of a unitary operator.

Theorem 5.2 *Let A be a selfadjoint operator in $\mathcal{H}$. There is a uniquely determined spectral measure $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ such that*

$$\langle \psi | A\varphi \rangle = \int_{\mathbb{R}} \lambda d\mathbf{E}_{\psi, \varphi}(\lambda) \quad (5.8)$$

for all $\varphi \in \mathcal{D}(A)$ and $\psi \in \mathcal{H}$. Moreover, $\mathcal{D}(A)$ is the set of those $\varphi \in \mathcal{H}$ for which the function $\lambda \mapsto \lambda^2$ on $\mathbb{R}$ is $\mathbf{E}_{\varphi, \varphi}$ -integrable, so that

$$A = \int_{\mathbb{R}} \lambda d\mathbf{E}(\lambda).$$

Proof We first prove the existence part. Let V be the Cayley transform of A . Since V is unitary by Theorem 5.1, Corollary 4.3 yields a spectral measure $\mathbf{E}_1 : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $V = \int_{\mathbb{T}} w d\mathbf{E}_1(w)$. Proposition 4.13 shows that

$$(I - V)\mathbf{E}_1(\{1\}) = \int_{\mathbb{T}} (1 - w)\chi_{\{1\}}(w) d\mathbf{E}_1(w) = 0,$$

and since $I - V$ is an injection by Proposition 5.10 (d), we must have $\mathbf{E}_1(\{1\}) = 0$. Therefore the formula $\tilde{\mathbf{E}}(B) = \mathbf{E}_1(B)$, $B \in \mathcal{B}(\mathbb{T} \setminus \{1\})$, defines a spectral measure $\tilde{\mathbf{E}} : \mathcal{B}(\mathbb{T} \setminus \{1\}) \rightarrow \mathcal{L}(\mathcal{H})$ satisfying

$$\int_{\mathbb{T} \setminus \{1\}} w d\tilde{\mathbf{E}}(w) = V. \quad (5.9)$$

We define $g(w) = i(1 + w)(1 - w)^{-1}$ for all $w \in \mathbb{T} \setminus \{1\}$. Then $g : \mathbb{T} \setminus \{1\} \rightarrow \mathbb{R}$ is a homeomorphism and $g^{-1}(x) = (x - i)(x + i)^{-1}$. We now define $\tilde{\mathbf{E}}(B) = \tilde{\mathbf{E}}(g^{-1}(B))$, $B \in \mathcal{B}(\mathbb{R})$. Then $\tilde{\mathbf{E}} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure, and for a vector $\varphi \in \mathcal{H}$ the function $|g|^2$ is $\tilde{\mathbf{E}}_{\varphi, \varphi}$ -integrable if and only if the function $\lambda \mapsto \lambda^2$ is $\mathbf{E}_{\varphi, \varphi}$ -integrable; we let $\mathcal{D}_g$ denote the set of all such vectors φ . We define $L(g, \tilde{\mathbf{E}}) : \mathcal{D}_g \rightarrow \mathcal{H}$, as usual, to be the operator satisfying

$$\langle \psi | L(g, \tilde{\mathbf{E}})\varphi \rangle = \int_{\mathbb{T} \setminus \{1\}} g d\tilde{\mathbf{E}}_{\psi, \varphi} \quad (5.10)$$

for all $\varphi \in \mathcal{D}_g$, $\psi \in \mathcal{H}$ (see Proposition 5.8). Since g is real valued, the operator $L(g)$ is selfadjoint (Proposition 5.9), and since $g(w)(1 - w) = i(1 + w)$, using Proposition 5.8 we see that

$$L(g)(I - V) = i(I + V); \quad \mathcal{R}(I - V) \subset \mathcal{D}(L(g)). \quad (5.11)$$

On the other hand,

$$A(I - V) = i(I + V); \quad \mathcal{D}(A) = \mathcal{R}(I - V) \subset \mathcal{D}(L(g)) \quad (5.12)$$

by Proposition 5.10 (d). Thus in view of (5.11) and (5.12) $L(g)$ is a selfadjoint extension of A , implying by Proposition 5.3 that $A = L(g)$. Since $g : \mathbb{T} \setminus \{1\} \rightarrow \mathbb{R}$ is a homeomorphism, by (5.10) we get $\langle \psi | A\varphi \rangle = \int_{\mathbb{R}} x d\mathbf{E}_{\psi, \varphi}(x)$ for all $\varphi \in \mathcal{D}(A)$, $\psi \in \mathcal{H}$. Thus (5.8) holds, and also $A = \int_{\mathbb{R}} \lambda d\hat{\mathbf{E}}(\lambda)$, since $\mathcal{D}_g = \mathcal{D}(L(\text{id}, \mathbf{E}))$.

We now prove the uniqueness claim. Let $\hat{\mathbf{E}} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be an arbitrary spectral measure such that $\langle \psi | A\varphi \rangle = \int_{\mathbb{R}} \lambda d\hat{\mathbf{E}}_{\psi, \varphi}$ for all $\varphi \in \mathcal{D}(A)$, $\psi \in \mathcal{H}$. Let $g : \mathbb{T} \setminus \{1\} \rightarrow \mathbb{R}$ be the homeomorphism defined above. The formula $\mathbf{E}_2(X) = \hat{\mathbf{E}}(g(X \setminus \{1\}))$ defines a spectral measure $\mathbf{E}_2 : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$. We define $V_1 = \int_{\mathbb{T}} \lambda d\mathbf{E}_2(\lambda)$ ($\in \mathcal{L}(\mathcal{H})$). Using Proposition 4.14 we see that

$$V_1 = \int_{\mathbb{T} \setminus \{1\}} \lambda d\mathbf{E}_2(\lambda) = \int_{\mathbb{R}} (x - i)(x + i)^{-1} d\hat{\mathbf{E}}(x).$$

Let $\varphi \in \mathcal{D}(A)$. If $\psi \in \mathcal{H}$, by assumption the function $x \mapsto x$ and hence $x \mapsto |x|$ is $\hat{\mathbf{E}}_{\psi, \varphi}$ -integrable (see Proposition 4.5(c)) so that the dominated convergence theorem, Theorem 4.1, implies that $\int_{\mathbb{R}} x d\hat{\mathbf{E}}_{\psi, \varphi}(x) = \lim_{n \rightarrow \infty} \langle \psi | A_n \varphi \rangle$ where

$$A_n = \int_{\mathbb{R}} x \chi_{[-n, n]}(x) d\hat{\mathbf{E}}(x) \quad (\in \mathcal{L}(\mathcal{H})).$$

We also have $\langle \psi | I\varphi \rangle = \langle \psi | \varphi \rangle = \lim_{n \rightarrow \infty} \langle \psi | I_n \varphi \rangle$ where $I_n = \int_{\mathbb{R}} \chi_{[-n, n]} d\hat{\mathbf{E}}$. Hence by Proposition 4.13(b)

$$\begin{aligned} \langle \psi | V_1(A + iI)\varphi \rangle &= \langle V_1^* \psi | (A + iI)\varphi \rangle = \lim_{n \rightarrow \infty} \langle V_1^* \psi | (A_n + iI)\varphi \rangle \\ &= \lim_{n \rightarrow \infty} \langle \psi | V_1(A_n + iI_n)\varphi \rangle \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \chi_{[-n, n]}(x) \frac{t - i}{t + i} (t + i) d\hat{\mathbf{E}}_{\psi, \varphi}(x) \\ &= \langle \psi | (A - iI)\varphi \rangle = \langle \psi | V(A + iI)\varphi \rangle \end{aligned}$$

for all $\psi \in \mathcal{H}$, $\varphi \in \mathcal{D}(A)$ (see Proposition 5.10). From the uniqueness part of Corollary 4.3 it thus follows that $\mathbf{E}_1 = \mathbf{E}_2$, and this implies that $\mathbf{E} = \hat{\mathbf{E}}$. $\square$

We call the spectral measure $\mathbf{E}$ in the preceding theorem *the spectral measure of the selfadjoint operator* A . In the case of a bounded (normal) operator we saw that the spectral measure is actually carried by the spectrum of the operator. In the next section we consider the analogue of this for unbounded selfadjoint operators.

5.6 The Support of the Spectral Measure of a Selfadjoint Operator

We assume in this section that A is a selfadjoint operator in $\mathcal{H}$ and $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is its spectral measure. In accordance with Theorem 5.2, $\mathbf{E}$ is characterised by the condition $A = \int_{\mathbb{R}} \lambda d\mathbf{E}(\lambda)$.

We recall from Definition 4.6 the notion of the support of a positive operator measure: Let U denote the union of all open sets $V \subset \mathbb{R}$ for which $\mathbf{E}(V) = 0$. Then $\mathbf{E}(U) = 0$, and the complement $\mathbb{R} \setminus U$ is the *support* of $\mathbf{E}$, denoted by $\text{supp}(\mathbf{E})$. Clearly $\text{supp}(\mathbf{E})$ is the smallest closed subset of $\mathbb{R}$ mapped to the identity operator I by $\mathbf{E}$. In the following we regard $\mathbb{R}$ in the usual way as a subset of $\mathbb{C}$, so that $\text{supp}(\mathbf{E})$ is also a (closed) subset of $\mathbb{C}$. We denote $B(z, \epsilon) = \{w \in \mathbb{C} \mid |w - z| < \epsilon\}$ for $z \in \mathbb{C}$, $\epsilon > 0$.

Theorem 5.3 *For a number $z \in \mathbb{C}$ the following conditions are equivalent:*

- (i) $\mathbf{E}(\mathbb{R} \cap B(z, \epsilon)) = 0$ for some $\epsilon > 0$;
- (ii) $z \in \mathbb{C} \setminus \text{supp}(\mathbf{E})$;
- (iii) *there is an operator $B \in \mathcal{L}(\mathcal{H})$ such that $B(zI - A) = I|_{\mathcal{D}(A)}$ and $(zI - A)B = I$.*

Proof The implication (i) $\implies$ (ii), as well as its converse, follows immediately from the definition.

Assume now (ii). Then the function $\lambda \mapsto (z - \lambda)^{-1}$ is defined, continuous and bounded on the closed set $\text{supp}(\mathbf{E})$. Define

$$B = \int_{\text{supp}(\mathbf{E})} (z - \lambda)^{-1} d\mathbf{E}(\lambda) \quad (\in \mathcal{L}(\mathcal{H})).$$

From Proposition 5.8 it follows that $(zI - A)B = I$ and $B(zI - A) = I|_{\mathcal{D}(A)}$. Thus (iii) holds.

Now assume (iii). Suppose that (ii) does not hold. Then z is real and (by the equivalence of (i) and (ii)) for each $n \in \mathbb{N}$ $\mathbf{E}((z - 1/n, z + 1/n))$ is a nonzero projection, so there is a $\varphi_n \in \mathcal{H}$ such that $\|\varphi_n\| = 1$ and $\mathbf{E}((z - 1/n, z + 1/n))\varphi_n = \varphi_n$. According to Corollary 5.1 (c) $\varphi_n \in \mathcal{D}(A)$. Now

$$\begin{aligned} 1 &= \|\varphi_n\| = \|B(zI - A)\mathbf{E}((z - 1/n, z + 1/n))\varphi_n\| \\ &= \left\| B \left(\int_{(z-1/n, z+1/n)} (z - \lambda) d\mathbf{E}(\lambda) \right) \varphi_n \right\| \leq \|B\| \frac{1}{n} \|\varphi_n\| \rightarrow 0, \end{aligned}$$

when $n \rightarrow \infty$ (see Propositions 5.8, 4.13). This contradiction shows that $z \notin \text{supp}(\mathbf{E})$. $\square$

The set of those $z \in \mathbb{C}$ for which the condition (iii) in the above theorem does not hold, is called the *spectrum* of A and denoted by $\sigma(A)$. We have thus proved that $\sigma(A) = \text{supp}(\mathbf{E})$. In particular, $\sigma(A) \subset \mathbb{R}$ for the selfadjoint operator A , and the spectral representation of A may be written in the form $A = \int_{\sigma(A)} \lambda d\mathbf{E}(\lambda)$. As in the case of a bounded normal operator, the spectral measure $\mathbf{E}$ restricted to the Borel σ -algebra of $\sigma(A)$ may also be called *the spectral measure of A* . The complement $\mathbb{C} \setminus \sigma(A)$ is called the *resolvent set* of A and denoted by $\rho(A)$.

It is clear that this definition of $\sigma(A)$ and $\rho(A)$ is consistent with that which was used in connection with bounded operators. (See the discussion preceding Theorem 3.4.)

For general (bounded) operators $T \in \mathcal{L}(\mathcal{H})$ we considered the subdivision of the spectrum of T into three parts: $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$. We can generalise this to the case of the not necessarily bounded selfadjoint operator A , but in this case the residual spectrum $\sigma_r(A)$ turns out to be empty. To see this we first prove a lemma.

- Lemma 5.3** (a) *For the selfadjoint operator A and $z \in \mathbb{R}$, denote $A_z = zI - A$. If A_z is an injective operator, then its range is dense in $\mathcal{H}$.*
 (b) *If $T : \mathcal{D}(T) \rightarrow \mathcal{L}(\mathcal{H})$ is a closed operator which is continuous on $\mathcal{D}(T)$, then $\mathcal{D}(T)$ is closed.*

Proof (a) A vector $\psi \in \mathcal{H}$ is orthogonal to $\mathcal{R}(A_z)$ if and only if $\langle \psi | A_z \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}(A_z) = \mathcal{D}(A)$, i.e. if and only if $\psi \in \mathcal{D}((A_z)^*)$ and $(A_z)^* \psi = 0$. Since clearly $A_z = (A_z)^*$, this is equivalent to $\psi \in \mathcal{D}(A_z)$ and $(A_z) \psi = 0$. Since A_z is injective, $\psi = 0$. Thus $\mathcal{R}(A_z)$ is dense in $\mathcal{H}$.

(b) Suppose (φ_n) is a sequence in $\mathcal{D}(T)$ converging to some $\varphi \in \mathcal{H}$. Since T is continuous $(T\varphi_n)$ is a Cauchy sequence, and so there is some $\psi \in \mathcal{H}$ such that $\lim_{n \rightarrow \infty} T\varphi_n = \psi$. Since T is closed, $(\varphi, \psi) \in \mathcal{G}(T)$, so $\varphi \in \mathcal{D}(T)$. $\square$

Proposition 5.12 (a) *A number $z \in \mathbb{C}$ belongs to $\rho(A)$ if and only if $zI - A$ is injective and its inverse $(zI - A)^{-1} : \mathcal{R}(zI - A) \rightarrow \mathcal{D}(A)$ is continuous. If this is the case, then $\mathcal{R}(zI - A) = \mathcal{H}$.*

(b) *If $z \in \sigma(A)$, then one and only one of the following two conditions holds:*

- (i) *there is some $\varphi \in \mathcal{D}(A) \setminus \{0\}$ satisfying $A\varphi = z\varphi$;*
- (ii) *the operator $zI - A$ is injective, its range is dense in $\mathcal{H}$ and the inverse $(zI - A)^{-1} : \mathcal{R}(zI - A) \rightarrow \mathcal{D}(A)$ is discontinuous.*

Proof (a) Suppose $z \in \mathbb{R}$ and $zI - A$ is injective. By Lemma 5.3 (a) $\mathcal{D}(zI - A)$ is dense in $\mathcal{H}$. It is easily seen that $zI - A$ is closed (since A is) and so is its inverse (since the map $(\xi, \eta) \mapsto (\eta, \xi)$ is a homeomorphism on $\mathcal{H} \times \mathcal{H}$). Thus if $(zI - A)^{-1}$ is continuous, its domain is all of $\mathcal{H}$ by Lemma 5.3 (b), and it follows that $z \in \rho(A)$. The rest of the claim is covered by Theorem 5.3.

(b) This follows immediately from (a). $\square$

The set of those points in $\sigma(A)$ which satisfy (i) in Proposition 5.12 (b) is called the *point spectrum* of A . Its elements are the *eigenvalues* of A , and a vector $\varphi \in \mathcal{D}(A) \setminus \{0\}$ satisfying $A\varphi = z\varphi$ is an *eigenvector* belonging to the eigenvalue z . The set $\sigma_c(A) = \sigma(A) \setminus \sigma_p(A)$ is called the *continuous spectrum* of A . Thus $\sigma_c(A)$ is characterised by condition (ii) in Proposition 5.12 (b). In view of Proposition 5.12 this terminology is consistent with the usage of Sect. 3.2. In particular, the residual spectrum of a selfadjoint operator is empty.

Proposition 5.13 *Suppose that $x \in \mathbb{R}$. The following conditions are equivalent:*

- (i) $E(\{x\}) \neq 0$;
- (ii) $E_{\varphi, \varphi}(\{x\}) = 1$ for some $\varphi \in \mathcal{D}(A)$, $\|\varphi\| = 1$;
- (iii) $x \in \sigma_p(A)$.

Proof Assume first (i). By Corollary 5.1 (c) $E(\{x\})$ maps $\mathcal{H}$ into $\mathcal{D}(A)$, so by choosing $\varphi \in E(\{x\})(\mathcal{H})$, $\|\varphi\| = 1$, we get a vector satisfying (ii).

Next assume (ii). From Lemma 5.2 it follows that

$$\|A\varphi - x\varphi\|^2 = \int_{\mathbb{R}} |\lambda - x|^2 dE_{\varphi, \varphi}(\lambda) \quad (5.13)$$

for all $\varphi \in \mathcal{D}(A)$. Since $E_{\varphi, \varphi}(\{x\}) = 1 = \|\varphi\|^2 = E_{\varphi, \varphi}(\mathbb{R})$, we have $E_{\varphi, \varphi}(\mathbb{R} \setminus \{x\}) = 0$, so (5.13) implies that $A\varphi = x\varphi$. Hence $x \in \sigma_p(A)$.

Finally assume (iii). Choose $\varphi \in \mathcal{D}(A) \setminus \{0\}$ such that $A\varphi = x\varphi$. By (5.13) $\int_{\mathbb{R}} |\lambda - x|^2 dE_{\varphi, \varphi}(\lambda) = 0$, so the function $\lambda \mapsto \lambda - x$ vanishes $E_{\varphi, \varphi}$ -almost everywhere. Thus $E_{\varphi, \varphi}(\mathbb{R} \setminus \{x\}) = 0$, so that $\|E(\{x\})\varphi\|^2 = E_{\varphi, \varphi}(\{x\}) = E_{\varphi, \varphi}(\mathbb{R}) = \|\varphi\|^2 > 0$. Thus (i) holds. $\square$

Corollary 5.2 *If x is an isolated point of $\sigma(A)$, then x is an eigenvalue of A .*

Proof If x is an isolated point of $\sigma(A)$, there is some $\epsilon > 0$ such that $\sigma(A) \cap (x - \epsilon, x + \epsilon) = \{x\}$. Then $E(\{x\}) \neq 0$ by Theorem 5.3, and the above proposition implies the claim. $\square$

Simple examples show that the converse of the above corollary does not hold.

5.7 Applying a Borel Function to a Selfadjoint Operator

We assume in this section that A is a selfadjoint operator in $\mathcal{H}$ and $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is its spectral measure.

Definition 5.6 If $X \in \mathcal{B}(\mathbb{R})$ is a set such that $E(X) = I$, and $f : X \rightarrow \mathbb{C}$ is a $\mathcal{B}(X)$ -measurable function, we denote $f(A) = \int_X f(\lambda) dE(\lambda)$.

The notation $f(A)$ is clearly unambiguous in the sense that for the choice of another $Y \in \mathcal{B}(\mathbb{R})$ with $\mathbf{E}(Y) = I$ and a $\mathcal{B}(Y)$ -measurable function $g : Y \rightarrow \mathbb{C}$ such that f and g agree on $X \cap Y$, we have $\int_X f(\lambda) d\mathbf{E}(\lambda) = \int_Y g(\lambda) d\mathbf{E}(\lambda)$. We may also denote $f(A) = L(f, \mathbf{E})$ consistently with our earlier usage (where f was defined on all of $\mathbb{R}$).

Example 5.1 (a) If $A \geq 0$ and $f : [0, \infty) \rightarrow \mathbb{R}$ is the function defined by $f(\lambda) = \sqrt{\lambda}$, we denote $f(A) = A^{\frac{1}{2}}$; this is the *square root* of A . In Example 5.2 it will be shown that $A^{\frac{1}{2}}$ is the only positive operator whose square is A . Thus this definition of $A^{\frac{1}{2}}$ is consistent with the notion of the square root of a bounded positive operator considered earlier.

(b) If $f(\lambda) = \lambda^2$, noting that $|\lambda^2|^2 \geq |\lambda|^2$ when $|\lambda| \geq 1$, we see from Proposition 5.8 that $f(A) = A^2$. By induction it is seen that $f(A) = A^n$ if $f(\lambda) = \lambda^n$, $n \in \mathbb{N}$. $\triangleleft$

According to Corollary 5.1 (a) $f(A)$ is always a closed operator, and if f is real valued, then $f(A)$ is selfadjoint (Proposition 5.9). The following result shows the connection between the spectral measures of A and $f(A)$ in case f is real valued.

Theorem 5.4 *Suppose $X \in \mathcal{B}(\mathbb{R})$ and let $f : X \rightarrow \mathbb{R}$ be a $\mathcal{B}(\mathbb{R})$ -measurable function. Assume that $\mathbf{E}(X) = I$. Let $\mathbf{E}^{f(A)} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be the spectral measure of the selfadjoint operator $f(A)$. Then*

- (a) $\mathbf{E}^{f(A)}(Y) = \mathbf{E}(f^{-1}(Y))$ for all $Y \in \mathcal{B}(\mathbb{R})$ and
- (b) if Z is a Borel set containing the set $f(X)$ and $g : Z \rightarrow \mathbb{C}$ is a $\mathcal{B}(Z)$ -measurable function, then $\mathbf{E}^{f(A)}(Z) = I$ and $g(f(A)) = (g \circ f)(A)$.

Proof Let $Z \subset f(X)$ be a Borel set and $g : Z \rightarrow \mathbb{C}$ a $\mathcal{B}(Z)$ -measurable function. Let $\mathbf{E}^f$ denote the spectral measure defined by $\mathbf{E}^f(Y) = \mathbf{E}(f^{-1}(Y))$ for all $Y \in \mathcal{B}(\mathbb{R})$. If $\varphi \in \mathcal{H}$, we have $\mathbf{E}_{\varphi, \varphi}^f(Y) = \mathbf{E}_{\varphi, \varphi}(f^{-1}(Y))$, and so using the well-known theorem on transformations of integrals for positive measures we see that the function $|f|^2$ is $\mathbf{E}_{\varphi, \varphi}^f$ -integrable if and only if the composite function $|g \circ f|^2 = |g|^2 \circ f$ is $\mathbf{E}_{\varphi, \varphi}^f$ -integrable. (As $\mathbf{E}^f(Z) = \mathbf{E}(f^{-1}(Z)) = \mathbf{E}(X) = I$, we may consider $\mathbf{E}^f$ restricted to $\mathcal{B}(Z)$.) It follows that $\mathcal{D}_g^f = \mathcal{D}_{g \circ f}$ where the (obvious) notation $\mathcal{D}_g^f$ refers to the spectral measure $\mathbf{E}^f$. Let now $\varphi \in \mathcal{D}_g^f = \mathcal{D}_{g \circ f}$. Choose a sequence of bounded $\mathcal{B}(Z)$ -measurable functions g_n such that $|g_n| \leq |g|$ and $\lim_{n \rightarrow \infty} g_n = g$ pointwise. For every $\psi \in \mathcal{H}$ we then get

$$\begin{aligned} \langle \psi | (\int_Z g d\mathbf{E}^f) \varphi \rangle &= \int_Z g d\mathbf{E}_{\psi, \varphi}^f = \lim_{n \rightarrow \infty} \int_Z g_n d\mathbf{E}_{\psi, \varphi}^f \\ &= \lim_{n \rightarrow \infty} \int_X (g_n \circ f) d\mathbf{E}_{\psi, \varphi}^f = \langle \psi | (\int_X (g \circ f) d\mathbf{E}) \varphi \rangle \end{aligned}$$

by the dominated convergence theorem and Proposition 4.14, for we have $|g_n \circ f| \leq |g \circ f|$ where the function $|g \circ f|$ is $\mathbf{E}_{\psi, \varphi}$ -integrable (see Proposition 5.5). Thus we have seen that

$$\int_Z g d\mathbf{E}^f = \int_X (g \circ f) d\mathbf{E}. \quad (5.14)$$

With the choice $Z = \mathbb{R}$ and $g(\lambda) = \lambda$ we get $\int_{\mathbb{R}} \lambda d\mathbf{E}^f(\lambda) = f(A)$, and so by Theorem 5.2 $\mathbf{E}^f$ is the spectral measure of $f(A)$. This proves (a), and (b) follows from (5.14), since $g(f(A)) = \int_Z g d\mathbf{E}^{f(A)} = \int_Z g d\mathbf{E}^f$. $\square$

Example 5.2 Suppose that $A \geq 0$. The operator denoted by $A^{\frac{1}{2}}$ in Example 5.1 satisfies $\langle \varphi | A^{\frac{1}{2}} \varphi \rangle = \int_{[0, \infty)} \sqrt{\lambda} d\mathbf{E}_{\varphi, \varphi}(\lambda)$ for all $\varphi \in \mathcal{D}(A^{\frac{1}{2}})$, i.e. $A^{\frac{1}{2}} \geq 0$. Moreover, in view of Theorem 5.4 (b) and Example 5.1 (b), $(A^{\frac{1}{2}})^2 = \int (\sqrt{\lambda})^2 d\mathbf{E}(\lambda) = A$. Let us assume that we have another operator $B \geq 0$ satisfying $B^2 = A$. Let $f : [0, \infty) \rightarrow [0, \infty)$ be the function defined by $f(t) = t^2$ and $g : [0, \infty) \rightarrow [0, \infty)$ the function defined by $g(t) = t^{\frac{1}{2}}$. By Theorem 5.4 (b) we have $B = g(f(B)) = g(A) = A^{\frac{1}{2}}$ with the notation of Example 5.1. Thus A has a uniquely defined positive square root. $\triangleleft$

5.8 One-Parameter Unitary Groups and Stone's Theorem

If $\mu : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{C}$ is a complex measure, its Fourier-Stieltjes transform $\hat{\mu} : \mathbb{R} \rightarrow \mathbb{C}$ is usually defined by the formula

$$\hat{\mu}(x) = \int_{\mathbb{R}} e^{-ixt} d\mu(t).$$

The right-hand side makes sense if μ is replaced by a positive operator measure. In this section we assume that $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure and denote

$$\hat{\mathbf{E}}(x) = \int_{\mathbb{R}} e^{ixt} d\mathbf{E}(t), \quad x \in \mathbb{R}.$$

(Thus we use e^{ixt} instead of e^{-ixt} , this is largely a matter of taste.) Since $|e^{ixt}| = 1$, each $\hat{\mathbf{E}}(x)$ belongs to $\mathcal{L}(\mathcal{H})$. As $\hat{\mathbf{E}}(x)^* = \int_{\mathbb{R}} e^{-ixt} d\mathbf{E}(t)$ and

$$\hat{\mathbf{E}}(x)^* \hat{\mathbf{E}}(x) = \hat{\mathbf{E}}(x) \hat{\mathbf{E}}(x)^* = \int_{\mathbb{R}} e^{ixt} e^{-ixt} d\mathbf{E}(t) = I$$

by Propositions 4.12(b) and 4.13(b), each $\hat{\mathbf{E}}(x)$ is a unitary operator. If $x, y \in \mathbb{R}$, using Proposition 4.13(b) we get $\hat{\mathbf{E}}(x+y) = \int_{\mathbb{R}} e^{ixt} e^{iyt} d\mathbf{E}(t) = \hat{\mathbf{E}}(x) \hat{\mathbf{E}}(y)$, so that $\hat{\mathbf{E}}$ is a group homomorphism. In particular, $\hat{\mathbf{E}}(0) = I$. We now denote $A = \int_{\mathbb{R}} t d\mathbf{E}(t)$, so that A is a selfadjoint operator and $\mathbf{E}$ is its spectral measure. In this kind of situations also in the sequel we use the notation $\hat{\mathbf{E}}(x) = e^{ixA}$ which is consistent with the usage in the previous section when we consider the function $e^{ix(\cdot)}$.

Theorem 5.5 (a) For all $\varphi \in \mathcal{H}$ and $x_0 \in \mathbb{R}$ we have $\lim_{x \rightarrow x_0} \hat{\mathbf{E}}(x)\varphi = \hat{\mathbf{E}}(x_0)\varphi$.
 (b) For all $\varphi \in \mathcal{D}(A)$ we have $\lim_{h \rightarrow 0} \frac{1}{h} [\hat{\mathbf{E}}(h)\varphi - \varphi] = iA\varphi$.

(c) If $\varphi \in \mathcal{H}$ is such that the limit $\lim_{h \rightarrow 0} \frac{1}{h} [\hat{E}(h)\varphi - \varphi]$ exists, then $\varphi \in \mathcal{D}(A)$.

Proof It is enough to show that for every sequence (x_n) in $\mathbb{R}$ satisfying $\lim_{n \rightarrow \infty} x_n = x_0$ we have $\lim_{n \rightarrow \infty} \hat{E}(x_n)\varphi = \hat{E}(x_0)\varphi$. But this follows from the dominated convergence theorem, for by Lemma 5.2

$$\left\| \hat{E}(x_n)\varphi - \hat{E}(x_0)\varphi \right\|^2 = \int_{\mathbb{R}} |e^{ix_nt} - e^{ix_0t}|^2 d\mathbf{E}_{\varphi, \varphi}(t) \rightarrow 0,$$

since $e^{ix_nt} \rightarrow e^{ix_0t}$ and $|e^{ix_nt} - e^{ix_0t}|^2 \leq 4$.

(b) For any $\varphi \in \mathcal{D}$ the function $t \mapsto t^2$ is $\mathbf{E}_{\varphi, \varphi}$ -integrable. Using Lemma 5.2 and the dominated convergence theorem we see that

$$\left\| \frac{1}{h_n} [\hat{E}(h_n)\varphi - \varphi] - iA\varphi \right\|^2 = \int_{\mathbb{R}} \left| \frac{1}{h_n} (e^{ih_nt} - 1) - it \right|^2 d\mathbf{E}_{\varphi, \varphi}(t)$$

converges to zero whenever $h_n \rightarrow 0$. Indeed, $\lim_{n \rightarrow \infty} \frac{1}{h_n} (e^{ih_nt} - 1) = it$ and a direct calculation shows that

$$\left| \frac{1}{h_n} (e^{ih_nt} - 1) - it \right|^2 = t^2 \left[\frac{1}{h_nt} (\cos h_nt - 1) \right]^2 + t^2 \left[\frac{\sin h_nt}{h_nt} - 1 \right]^2,$$

where $\left| \frac{1}{h_nt} (\cos h_nt - 1) \right| \leq 1$ by the mean value theorem and $-1 \leq \frac{\sin h_nt}{h_nt} - 1 \leq 0$, since $|\sin h_nt| \leq |h_nt|$ and $\sin h_nt$ and h_nt have the same sign.

(c) Consider the vector subspace

$$\mathcal{D} = \left\{ \varphi \in \mathcal{H} \mid \lim_{h \rightarrow 0} \frac{1}{h} [\hat{E}(h)\varphi - \varphi] \text{ exists} \right\}$$

and define the linear map $B : \mathcal{D} \rightarrow \mathcal{H}$ by the formula $B\psi = \frac{1}{i} \lim_{h \rightarrow 0} \frac{1}{h} [\hat{E}(h)\psi - \psi]$. Then $B \supset A$ by (b). Since A is selfadjoint, to see that $A = B$, it is in view of Proposition 5.3 enough to show that B is symmetric. But for any $\varphi, \psi \in \mathcal{D}$ we get

$$\begin{aligned} \langle \varphi \mid B\psi \rangle &= \frac{1}{i} \lim_{h \rightarrow 0} \frac{1}{h} \langle \varphi \mid (\hat{E}(h) - I)\psi \rangle = \left\langle i \lim_{h \rightarrow 0} \frac{1}{h} (\hat{E}^*(h) - I)\varphi \mid \psi \right\rangle \\ &= \left\langle -i \lim_{h \rightarrow 0} \frac{1}{-h} (\hat{E}(-h) - I)\varphi \mid \psi \right\rangle = \langle B\varphi \mid \psi \rangle. \end{aligned} \quad \square$$

Definition 5.7 The function $x \mapsto U(x) \in \mathcal{L}(\mathcal{H})$ on $\mathbb{R}$ is a *strongly continuous one-parameter unitary group* (in $\mathcal{L}(\mathcal{H})$) if each $U(x)$ is a unitary operator, $U(x+y) = U(x)U(y)$ for all $x, y \in \mathbb{R}$ and $\lim_{x \rightarrow x_0} U(x)\varphi = U(x_0)\varphi$ for all $x_0 \in \mathbb{R}$ and $\varphi \in \mathcal{H}$.

The function $x \mapsto \hat{E}(x) = e^{ixA}$ was above been seen to be a strongly continuous one-parameter unitary group. We now prove the converse result due to M.H. Stone.

Theorem 5.6 *If the map $x \rightarrow U(x)$ is a strongly continuous one-parameter unitary group in $\mathcal{L}(\mathcal{H})$, then there is a unique selfadjoint operator A in $\mathcal{H}$ such that $U(x) = e^{ixA}$ for all $x \in \mathbb{R}$.*

Proof The uniqueness of A follows from Theorem 5.5 which also gives a hint for the construction of A : for suitable vectors $\varphi \in \mathcal{H}$, differentiate the function $x \rightarrow U(x)\varphi$ at the origin. We let $C_c^\infty(\mathbb{R})$ denote the vector space of all infinitely differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with compact support. If $f \in C_c^\infty(\mathbb{R})$ and $\varphi \in \mathcal{H}$, we denote

$$\varphi_f = \int_{\mathbb{R}} f(x)U(x)\varphi dx.$$

(This Bochner integral with respect to the Lebesgue measure exists since the integrand, denote it by F , is continuous and vanishes outside some compact set K , and so it can be approximated uniformly by integrable simple functions: for each $n \in \mathbb{N}$ cover the compact set $F(K) \subset \mathcal{H}$ by a finite number of balls with diameter $< \frac{1}{n}$ and construct a suitable simple function by using their inverse images. The integral can also be interpreted in the sense of Riemann.) Let $\mathcal{D}$ be the vector space spanned by all the vectors φ_f where $\varphi \in \mathcal{H}$ and $f \in C_c^\infty(\mathbb{R})$. For each $n \in \mathbb{N}$ it is possible (see e.g. [1]) to choose a function $j_n \in C_c^\infty(\mathbb{R})$ such that $j_n \geq 0$, j_n vanishes outside the interval $[-\frac{1}{n}, \frac{1}{n}]$ and $\int_{\mathbb{R}} j_n(x)dx = 1$. For each $\varphi \in \mathcal{H}$

$$\begin{aligned} \|\varphi_{j_n} - \varphi\| &= \left\| \int_{\mathbb{R}} j_n(x)(U(x)\varphi - \varphi)dx \right\| \\ &\leq \sup_{-\frac{1}{n} \leq x \leq \frac{1}{n}} \|U(x)\varphi - \varphi\| \int_{\mathbb{R}} j_n(t)dt \rightarrow 0, \end{aligned}$$

when $n \rightarrow \infty$, because $\lim_{x \rightarrow 0} U(x)\varphi = \varphi$. It follows that $\mathcal{D}$ is dense in $\mathcal{H}$. If $\varphi \in \mathcal{H}$ and $f \in C_c^\infty(\mathbb{R})$, we get

$$\begin{aligned} \frac{1}{h}[U(h) - I]\varphi_f &= \int_{\mathbb{R}} f(x) \frac{1}{h}[U(h)U(x) - U(x)]\varphi dx \\ &= \int_{\mathbb{R}} \frac{1}{h}f(x)U(h+x)\varphi dx - \int_{\mathbb{R}} \frac{1}{h}f(x)U(x)\varphi dx \\ &= \int_{\mathbb{R}} \frac{1}{h}f(x-h)U(x)\varphi dx - \int_{\mathbb{R}} \frac{1}{h}f(x)U(x)\varphi dx \\ &= \int_{\mathbb{R}} \frac{1}{h}[f(x-h) - f(x)]U(x)\varphi dx, \end{aligned}$$

for integration commutes with the continuous linear operator $U(h) - I$ and is translation invariant (which can be seen by approximating with simple functions). By the mean value theorem

$$|h^{-1}[f(x-h) - f(x)] - (-f'(x))| = |-f'(\xi) + f'(x)|$$

for some ξ with $|\xi - x| < |h|$, and since f' is uniformly continuous, we have $\lim_{h \rightarrow 0} \frac{1}{h} [f(x-h) - f(x)] = -f'(x)$ uniformly in x . In view of the above calculation we thus get

$$\lim_{h \rightarrow 0} \frac{1}{h} [U(h) - I]\varphi_f = \int_{\mathbb{R}} -f'(x)U(x)\varphi dx = \varphi_{-f'}. \quad (5.15)$$

Thus for any $\psi \in \mathcal{D}$ we may define

$$B\psi = \frac{1}{i} \lim_{h \rightarrow 0} [U(h) - I]\psi. \quad (5.16)$$

If $\varphi, \psi \in \mathcal{D}$, an argument used in the proof of Theorem 5.5 (c) shows that $\langle \varphi | B\psi \rangle = \langle B\varphi | \psi \rangle$, i.e. B is symmetric: $B \subset B^*$. Since B^* is a closed operator by Proposition 5.2 (b), the closure in $\mathcal{H} \times \mathcal{H}$ of the graph of B is contained in the graph of B^* and is thus the graph of an operator $\overline{B}$. Denote $A = \overline{B}$. It is easily seen that the domain $\mathcal{D}(A)$ of A is a vector subspace of $\mathcal{H}$ and A is a linear map. From Proposition 5.2 (a) it follows that

$$\mathcal{G}(A^*) = \{(-A\xi, \xi) \mid \xi \in \mathcal{D}(A)\}^\perp = \{(-B\xi, \xi) \mid \xi \in \mathcal{D}(B)\}^\perp = \mathcal{G}(B^*),$$

because the map $(\varphi, \psi) \mapsto (-\psi, \varphi)$ is a homeomorphism and maps therefore $\mathcal{G}(A)$ onto the closure of the image of $\mathcal{G}(B)$.

We next show, using Proposition 5.11, that A is selfadjoint. By its definition A is closed. Let $\xi \in \mathcal{D}(A^*) = \mathcal{D}(B^*)$ be such that $A^*\xi = i\xi$ (so that $B^*\xi = i\xi$). For any $\varphi \in \mathcal{D}(B) = \mathcal{D}$ we have

$$\begin{aligned} \frac{d}{dx} \langle U(x)\varphi | \xi \rangle &= \left\langle \lim_{h \rightarrow 0} h^{-1} [U(x+h) - U(x)]\varphi \middle| \xi \right\rangle \\ &= \langle iBU(x)\varphi | \xi \rangle = -i \langle U(x)\varphi | B^*\xi \rangle \\ &= -i \langle U(x)\varphi | i\xi \rangle = \langle U(x)\varphi | \xi \rangle. \end{aligned}$$

(Here we have $U(x)\varphi \in \mathcal{D}$, for

$$U(x)\varphi_f = (U(x)\varphi)_f, \quad (5.17)$$

since integration commutes with the continuous linear map $U(x)$.) It follows that the complex valued function f , $f(x) = \langle U(x)\varphi | \xi \rangle$, satisfies the differential equation $f' = f$, and so $f(x) = f(0)e^x$ for all $x \in \mathbb{R}$. As $\|U(x)\| = 1$, f is bounded and we must have $\langle \varphi | \xi \rangle = f(0) = 0$. Since $\mathcal{D}$ is dense in $\mathcal{H}$, this implies that $\xi = 0$. In a similar way it is shown that zero is the only vector $\xi \in \mathcal{D}(A^*)$ satisfying $A^*\xi = -i\xi$. Thus Proposition 5.11 shows that A is selfadjoint.

Let now $V(x) = e^{ixA}$. We show that $U(x) = V(x)$. Let $\varphi \in \mathcal{D}$. As $\varphi \in \mathcal{D}(A)$, we have $V(x)\varphi \in \mathcal{D}(A)$ (see Corollary 5.1 (a)). By Theorem 5.5 (b)

$$\begin{aligned} \frac{d}{dx} V(x)\varphi &= \lim_{h \rightarrow 0} \frac{1}{h} [V(x+h) - V(x)]\varphi \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [V(h) - I]V(x)\varphi = iAV(x)\varphi. \end{aligned}$$

For all $x \in \mathbb{R}$ we have $U(x)\varphi \in \mathcal{D} \subset \mathcal{D}(A)$ by (5.17). Let $u(x) = U(x)\varphi - V(x)\varphi$. Then $\frac{d}{dx}u(x) = iBU(x)\varphi - iAV(x)\varphi = iAu(x)$. It follows that

$$\begin{aligned} \frac{d}{dx} \|u(x)\|^2 &= \frac{d}{dx} \langle u(x) | u(x) \rangle \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [\langle u(x+h) | u(x+h) \rangle - \langle u(x+h) | u(x) \rangle \\ &\quad + \langle u(x+h) | u(x) \rangle - \langle u(x) | u(x) \rangle] \\ &= \left\langle u(x) \left| \frac{d}{dx} u(x) \right. \right\rangle + \left\langle \frac{d}{dx} u(x) \left| u(x) \right. \right\rangle \\ &= \langle u(x) | iAu(x) \rangle + \langle iAu(x) | u(x) \rangle = 0 \end{aligned}$$

for all $x \in \mathbb{R}$, and since $u(0) = 0$, we have $u(x) = 0$ for all $x \in \mathbb{R}$. Thus $U(x)\varphi - V(x)\varphi = 0$ for all $x \in \mathbb{R}$, $\varphi \in \mathcal{D}$. Since $\mathcal{D}$ is dense in $\mathcal{H}$, we have $U(x) = V(x) = e^{ixA}$ for all $x \in \mathbb{R}$. $\square$

5.9 Taking Stock: Hilbert Space Theory and Its Use in Quantum Mechanics

To summarise (at the risk of slight repetition) the discussion so far, Chaps. 2–4 and this chapter develop some central Hilbert space theory, with an emphasis on the material that is needed in the formulation of quantum mechanics. The presentation is self-contained and does not require any particular preparation from the reader apart from a few basic results of functional analysis and measure theory. Chapters 2 and 3 are free from measure theory. They culminate in the theory of certain Hilbert space operator ideals, especially the trace class, used to describe the states in quantum mechanics. Here a crucial tool is the spectral representation of a compact selfadjoint operator. Chapter 4 and this chapter, on the other hand, depend heavily on measure theory, and one aim there is to develop the language for the representation of observables in quantum mechanics.

In the most traditional formulation, observables are described as (generally unbounded) selfadjoint operators in a Hilbert space $\mathcal{H}$. In this chapter it is shown that these operators correspond bijectively, on the one hand, to Borel spectral measures on the real line, and to strongly continuous one-parameter unitary groups on the

other. The latter correspondence, the theorem of Stone, is essential in the study of the dynamical evolution and the notion of symmetry of a quantum system.

The spectral representation of an unbounded selfadjoint operator is of particular importance in this context. In our approach, this result is reduced via the Cayley transform to the spectral representation of a unitary operator (or more generally, a normal operator), which in turn depends on the spectral representation of a bounded selfadjoint operator. Unlike many other modern treatments, our approach towards obtaining this result avoids the theory of Banach algebras. In a sense, the task often reserved for the Gelfand-Naimark theorem for commutative C^* -algebras is here essentially handled by Lemma 4.8, which [1] attributes to von Neumann's unpublished

lecture notes.

Although the spectral theory of unbounded selfadjoint operators is important, not least for the reasons detailed above, its presence in the remainder of the present text is rather evanescent. In this work, quantum observables are generally described as normalised positive operator measures, which include the spectral measures, whereas selfadjoint operators serve a more ancillary function. This said, it is to be emphasised, however, that in many concrete applications of quantum mechanics in defining, say, the energy operator and thus the energy observable of the physical system under consideration, a fundamental difficulty is to show that a naturally given densely defined symmetric operator H_0 , the sum of kinetic and potential energies, is essentially selfadjoint, having thus a unique selfadjoint extension H and hence also the spectral measure E^H together with the spectral integrals, including the unitary group e^{itH} , $t \in \mathbb{R}$. The theory of extending symmetric operators to selfadjoint operators has its origin in the work of von Neumann [2] and it is extensively developed in classical treatises like Dunford and Schwartz [3], Kato [4], or Reed and Simon [5, 6].

Most of the material in Chaps. 2–4 and this chapter follows a well-established tradition in the presentation of the mathematical foundations of quantum mechanics that is elucidated in many books, and we make no attempt to thoroughly trace its development to the original literature. However, at this point we wish to indicate some sources insofar as they are relevant for the specific perspective pursued in this text.

Our presentation of the Hilbert-Schmidt and trace classes to a certain extent follows that of [7], and the proof of Theorem 3.4 was influenced by [8]. The proof of and some developments leading to the spectral representation theorem of an unbounded selfadjoint operator were originally inspired by the treatment in [9]. It should be pointed out, however, that the integration of (bounded and unbounded) measurable functions is in our presentation mainly treated in the context of general positive operator measures. The specialisation to spectral measures is only made at times to exhibit the distinctive features that then arise (and are needed especially in spectral representation theory). Our key sources for some technical developments in the integration theory with respect to positive operator measures are [10, 11]. The proof of the Stone representation Theorem 5.6 follows the presentation in [5]. Finally,

we note that most of the developments of Chaps. 2–4 and this chapter can be read in an early Finnish version in [12].

5.10 Exercises

1. Let R , S and T be operators in $\mathcal{H}$. Show that $R(ST) = (RS)T$.
2. Let $g : \mathbb{N} \rightarrow \mathbb{C}$ be any function. Define the operator T_g in ℓ^2 as follows: The domain $\mathcal{D}(T_g)$ is the set of those $f \in \ell^2$ for which $gf \in \ell^2$, and $T_g f = gf$ for all $f \in \mathcal{D}(T_g)$.
 - (a) Show that T_g is densely defined.
 - (b) Describe the adjoint of T_g .
 - (c) Is T_g necessarily closed?
3. Let T be an operator in $\mathcal{H}$. Show that T is closable if and only if the following condition holds: $\psi = 0$ whenever (φ_n) is a sequence in $\mathcal{D}(T)$ satisfying $\lim_{n \rightarrow \infty} \varphi_n = 0$ and $\psi = \lim_{n \rightarrow \infty} T\varphi_n$.
4. Give an example of an operator which is not closable.

In Exercises 5–7 and 9 we consider the Hilbert space $L^2 = L^2(\Omega, \mathcal{A}, \mu)$ defined in terms of the measure space $(\Omega, \mathcal{A}, \mu)$, and for each $X \in \mathcal{A}$ we denote $M_X = \{f \in L^2 \mid f(x) = 0 \text{ for } \mu\text{-almost every } x \in \Omega \setminus X\}$.

5. (a) Show that the orthogonal complement of M_X in L^2 is $M_{\Omega \setminus X}$.
 (b) Show that M_X is a closed subspace of L^2 .
6. (a) Show that if P is the orthogonal projection of L^2 onto M_X , then $Pf = \chi_X f$ for every $f \in L^2$.
 (b) Define $\mathbf{E}(X)f = \chi_X f$ for all $X \in \mathcal{A}$ and $f \in L^2$. Show that in this way we get a spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$.
7. Let $g : \Omega \rightarrow \mathbb{C}$ be a measurable function. Define the operator T_g in L^2 as follows: The domain $\mathcal{D}(T_g)$ is the set of those $f \in L^2$ for which $gf \in L^2$, and $T_g f = gf$ for all $f \in \mathcal{D}(T_g)$.
 - (a) Show that T_g is densely defined.
 - (b) Describe the adjoint of T_g .
 - (c) Is T_g necessarily closed?
8. Let $\mathcal{A} \subset 2^\Omega$ be a σ -algebra and $\mu : \mathcal{A} \rightarrow \mathbb{C}$ a complex measure. Let $g : \Omega \rightarrow \mathbb{C}$ be a μ -integrable $\mathcal{A}$ -measurable function (so that by exercise 18 in Chap. 4 we get a complex measure $g \cdot \mu$ whose value for $X \in \mathcal{A}$ is $\int_X g d\mu$). Let $f : \Omega \rightarrow \mathbb{C}$ be an $\mathcal{A}$ -measurable function. Show that f is $g \cdot \mu$ -integrable if and only if fg is μ -integrable. Show that if these equivalent conditions hold, then $\int f d(g \cdot \mu) = \int fg d\mu$. (The corresponding result for positive μ and g may be assumed to be known.)
9. Let $\mathbf{E}$ and T_g be as in Exercises 6 and 7. Show that $T_g = \int_\Omega g d\mathbf{E}$.

10. Let A be a selfadjoint operator in $\mathcal{H}$, $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ its spectral measure, and $U(x) = e^{ixA}$ for all $x \in \mathbb{R}$. Show that for an operator $S \in \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:
- (i) $SE(X) = E(X)S$ for all $X \in \mathcal{B}(\mathbb{R})$;
 - (ii) $SU(x) = U(x)S$ for all $x \in \mathbb{R}$.
11. Let A and U be as in the preceding exercise. Let $\varphi \in \mathcal{H} \setminus \{0\}$ and $\lambda \in \mathbb{R}$. Show that the following conditions are equivalent:
- (i) λ is an eigenvalue of A and φ is an eigenvector belonging to it;
 - (ii) $U(x)\varphi = e^{ix\lambda}\varphi$ for all $x \in \mathbb{R}$.

References

1. Lang, S.: Analysis II. Addison-Wesley, Reading (1969)
2. von Neumann, J.: Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Math. Ann. **102**, 49–131 (1930)
3. Dunford, N., Schwartz, J.T.: Linear Operators. Part II: Spectral Theory. Self Adjoint Operators in Hilbert Space. With the assistance of William G.B., Robert, G.B. (eds.) Interscience Publishers Wiley, New York (1963)
4. Kato, T.: Perturbation Theory for Linear Operators. Classics in Mathematics. Springer, Berlin (1995). Reprint of the 1980 edition
5. Reed, M., Simon, B.: Methods of Modern Mathematical Physics, I. Functional Analysis. Academic Press, New York (1972)
6. Reed, M., Simon, B.: Methods of Modern Mathematical Physics, II. Fourier Analysis, Self-adjointness. Academic Press, New York (1975)
7. S. Sakai. *C*-Algebras and W*-Algebras*. Springer, New York (1971). Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 60
8. Sugiura, M.: Unitary representations and harmonic analysis. North-Holland Mathematical Library, Vol. 44. North-Holland Publishing Co., Amsterdam, Kodansha Ltd., Tokyo (1990)
9. Rudin, W.: Functional Analysis. International Series in Pure and Applied Mathematics, 2nd edn. McGraw-Hill Inc, New York (1991)
10. Lahti, P., Mączyński, M., Ylinen, K.: The moment operators of phase space observables and their number margins. Rep. Math. Phys. **41**(3), 319–331 (1998)
11. Lahti, P., Pellonpää, J.-P., Ylinen, K.: Operator integrals and phase space observables. J. Math. Phys. **40**(4), 2181–2189 (1999)
12. Lahti, P., Ylinen, K.: Johdatus kvanttimekaniikkaan. Fysiikan Kustannus Oy (1989)

Chapter 6

Miscellaneous Algebraic and Functional Analytic Techniques

The level of mathematical sophistication assumed of the reader cannot be perfectly uniform throughout the work, although the present chapter has been designed to alleviate this situation. In the sequel we at times need to refer to outside literature for some rather deep tools. This chapter contains some such results, though many statements are given with full proofs.

Our emphasis is on topological considerations related to $\mathcal{L}(\mathcal{H})$ along with the basic theory of C^* -algebras and their representations. We also recall the notion of the algebraic tensor product of two vector spaces which will make a brief appearance in the next chapter on dilation theory.

6.1 Normal and Positive Linear Maps on $\mathcal{L}(\mathcal{H})$

Up to this point, we have not entered a discussion of a range of natural topologies one may use on the space $\mathcal{L}(\mathcal{H})$, apart from some use of the strong and weak convergence of nets of operators. From now on, we assume some elementary knowledge of the theory of locally convex topological vector spaces, using without explicit reference basic definitions and results that can be found in any introductory text on the subject. We take [1] as our key reference.

In addition to the norm topology, we consider three locally convex topologies on $\mathcal{L}(\mathcal{H})$. The *strong (operator) topology* is generated by the family of seminorms $S \mapsto \|S\varphi\|$ where $\varphi \in \mathcal{H}$. Convergence with respect to this topology is just the strong convergence discussed earlier. A similar statement holds for the *weak operator topology*, which is generated by the family of seminorms $S \mapsto |\langle \psi | S\varphi \rangle|$ where $\varphi, \psi \in \mathcal{H}$.

The σ -weak or *ultraweak topology* of $\mathcal{L}(\mathcal{H})$ is generated by the family of seminorms $S \mapsto |\operatorname{tr}[ST]|$ where $T \in \mathcal{T}(\mathcal{H})$. When $\mathcal{L}(\mathcal{H})$ is regarded as the dual of the Banach space $\mathcal{T}(\mathcal{H})$ as explained in Theorem 3.13, the σ -weak topology is the same as the w^* -topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$, and it is easy to see that the weak operator topology on $\mathcal{L}(\mathcal{H})$ is the same as the coarser topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{F}(\mathcal{H}))$. We sometimes

use the notation $\mathcal{L}(\mathcal{H})_*$ for the space of those linear functionals on $\mathcal{L}(\mathcal{H})$ which are σ -weakly continuous. Then $\mathcal{L}(\mathcal{H})_*$ may be (isometrically) identified with $\mathcal{T}(\mathcal{H})$ and we write accordingly $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H})_*) = \sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$. The space $\mathcal{L}(\mathcal{H})_*$ is called the *predual* of $\mathcal{L}(\mathcal{H})$. (It is known that $\mathcal{L}(\mathcal{H})_*$ is up to isometric isomorphism the only Banach space having $\mathcal{L}(\mathcal{H})$ with the operator norm as its dual.)

It is a standard fact that a net $(A_i)_{i \in \mathcal{I}}$ in $\mathcal{L}(\mathcal{H})$ converges to an operator $A \in \mathcal{L}(\mathcal{H})$ ultraweakly, i.e. with respect to the topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$, if and only if $\lim_i \operatorname{tr}[TA_i] = \operatorname{tr}[TA]$ for each $T \in \mathcal{T}(\mathcal{H})$. In this case we write $A = \operatorname{uw}\text{-}\lim_i A_i$, or $A_i \rightarrow^{\operatorname{uw}} A$. Ultraweak convergence clearly implies weak convergence. On a norm bounded set of operators the converse is also true.

Lemma 6.1 *Let $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$ be a norm bounded set. A net $(A_i)_{i \in \mathcal{I}}$ in $\mathcal{A}$ converges ultraweakly to an operator $A \in \mathcal{A}$ if and only if it converges to A weakly.*

Proof Let $\operatorname{w}\text{-}\lim A_i = A$, and assume that $\mathcal{A}$ is norm bounded by $M > 0$, that is, $\|B\| \leq M$ for any $B \in \mathcal{A}$. From $\operatorname{w}\text{-}\lim A_i = A$ it follows that $\lim \operatorname{tr}[FA_i] = \operatorname{tr}[FA]$ for any $F \in \mathcal{F}(\mathcal{H})$. Let $T \in \mathcal{T}(\mathcal{H})$ and $\varepsilon > 0$. There is an $F \in \mathcal{F}(\mathcal{H})$ such that $\|T - F\|_1 < \varepsilon/2M$ (see Theorem 3.11 (e)). By assumption there is an $i_0 \in \mathcal{I}$ such that $|\operatorname{tr}[FA_i] - \operatorname{tr}[FA]| \leq \frac{\varepsilon}{2}$ whenever $i \geq i_0$. Thus

$$\begin{aligned} |\operatorname{tr}[TA_i] - \operatorname{tr}[TA]| &= |\operatorname{tr}[T(A_i - A)]| \\ &\leq |\operatorname{tr}[(T - F)(A_i - A)]| + |\operatorname{tr}[F(A_i - A)]| \\ &\leq \|T - F\|_1 \|A_i - A\| + |\operatorname{tr}[F(A_i - A)]| \\ &\leq \frac{\varepsilon}{2M} 2M + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

for any $i \geq i_0$. □

In the rest of this section, we let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces.

Proposition 6.1 *For a linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ the following conditions are equivalent:*

- (i) Φ is continuous with respect to the σ -weak topologies of $\mathcal{L}(\mathcal{H})$ and $\mathcal{L}(\mathcal{K})$;
- (ii) there is a bounded linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ such that $\operatorname{tr}[S\Psi(T)] = \operatorname{tr}[T\Phi(S)]$ for all $S \in \mathcal{L}(\mathcal{H})$, $T \in \mathcal{T}(\mathcal{K})$.

If these conditions hold, then Φ is norm continuous.

Proof Since we may identify e.g. $\mathcal{L}(\mathcal{H})$ with the dual of the Banach space $\mathcal{T}(\mathcal{H})$, the claim follows from the basic duality theory of Banach spaces. □

In the above situation the maps Φ and Ψ are the *transposes* (or *dual maps*) of each other (with respect to the dualities $\langle \mathcal{T}(\mathcal{H}), \mathcal{L}(\mathcal{H}) \rangle$ and $\langle \mathcal{T}(\mathcal{K}), \mathcal{L}(\mathcal{K}) \rangle$). A linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ satisfying the equivalent conditions in this proposition is said to be *normal*. As a special case, the σ -weakly continuous linear functionals on $\mathcal{L}(\mathcal{H})$

are said to be normal. These form the predual of $\mathcal{L}(\mathcal{H})$ in the terminology introduced above, and they are the same as the functionals $S \mapsto \text{tr}[ST]$ where $T \in \mathcal{T}(\mathcal{H})$.

A linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is said to be *positive* if $\Phi(S) \geq 0$ whenever $S \geq 0$. The positivity of a linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ is defined similarly. We use the same terminology for real linear maps in the situation of the following lemma. We use the notation $\mathcal{T}(\mathcal{K})_1^+$ for the set of the positive trace class operators on $\mathcal{K}$ with trace one.

Lemma 6.2 *A positive real linear map $\Psi : \mathcal{T}_s(\mathcal{K}) \rightarrow \mathcal{T}_s(\mathcal{H})$ is norm bounded and*

$$\|\Psi\| = \sup \{ \text{tr}[\Psi(T)] \mid T \in \mathcal{T}(\mathcal{K})_1^+ \}.$$

In particular, any positive linear functional $\mathcal{T}_s(\mathcal{K}) \rightarrow \mathbb{R}$ is bounded.

Proof We may represent any selfadjoint trace class operator $T \in \mathcal{T}_s(\mathcal{K})$ in the form $T = T^+ - T^-$ where $T^+, T^- \in \mathcal{T}(\mathcal{K})^+$. (Take, for instance, $T^\pm = |T| \pm T$.) We have now by the linearity and positivity of Ψ

$$\begin{aligned} \|\Psi\| &= \sup \{ \|\Psi(T)\|_1 \mid T \in \mathcal{T}_s(\mathcal{K}), \|T\|_1 = 1 \} \\ &= \sup \{ \text{tr}[\Psi(T^+ + T^-)] \mid T^\pm \in \mathcal{T}(\mathcal{K})^+, \text{tr}[T^+ + T^-] = 1 \} \\ &= \sup \{ \text{tr}[\Psi(T)] \mid T \in \mathcal{T}(\mathcal{K})_1^+ \}. \end{aligned}$$

If the set $\{ \text{tr}[\Psi(T)] \mid T \in \mathcal{T}(\mathcal{K})_1^+ \}$ is not bounded (from above), one finds for each $n \in \mathbb{N}$ an operator $T_n \in \mathcal{T}(\mathcal{K})_1^+$ such that $\text{tr}[\Psi(T_n)] \geq 4^n$. Since $\mathcal{T}(\mathcal{K})_1^+$ is complete, we may define $T = \sum_{n=1}^{\infty} 2^{-n} T_n \in \mathcal{T}(\mathcal{K})_1^+$, and since $0 \leq 2^{-n} T_n \leq T$ and Ψ is positive, we have $0 \leq 2^{-n} \Psi(T_n) \leq \Psi(T)$. Thus

$$0 \leq 2^n \quad (= 4^n 2^{-n}) \leq 2^{-n} \text{tr}[\Psi(T_n)] \leq \text{tr}[\Psi(T)]$$

for every $n \in \mathbb{N}$. As $\Psi(T) \in \mathcal{T}_s(\mathcal{H})$, i.e. $\text{tr}[\Psi(T)] < \infty$, we arrive at a contradiction. Thus Ψ has to be norm bounded. The last claim follows from the case $\mathcal{H} = \mathbb{C}$. $\square$

Proposition 6.2 (a) *Any positive linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is bounded.*

(b) *Any positive linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ is bounded.*

Proof (a) Any operator $A \in \mathcal{L}(\mathcal{H})$ may be written as $A_1 + iA_2$ with two selfadjoint operators A_1 and A_2 such that $\|A_j\| \leq \|A\|$. Using the spectral theorem we may furthermore write $A_j = A_j^+ - A_j^-$ where A_j^+ and A_j^- are positive operators with norms at most $\|A\|$. It is therefore sufficient to show that Φ is bounded on the positive part of the unit ball of $\mathcal{L}(\mathcal{H})$. Thus assume that $A \geq 0$ with norm at most one. Since $-I \leq A \leq I$, we have $-\Phi(I) \leq \Phi(A) \leq \Phi(I)$, and so $\|A\| \leq \|\Phi(I)\|$ by Proposition 2.2.

(b) Again the claim may easily be reduced to the selfadjoint part of the trace class, which was handled in Lemma 6.2. $\square$

We prepare the proof of Theorem 6.1 with a lemma.

Lemma 6.3 (a) For any $A \geq 0$ in $\mathcal{L}(\mathcal{H})$ there is an increasing net of finite rank operators $A_i \geq 0$ such that $A = \text{w-lim } A_i = \sup_i A_i$.

(b) If $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is a normal linear map taking the value zero for each finite rank operator in $\mathcal{L}(\mathcal{H})$, then $\Phi = 0$.

Proof (a) Let K be an orthonormal basis of $\mathcal{H}$ and denote by $\mathcal{I}$ the directed set of the finite subsets of K (ordered by inclusion). For any $i \in \mathcal{I}$, we let P_i denote the projection onto the (finite-dimensional, hence closed) subspace spanned by i . From Theorems 3.3 (a) and 2.20 it follows that the net $(P_i)_{i \in \mathcal{I}}$ converges strongly to the identity operator I . Thus the increasing net (A_i) with $A_i = A^{1/2} P_i A^{1/2}$ satisfies the requirements in (a).

(b) Let (P_i) be as in the proof of (a). For any $S \in \mathcal{L}(\mathcal{H})$, the net $(P_i S)_{i \in \mathcal{I}}$ converges weakly to S and hence also ultraweakly by Lemma 6.1. Since $\Phi(P_i S) = 0$ for all $i \in \mathcal{I}$, we thus have $\Phi = 0$. $\square$

Theorem 6.1 For a positive linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ the following conditions are equivalent:

- (i) Φ is normal;
- (ii) for any increasing net (A_i) in $\mathcal{L}_s(\mathcal{H})$, bounded from above with $\sup(A_i) = A$, $\text{w-lim } \Phi(A_i) = \Phi(A)$.

Proof The implication (i) $\implies$ (ii) is a consequence of the fact that A is the weak limit and hence by Lemma 6.1 also the ultraweak limit of the net (A_i) .

Assume now (ii). To prove (i) it is enough to show that for each $T \in \mathcal{T}(\mathcal{K})$ the bounded linear functional f_T on $\mathcal{L}(\mathcal{H})$ defined by $f_T(S) = \text{tr}[T\Phi(S)]$ is normal. Denote by g the restriction of f_T to $\mathcal{C}(\mathcal{H})$, the space of the compact operators on $\mathcal{H}$. By Theorem 3.12 there is some $R \in \mathcal{T}(\mathcal{H})$ such that $g(S) = \text{tr}[RS]$ for all $S \in \mathcal{C}(\mathcal{H})$. We denote by $\tilde{g}$ the normal linear functional on $\mathcal{L}(\mathcal{H})$ extending g , i.e. defined by $\tilde{g}(S) = \text{tr}[RS]$ for all $S \in \mathcal{C}(\mathcal{H})$. We claim that $f_T = \tilde{g}$. It is enough to show that $f_T(A) = \tilde{g}(A)$ for any $A \geq 0$ in $\mathcal{L}(\mathcal{H})$. Using Lemma 6.3 we find an increasing net of finite rank operators $A_i \geq 0$ such that $A = \text{w-lim } A_i = \sup_i A_i$. Hence $\tilde{g}(A) = \lim \tilde{g}(A_i) = \lim f_T(A_i) = f_T(A)$ where the last equality follows from Lemma 6.1. $\square$

6.2 Basic Notions of the Theory of C^* -algebras and Their Representations

In this section we present basic definitions and results from the theory of C^* -algebras. Relevant proofs can be found in [2].

Definition 6.1 An algebra A over the field $\mathbb{C}$ is a vector space A equipped with such a mapping $(x, y) \mapsto xy$ (called *multiplication*) from the Cartesian product $A \times A$ to A that for all $x, y, z \in A$ and $\lambda \in \mathbb{C}$ the following hold:

- (i) $x(yz) = (xy)z$,
- (ii) $x(y + z) = xy + xz$, $(x + y)z = xz + yz$,
- (iii) $(\lambda x)y = \lambda(xy) = x(\lambda y)$.

An algebra A is *commutative* if $xy = yx$ whenever $x, y \in A$. If there is an element $e \in A \setminus \{0\}$ such that $ex = xe = x$ for all $x \in A$, e is called the *unit (element)* or the *identity (element)* of A , and A is said to be *unital*. A set $B \subset A$, $B \neq \emptyset$, is a *subalgebra* of A , if $x + y \in B$, $xy \in B$ and $\lambda x \in B$ whenever $x, y \in B$ and $\lambda \in \mathbb{C}$. A set $B \subset A$ is called a *left ideal*, if it is a vector subspace of A and $xy \in B$ whenever $x \in A$, $y \in B$. A *right ideal* is defined in an analogous way. If $B \subset A$ is both a left and right ideal of A , B is a *two-sided ideal* of A . An (automatically) two-sided ideal of a commutative algebra is just called an *ideal*.

According to the above definition, only a non-zero element of an algebra can be an identity element. An algebra can contain at most one unit element, for if e and e' are unit elements, we have $e' = e'e = e$. In the sequel we usually denote by 1 the identity element of any algebra.

Definition 6.2 Let A and B be algebras. A map $\phi : A \rightarrow B$ is an (*algebra*) *homomorphism* if ϕ is linear and $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in A$. If, in addition, ϕ is a bijection, ϕ is called an (*algebra*) *isomorphism*. If both A and B have identity elements (denoted by 1) and $\phi : A \rightarrow B$ is a homomorphism such that $\phi 1 = 1$ then ϕ is called *unital*. The set $\phi^{-1}(0) = \{x \in A \mid \phi(x) = 0\}$ is the *kernel* of the homomorphism ϕ ; we denote $\phi^{-1}(0) = \ker(\phi)$.

Definition 6.3 Let $(A, \|\cdot\|)$ be a normed vector space over $\mathbb{C}$. If A is endowed with a multiplication $(x, y) \mapsto xy$ such that A is an algebra and $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in A$, we say that A is a *normed algebra*. If, in addition, A is complete, i.e. a Banach space, A is called a *Banach algebra*.

Definition 6.4 Let A be an algebra (with complex coefficients). A map $x \mapsto x^*$ from A to A is called an *involution*, if

- (i) $x^{**} = x$ (notation: $(x^*)^* = x^{**}$),
- (ii) $(x + y)^* = x^* + y^*$,
- (iii) $(\lambda x)^* = \overline{\lambda}x^*$,
- (iv) $(xy)^* = y^*x^*$

for all $x, y \in A$ and $\lambda \in \mathbb{C}$. An algebra equipped with an involution is called a **-algebra*.

Definition 6.5 Let A be a normed algebra with an involution $x \mapsto x^*$. If $\|x\| = \|x^*\|$ for all $x \in A$, A is a *normed *-algebra*. If, in addition, A is complete, A is called a *Banach *-algebra*.

Definition 6.6 (a) If B is a subalgebra of a *-algebra A , such that B is closed under the involution $x \mapsto x^*$ (i.e. $x^* \in B$ for all $x \in B$), B is a **-subalgebra* of A .

- (b) Let A and B be $*$ -algebras. If $\Phi : A \rightarrow B$ is an algebra homomorphism (respectively algebra isomorphism) such that $\Phi(x^*) = (\Phi x)^*$ for all $x \in A$, the homomorphism (isomorphism) Φ is called a $*$ -homomorphism (respectively $*$ -isomorphism). If there is a $*$ -isomorphism $\Phi : A \rightarrow B$, A and B are mutually $*$ -isomorphic.

Definition 6.7 A Banach algebra A equipped with an involution $x \mapsto x^*$ is a C^* -algebra, if

$$\|x\|^2 = \|xx^*\| \quad (6.1)$$

for all $x \in A$.

Remark 6.1 Equation (6.1) implies $\|x\|^2 \leq \|x^*\| \|x\|$ so that if $x \neq 0$, one has $\|x\| \leq \|x^*\|$ (the same holds even if $x = 0$). Similarly, $\|x^*\| \leq \|x^{**}\| = \|x\|$. Thus $\|x\| = \|x^*\|$ so that any C^* -algebra is a Banach $*$ -algebra. Using this observation one sees that the equation $\|x\|^2 = \|xx^*\|$ of Definition 6.7 could be replaced with $\|x\|^2 = \|x^*x\|$. $\triangleleft$

Let $\mathcal{H}$ be a Hilbert space with the inner product $\langle \cdot | \cdot \rangle$. Clearly $\mathcal{L}(\mathcal{H})$ is a Banach algebra. Equipped with the involution $T \mapsto T^*$, $\mathcal{L}(\mathcal{H})$ has the structure of a C^* -algebra. Clearly every closed $*$ -algebra of $\mathcal{L}(\mathcal{H})$ is a C^* -algebra, too. A theorem due to Gelfand and Naimark states that there are essentially no other C^* -algebras, i.e. every C^* -algebra is isometrically $*$ -isomorphic to a closed $*$ -subalgebra of some $\mathcal{L}(\mathcal{H})$.

Let S be a locally compact Hausdorff space. The space $C_0(S)$ of continuous functions vanishing at infinity (with pointwise operations) is a commutative Banach algebra. We immediately see that the complex conjugation $f \mapsto \bar{f}$ is an involution on $C_0(S)$. Since

$$\|f\bar{f}\| = \sup_{x \in S} |f(x)|^2 = \left(\sup_{x \in S} |f(x)| \right)^2 = \|f\|^2,$$

$C_0(S)$ is a C^* -algebra. In fact any commutative C^* -algebra is isometrically $*$ -isomorphic to some $C_0(S)$. This is another important result due to Gelfand and Naimark.

Remark 6.2 A $*$ -subalgebra $\mathcal{M}$ of $\mathcal{L}(\mathcal{H})$ is called a *von Neumann algebra* if $\mathcal{M}$ contains the identity and is closed in the $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$ topology. (Here $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$ could be replaced e.g. by the weak or the strong operator topology.) The space of the $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$ -continuous linear functionals on $\mathcal{M}$ is (as in the case of $\mathcal{L}(\mathcal{H})$) called the *predual* of $\mathcal{M}$ and denoted by $\mathcal{M}_*$. It is a closed subspace of $\mathcal{M}^*$, and again it is up to isometric isomorphism the only Banach space having $\mathcal{M}$ with the operator norm as its dual. These basic results may be found, e.g., in [2]. In the context of general von Neumann algebras $\mathcal{M}$ in place of $\mathcal{L}(\mathcal{H})$ we also talk about normal linear maps when the topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{H})_*)$ is replaced by $\sigma(\mathcal{M}, \mathcal{M}_*)$. $\triangleleft$

The following definition generalises the terminology used in connection with $\mathcal{L}(\mathcal{H})$.

Definition 6.8 Let A be a $*$ -algebra and $x \in A$. If $x^* = x$, x is *selfadjoint*. If $x^*x = xx^*$, x is *normal*. If A has a unit element 1 and $x^*x = xx^* = 1$, then x is *unitary*.

Definition 6.9 Let A be a $*$ -algebra and $\mathcal{H}$ a Hilbert space. A $*$ -homomorphism $\pi : A \rightarrow \mathcal{L}(\mathcal{H})$ (or the pair $\{\pi, \mathcal{H}\}$) is called a *$*$ -representation* of A (in $\mathcal{H}$). (We often simplify our usage and talk about *representations* instead of *$*$ -representations*.) The closed subspace of $\mathcal{H}$ generated by the set $\{\pi(x)\varphi \mid x \in A, \varphi \in \mathcal{H}\}$ is called the *essential subspace* of π . If this essential subspace is the whole of $\mathcal{H}$ then π is said to be *essential* or *nondegenerate*. If there is some $\xi \in \mathcal{H}$ such that $\mathcal{H}$ is the closed linear span of the set $\{\pi(x)\xi \mid x \in A\}$ then the representation π is said to be *cyclic* and such a ξ is a *cyclic vector*. We may use the notation $\{\pi, \mathcal{H}, \xi\}$ for a cyclic representation with the cyclic vector ξ specified. Two $*$ -representations $\{\pi_1, \mathcal{H}_1\}$ and $\{\pi_2, \mathcal{H}_2\}$ of A are *unitarily equivalent* if there is an isometric isomorphism $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ satisfying $U\pi_1(x) = \pi_2(x)U$ for all $x \in A$.

Proposition 6.3 If $\{\pi_1, \mathcal{H}_1, \xi_1\}$ and $\{\pi_2, \mathcal{H}_2, \xi_2\}$ are two cyclic representations of a $*$ -algebra A such that $\langle \xi_1 \mid \pi_1(x)\xi_1 \rangle = \langle \xi_2 \mid \pi_2(x)\xi_2 \rangle$ for all $x \in A$, then there is an isometric isomorphism $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ such that $U\xi_1 = \xi_2$ and $U\pi_1(x) = \pi_2(x)U$ for all $x \in A$.

Proof See [2].

Let A be a $*$ -algebra, $\mathcal{H}$ a Hilbert space and $\pi : A \rightarrow \mathcal{L}(\mathcal{H})$ a representation. If $M \subset \mathcal{H}$ is a closed subspace such that $\pi(x)M \subset M$ for all $x \in A$ then M is said to be *invariant* (under π). It then easily follows that the orthogonal complement $M^\perp$ is also invariant. Writing e.g. $\pi_M(x) : \pi(x)|_M \rightarrow M$, we can then say that π is in an obvious sense the *direct sum* of the *subrepresentations* π_M and $\pi_{M^\perp}$. In particular, the essential subspace of π is invariant, and we may express π as a direct sum of a nondegenerate representation and the zero representation (in the orthogonal complement of the essential subspace). We may also define the direct sum of an arbitrary number of representations if a boundedness condition is met as follows.

Assume that A is a $*$ -algebra and $\{\pi_i, \mathcal{H}_i\}$ is a $*$ -representation of A for each $i \in \mathcal{I}$, such that $\sup_{i \in \mathcal{I}} \|\pi_i(x)\| < \infty$ for all $x \in A$. (If e.g. A is a C^* -algebra, this condition always holds, since $\|\pi_i(x)\| \leq \|x\|$.) Consider the direct sum $\mathcal{H} = \sum_{i \in \mathcal{I}} \mathcal{H}_i$ and define $\pi((x_i)_{i \in \mathcal{I}})(\varphi_i)_{i \in \mathcal{I}} = (\pi_i(x_i)\varphi_i)_{i \in \mathcal{I}}$ for all $(\varphi_i)_{i \in \mathcal{I}} \in \mathcal{H}$. It is easy to verify that in this way we get a $*$ -representation $\pi : A \rightarrow \mathcal{L}(\mathcal{H})$; it is called the *direct sum* of the (family of) representations π_i .

Proposition 6.4 Any nondegenerate representation of a C^* -algebra is unitarily equivalent to the direct sum of a family of cyclic representations.

Proof See [2].

6.3 Algebraic Tensor Products of Vector Spaces

We have already defined the (complete normed) Hilbert (space) tensor product of two Hilbert spaces, and in the present treatise this is the main use of the term “tensor product”. In the chapter on dilation theory we need, however, an algebraic approach to tensor products, and we now present some relevant preliminaries. In this section we consider general vector spaces without assuming the existence of an inner product. The proofs omitted here may be found, e.g., in [3].

Definition 6.10 Let E, F and G be vector spaces.

(a) A mapping $\phi : E \times F \rightarrow G$ is *bilinear* if for each $x \in E$ the map $y \mapsto \phi(x, y)$ is linear on F and for each $y \in F$ the map $x \mapsto \phi(x, y)$ is linear on E .

(b) We say that the pair (G, ϕ) where $\phi : E \times F \rightarrow G$ is a bilinear map is a *tensor product* of E and F if the following two conditions hold:

(TP1) the linear span of the set $\{\phi(x, y) \mid x \in E, y \in F\}$ is the whole of G ;

(TP2) whenever H is a vector space and $\psi : E \times F \rightarrow H$ is a bilinear map there is a linear map $f : G \rightarrow H$, satisfying $\psi = f \circ \phi$.

Theorem 6.2 *The above conditions (TP1) and (TP2) together are equivalent to the following statement:*

(TP) *whenever H is a vector space and $\psi : E \times F \rightarrow H$ is a bilinear map, there is a unique linear map $f : G \rightarrow H$ such that $\psi = f \circ \phi$.*

Lemma 6.4 *If (G_1, ϕ_1) and (G_2, ϕ_2) are tensor products of the vector spaces E and F , there is a unique isomorphism $f : G_1 \rightarrow G_2$ such that $f(\phi_1(x, y)) = \phi_2(x, y)$ whenever $x \in E, y \in F$.*

In other words, the tensor product is essentially unique if it exists. The existence part is also true:

Theorem 6.3 *The vector spaces E and F have a tensor product.*

If (G, ϕ) is a tensor product of E and F , we usually denote $G = E \otimes F$ and $\phi(x, y) = x \otimes y$. (Note that if E and F are Hilbert spaces, this is in conflict with the usage in Sect. 2.9. When there might be a danger of confusion, we specify which usage is meant.) We still briefly consider the tensor product of linear mappings in our present purely algebraic context. Suppose that E_1, E_2, F_1 and F_2 are vector spaces and $S : E_1 \rightarrow E_2, T : F_1 \rightarrow F_2$ linear maps. The mapping $(x, y) \mapsto Sx \otimes Ty$ from the Cartesian product $E_1 \times F_1$ into the vector space $E_2 \otimes F_2$ is bilinear, and so there is a unique linear map $S \otimes T : E_1 \otimes F_1 \rightarrow E_2 \otimes F_2$ satisfying

$$(S \otimes T)(x \otimes y) = Sx \otimes Ty$$

for all $x \in E_1, y \in F_1$.

Definition 6.11 We say that the linear map $S \otimes T$ considered above is the *tensor product* of the linear maps S and T .

Proposition 6.5 Let $E_i, F_i, i = 1, 2, 3$, be vector spaces and $S_1 : E_1 \rightarrow E_2, S_2 : E_2 \rightarrow E_3, T_1 : F_1 \rightarrow F_2, T_2 : F_2 \rightarrow F_3$ linear maps. Then $(S_2 \otimes T_2) \circ (S_1 \otimes T_1) = S_2 S_1 \otimes T_2 T_1$.

Remark 6.3 One can also form, with essentially the same process as for two vector spaces, the tensor product $E_1 \otimes \cdots \otimes E_n$ of more than two vector spaces $E_1, \dots, E_n$. It is also possible to define $E_1 \otimes \cdots \otimes E_n$ by induction by using the notion of the tensor product of two vector spaces and the easily verified formula $(E_1 \otimes E_2) \otimes E_3 = E_1 \otimes (E_2 \otimes E_3)$. $\triangleleft$

6.4 Completions

We take for granted that for any normed space E there is a Banach space F with an isometric linear map $\theta : E \rightarrow F$ such that the image $\theta(E)$ is dense in F . Such a pair (F, θ) is essentially unique in the sense that if (F', θ') is another pair with these properties, then there is an isometric isomorphism $\eta : F \rightarrow F'$ such that $\theta' = \eta \circ \theta$. The pair (F, θ) (or just F) is called the *completion* of E . There are various ways of constructing the completion. One can start with the completion of a general metric space (constructed, for instance, as the set of equivalence classes of Cauchy sequences) and extend the Banach space operations by uniform continuity or embed E into its bidual (using the Hahn–Banach theorem). In our context the following result is particularly useful.

Proposition 6.6 If E is an inner product space, then its completion can be given a unique inner product extending that of E and yielding the norm of the completion.

We leave the proof as an exercise. We refer to the completion of E equipped with this inner product as the *Hilbert space completion* of E .

6.5 Exercises

1. Prove Proposition 6.5.
2. Complete the proof of Proposition 6.1.
3. Prove Proposition 6.6.
4. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Let $\mathcal{H}_{\otimes}$ denote their Hilbert tensor product (as given in Definition 2.7). Since the map $(\varphi, \psi) \mapsto \varphi \otimes \psi$ (in the notation of Sect. 2.9) on $\mathcal{H} \times \mathcal{K}$ is bilinear, it corresponds to a linear map θ from the algebraic tensor product of $\mathcal{H}$ and $\mathcal{K}$ into $\mathcal{H}_{\otimes}$. Show that θ is injective and its range is dense in $\mathcal{H}_{\otimes}$. (Thus $\mathcal{H}_{\otimes}$ may be regarded as the completion of the algebraic tensor product of $\mathcal{H}$ and $\mathcal{K}$ with respect to the norm inherited from $\mathcal{H}_{\otimes}$.)

References

1. Schaefer, H.H.: Topological Vector Spaces, vol. 3. Springer, New York (1971). (Corrected third printing, Graduate Texts in Mathematics)
2. Takesaki, M.: Theory of Operator Algebras. I, Encyclopaedia of Mathematical Sciences, vol. 124. Springer, Berlin (2002). (Reprint of the first (1979) edition, Operator Algebras and Non-commutative Geometry, 5)
3. Greub, W.H.: Multilinear Algebra. Die Grundlehren der Mathematischen Wissenschaften, Band, vol. 136. Springer New York, Inc., New York (1967)

Chapter 7

Dilation Theory

Naimark's work in the 1940s on the representation of positive operator measures in terms of projection valued measures acting in a larger Hilbert space may be seen as the starting point for the discovery of a vast variety of dilation theorems, an early highlight being Stinespring's dilation theorem from the mid 1950s. This chapter contains a unified approach to the sort of dilation theorems that are particularly useful for quantum mechanics. Besides the Naimark and Stinespring results, these include an analogous two-variable theory and the so-called Kraus representation in a general functional analytic setting. One also encounters the first inklings of these results in action: the sections on operations and instruments and measurement dilations point to the central quantum mechanical applications which are the main theme of this book.

7.1 Completely Positive Linear Maps

In this section, we let A denote an arbitrary $*$ -algebra. For $n \in \mathbb{N}$, let $M_n(A)$ be the linear space of $n \times n$ -matrices with entries in A , equipped with its natural $*$ -algebra structure, i.e., $(a_{ij})(b_{ij}) = (\sum_{k=1}^n a_{ik}b_{kj})$ and $(a_{ij})^* = (a_{ji}^*)$. For a Hilbert space $\mathcal{H}$, $M_n(\mathcal{L}(\mathcal{H}))$ will be identified with $\mathcal{L}(\mathcal{H} \oplus \cdots \oplus \mathcal{H})$, so that positivity in $M_n(\mathcal{L}(\mathcal{H}))$ is well defined.

Definition 7.1 For $n \in \mathbb{N}$, a linear map $\Phi : A \rightarrow \mathcal{L}(\mathcal{H})$ is said to be *n-positive*, if the linear map $\Phi_n : M_n(A) \rightarrow M_n(\mathcal{L}(\mathcal{H}))$ defined by

$$\Phi_n((a_{ij})) = (\Phi(a_{ij}))$$

satisfies $\Phi_n(C^*C) \geq 0$ for all $C \in M_n(A)$. If Φ is n -positive for all $n \in \mathbb{N}$, then Φ is said to be *completely positive*.

Remark 7.1 In case $A = \mathcal{L}(\mathcal{H}_0)$ for some Hilbert space $\mathcal{H}_0$, an operator $S \in A$ is positive if and only if $S = B^*B$ for some $B \in A$, as is seen using the square root of a positive operator for the less trivial direction (see Proposition 4.15). Thus in this case 1-positivity simply means positivity in the sense defined before. $\triangleleft$

Lemma 7.1 Let $\Phi : A \rightarrow \mathcal{L}(\mathcal{H})$ be a linear map. For any $n \in \mathbb{N}$ the following conditions are equivalent:

- (i) Φ is n -positive;
- (ii) $\Phi_n(C) \geq 0$ whenever $C \in M_n(A)$ is of the form $C = (a_i^* a_j)$ for some $a_1, \dots, a_n \in A$;
- (iii) $\sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | \Phi(x_i^* x_j) \xi_j \rangle \geq 0$ for all $x_1, \dots, x_n \in A$, $\xi_1, \dots, \xi_n \in \mathcal{H}$.

Proof Clearly, (ii) and (iii) are equivalent. The equivalence of (i) and (ii) follows from the fact that if $C = (a_i^* a_j)$ then $C = B^*B$ where $B = (b_{ij})$ with $b_{1j} = a_j$, $j = 1, \dots, n$, and $b_{ij} = 0$ elsewhere, and on the other hand, for any $D = (d_{ij}) \in M_n(A)$, $D^*D = \sum_{k=1}^n C_k$, with $C_k = (d_{ki}^* d_{kj})$. $\square$

In the next two propositions, $\mathcal{H}, \mathcal{K}$ are Hilbert spaces.

Proposition 7.1 Let $\Phi : A \rightarrow \mathcal{L}(\mathcal{H})$ and $\Psi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ be n -positive linear maps. Then the composition $\Psi \circ \Phi$ is n -positive.

Proof For any $C \in M_n(A)$, $\Phi_n(C^*C) = D^*D$ for some $D \in M_n(\mathcal{L}(\mathcal{H}))$ (as noted in Remark 7.1), and so $(\Psi \circ \Phi)_n(C^*C) = \Psi_n(\Phi_n(C^*C)) \geq 0$. $\square$

Remark 7.2 If $T : \mathcal{H} \rightarrow \mathcal{K}$ is a bounded linear map, we use the notation $T^* : \mathcal{K} \rightarrow \mathcal{H}$ for the bounded linear map satisfying $\langle T^*\xi | \varphi \rangle = \langle \xi | T\varphi \rangle$ for all $\varphi \in \mathcal{H}$, $\xi \in \mathcal{K}$. The existence of T^* is proved essentially in the same way as in the case where $\mathcal{K} = \mathcal{H}$, and we call T^* the *adjoint* of T also in the general case. We have previously used (and continue to use) the notation T^* also for the transpose (or dual map) for a bounded linear map T in the context of Banach spaces like $\mathcal{T}(\mathcal{H})$ and its dual $\mathcal{L}(\mathcal{H})$. This double use of the same notation should not cause confusion. $\triangleleft$

Proposition 7.2 Let $\Phi : A \rightarrow \mathcal{L}(\mathcal{H})$ be a linear map. Suppose that $V : \mathcal{H} \rightarrow \mathcal{K}$ a bounded linear map and $\pi : A \rightarrow \mathcal{L}(\mathcal{K})$ is a $*$ -representation such that

$$\Phi(x) = V^* \pi(x) V$$

for all $x \in A$. Then Φ is completely positive.

Proof For any $n \in \mathbb{N}$ and $x_1, \dots, x_n \in A$, $\xi_1, \dots, \xi_n \in \mathcal{H}$, we have

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | \Phi(x_i^* x_j) \xi_j \rangle &= \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | V^* \pi(x_i^* x_j) V \xi_j \rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | V^* \pi(x_i)^* \pi(x_j) V \xi_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \langle \pi(x_i) V \xi_i | \pi(x_j) V \xi_j \rangle \\
&= \left\| \sum_{i=1}^n V \pi(x_i) \xi_i \right\|^2 \geq 0.
\end{aligned}$$

□

A remarkable so-called *dilation theorem* due to Stinespring [1] says that under certain assumptions on A the converse of the above proposition is also true. There are now various versions of this result. In our treatment, Stinespring's theorem for a C^* -algebra with an identity will be a corollary of a bilinear dilation theorem to be considered next.

7.2 A Bilinear Dilation Theorem

In the following theorem, let $\mathcal{A}$ be an algebra of subsets of a set Ω and let A be the closed $*$ -subalgebra of the C^* -algebra of all bounded complex functions $f : \Omega \rightarrow \mathbb{C}$ (with the sup-norm) obtained as the closure of the space of the linear combinations of the characteristic functions of the sets in $\mathcal{A}$. (Thus if $\mathcal{A}$ is a σ -algebra, then A is the space of bounded $\mathcal{A}$ -measurable functions on Ω .) Moreover, let B be a C^* -algebra with identity 1.

If B_1 and B_2 are C^* -algebras and $F : B_1 \times B_2 \rightarrow \mathcal{L}(\mathcal{H})$ is a bilinear map such that $F(x^* x, y^* y) \geq 0$ for all $x \in B_1, y \in B_2$, then F is said to be *positive*. A bilinear map $F : B_1 \times B_2 \rightarrow \mathcal{L}(\mathcal{H})$ is said to be *bounded* if the set of the values it takes on the Cartesian product of the unit balls of B_1 and B_2 is norm bounded. It is equivalent to require that F be continuous with respect to the product topology of $B_1 \times B_2$.

Theorem 7.1 *Let $\Psi : A \times B \rightarrow \mathcal{L}(\mathcal{H})$ be a positive bilinear map such that for all $X \in \mathcal{A}$ the linear map $y \mapsto \Psi(\chi_X, y) \in \mathcal{L}(\mathcal{H})$ on B is completely positive.*

- (a) Ψ is bounded.
- (b) *There is a Hilbert space $\mathcal{K}$ with unital $*$ -representations $\pi : A \rightarrow \mathcal{L}(\mathcal{K})$ and $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ and a bounded linear map $V : \mathcal{H} \rightarrow \mathcal{K}$ such that $\pi(x)\varrho(y) = \varrho(y)\pi(x)$ and $\Psi(x, y) = V^* \pi(x) \varrho(y) V$ for all $x \in A, y \in B$.*

Proof (a) The claim is true even if A and B are not assumed to have identities (see, e.g., [2, p. 159]). In our present unital case it is clear since fixing $\varphi \in \mathcal{H}$ we can consider $\langle \varphi | \Psi(\cdot, \cdot) \varphi \rangle$ and use twice the fact that for a positive linear form f on a unital C^* -algebra $\|f\| = f(1)$ (see, e.g., [3, p. 38]).

(b) Consider the algebraic tensor product $K_0 = A \otimes B \otimes \mathcal{H}$. For $u = \sum_{i=1}^m x_{1i} \otimes y_{1i} \otimes \xi_i \in K_0$ and $v = \sum_{j=1}^n x_{2j} \otimes y_{2j} \otimes \eta_j \in K_0$ we define

$$\langle v | u \rangle_0 = \sum_{i=1}^m \sum_{j=1}^n \langle \eta_j | \Psi(x_{2j}^* x_{1i}, y_{2j}^* y_{1i}) \xi_i \rangle.$$

Using Definition 6.10 it is easy to see that in this way we get a well-defined sesquilinear form $\langle \cdot | \cdot \rangle_0$ on K_0 . We leave the details as an exercise.

We show that $\langle u | u \rangle_0 \geq 0$ for all $u \in K_0$. Choose $\xi_1, \dots, \xi_n \in \mathcal{H}$, $y_1, \dots, y_n \in B$, and let $f_1, \dots, f_n$ be linear combinations of characteristic functions of sets in $\mathcal{A}$. We claim that

$$\sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | \Psi(f_i^* f_j, y_i^* y_j) \xi_j \rangle \geq 0. \quad (7.1)$$

We can find disjoint sets $X_\nu \in \mathcal{A}$ and complex numbers $c_{i\nu}$ for $i = 1, \dots, n$, $\nu = 1, \dots, p$ such that $f_i = \sum_{\nu=1}^p c_{i\nu} \chi_{X_\nu}$. Since

$$f_i^* f_j = \left(\sum_{\nu=1}^p \bar{c}_{i\nu} \chi_{X_\nu} \right) \left(\sum_{\mu=1}^p c_{j\mu} \chi_{X_\mu} \right) = \sum_{\nu=1}^p \bar{c}_{i\nu} c_{j\nu} \chi_{X_\nu},$$

we get

$$\begin{aligned} \sum_{i,j=1}^n \left\langle \xi_i \left| \Psi \left(\sum_{\nu=1}^p \bar{c}_{i\nu} c_{j\nu} \chi_{X_\nu}, y_i^* y_j \right) \xi_j \right. \right\rangle &= \sum_{\nu=1}^p \sum_{i,j=1}^n \langle c_{i\nu} \xi_i | \Psi(c_{j\nu} \chi_{X_\nu}, y_i^* y_j) \xi_j \rangle \\ &= \sum_{\nu=1}^p \left\langle \sum_{i=1}^n c_{i\nu} \xi_i \left| \Psi(\chi_{X_\nu}, y_i^* y_j) \sum_{j=1}^n c_{j\nu} \xi_j \right. \right\rangle \geq 0, \end{aligned}$$

since each map $y \mapsto \Psi(\chi_{X_\nu}, y)$ on B is completely positive. As Ψ is bounded, (7.1) also holds for any $f_i \in A$ since f_i can be approximated in norm by linear combinations of characteristic functions of sets in $\mathcal{A}$. Thus $\langle u | u \rangle_0 \geq 0$ for all $u \in K_0$.

Denote $N = \{u \in K_0 | \langle u | u \rangle_0 = 0\}$. The Cauchy-Schwarz inequality (as noted in Remark 2.1) implies that $N = \{u \in K_0 | \langle u | v \rangle_0 = 0 \text{ for all } v \in K_0\}$. It follows that N is a vector subspace of K_0 and the quotient space $\tilde{K}_0 = K_0/N$ is an inner product space with the inner product defined by $\langle v + N | u + N \rangle = \langle v | u \rangle_0$. Let $\mathcal{K}$ denote the Hilbert space completion of $\tilde{K}_0$, i.e the (up to an isometric isomorphism unique) Hilbert space containing $\mathcal{K}$ as a dense subspace. For any $a \in A$ there is a unique linear map $T_a : K_0 \rightarrow K_0$ satisfying $T_a(x \otimes y \otimes \xi) = ax \otimes y \otimes \xi$ for all $x \in A$, $y \in B$ and $\xi \in \mathcal{H}$. Fix an element $u = \sum_{i=1}^m x_i \otimes y_i \otimes \xi_i \in K_0$, and define $\omega : A \rightarrow \mathbb{C}$ by $\omega(a) = \langle u | T_a u \rangle_0$. A direct computation shows that $\omega(a^*a) = \langle u | T_{a^*a} u \rangle_0 =$

$\langle T_a u | T_a u \rangle_0 \geq 0$, and so ω is a positive linear form on A and hence continuous [3, p. 38]. We get

$$\langle T_a u | T_a u \rangle_0 = \omega(a^* a) \leq \|a^* a\| \|\omega\| = \|a^* a\| \omega(1) = \|a\|^2 \langle u | u \rangle_0.$$

It follows that for each $x \in A$ we have a well-defined bounded linear map $\tilde{T}_x : \tilde{K}_0 \rightarrow \tilde{K}_0$ such that $\tilde{T}_x(u + N) = T_x u + N$ for all $u \in K_0$ and $\|\tilde{T}_x\| \leq \|x\|$. We denote by $\pi(x) : \mathcal{K} \rightarrow \mathcal{K}$ the continuous linear extension of $\tilde{T}_x$. Then a direct calculation shows that π is a $*$ -representation. Similarly we can construct a $*$ -representation $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ satisfying $\varrho(b)[(x \otimes y \otimes \xi) + N] = x \otimes b y \otimes \xi + N$ for all $x \in A$, $y \in B$, $\xi \in \mathcal{H}$. The mapping $V : \mathcal{H} \rightarrow \mathcal{K}$ is defined by setting $V\xi = 1 \otimes 1 \otimes \xi + N$, and it is easy to see that the claims in (b) are valid. $\square$

Remark 7.3 Conversely, a bilinear analogue of Proposition 7.2 can be proved by a direct calculation. $\triangleleft$

If, with the assumptions of Theorem 7.1, $\mathcal{K}$ is a Hilbert space, $\pi : A \rightarrow \mathcal{L}(\mathcal{K})$ and $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ are $*$ -representations and $V : \mathcal{H} \rightarrow \mathcal{K}$ is a bounded linear map such that $\pi(x)\varrho(y) = \varrho(y)\pi(x)$ and $\Psi(x, y) = V^*\pi(x)\varrho(y)V$ for all $x \in A$, $y \in B$, then we say that the quadruple $(\mathcal{K}, \pi, \varrho, V)$ is a *Stinespring type representation* for Ψ . It is called *unital* if the $*$ -representations π and ϱ are unital. The Stinespring type representation $(\mathcal{K}, \pi, \varrho, V)$ for Ψ is said to be *minimal* if the linear combinations of the vectors $\pi(x)\varrho(y)V\xi$ with $x \in A$, $y \in B$ and $\xi \in \mathcal{H}$ are dense in $\mathcal{K}$.

We keep the assumptions of Theorem 7.1 in the following elaboration of the result.

Theorem 7.2 (a) *There is a minimal Stinespring type representation for Ψ .*

- (b) *If $(\mathcal{K}, \pi, \varrho, V)$ is a minimal Stinespring type representation for Ψ , then it is unital, i.e. $\pi(1) = \varrho(1) = I_{\mathcal{K}}$.*
- (c) *If $(\mathcal{K}_1, \pi_1, \varrho_1, V_1)$ and $(\mathcal{K}_2, \pi_2, \varrho_2, V_2)$ are two minimal Stinespring type representations for Ψ , there is a unique unitary map $U : \mathcal{K}_1 \rightarrow \mathcal{K}_2$ such that $UV_1 = V_2$ and $U\pi_1(x) = \pi_2(x)U$, $U\varrho_1(y) = \varrho_2(y)U$ for all $x \in A$, $y \in B$.*
- (d) *If $(\mathcal{K}, \pi, \varrho, V)$ is a minimal Stinespring type representation for Ψ satisfying $\Psi(1, 1) = I_{\mathcal{H}}$, then V is isometric.*

Proof (a) Let $(\mathcal{K}, \pi, \varrho, V)$ be the Stinespring type representation for Ψ constructed in the proof of Theorem 7.1. Denote by $\tilde{\mathcal{K}}$ the closed linear span of the set of vectors $\pi(x)\varrho(y)V\xi$ where $x \in A$, $y \in B$ and $\xi \in \mathcal{H}$. Then $\tilde{\mathcal{K}}$ is invariant under $\pi(x)$ and $\varrho(y)$ for all $x \in A$ and $y \in B$, and we define $\tilde{\pi}(x) : \tilde{\mathcal{K}} \rightarrow \tilde{\mathcal{K}}$ by restricting $\pi(x)$ to $\tilde{\mathcal{K}}$. Since $\pi(1)\varrho(1)V\xi = V\xi$, we have $V\xi \in \tilde{\mathcal{K}}$ for all $\xi \in \mathcal{H}$, and we may choose $\tilde{V}$ to be V interpreted as a map from $\tilde{\mathcal{K}}$ into itself. Then $(\tilde{\pi}, \tilde{V}, \tilde{\mathcal{K}})$ is a minimal Stinespring type representation for Ψ .

(b) Since $\pi(1)\pi(x)\varrho(y)V\xi = \pi(1x)\varrho(y)V\xi$ it follows that $\pi(1) = I_{\mathcal{K}}$, and similarly $\varrho(y) = I_{\mathcal{K}}$.

(c) To see the uniqueness of U , note that

$$U\pi_1(x)\varrho_1(y)V_1\xi = \pi_2(x)U\varrho_1(y)V_1\xi = \pi_2(x)\varrho_2(y)UV_1\xi = \pi_2(x)\varrho_2(y)V_2\xi,$$

so that the values of U are determined on a dense subspace of $\mathcal{K}_1$. We now show the existence of U . For any $x_1, \dots, x_n \in A$, $y_1, \dots, y_n \in B$ and $\xi_1, \dots, \xi_n \in \mathcal{H}$,

$$\begin{aligned} \left\| \sum_{j=1}^n \pi_1(x_j) \varrho(y_j) V_1 \xi_j \right\|^2 &= \left\langle \sum_{i=1}^n \pi_1(x_i) \varrho(y_i) V_1 \xi_i \left| \sum_{j=1}^n \pi_1(x_j) \varrho(y_j) V_1 \xi_j \right. \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle V_1 \xi_i | \pi(x_i^* x_j) \varrho(y_i^* y_j) V_1 \xi_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | V_1^* \pi(x_i^* x_j) \varrho(y_i^* y_j) V_1 \xi_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | \Psi(x_i^* x_j, y_i^* y_j) \xi_j \rangle = \left\| \sum_{j=1}^n \pi_2(x_j) \varrho(y_j) V_2 \xi_j \right\|^2. \end{aligned}$$

This shows that we get a well-defined unitary map $U : \mathcal{K}_1 \rightarrow \mathcal{K}_2$ such that

$$U(\pi_1(x) \varrho(y) V_1 \xi) = \pi_2(x) \varrho(y) V_2 \xi.$$

In particular, using (b) we get $U V_1 \xi = U(\pi_1(1) \varrho_1(1) V_1 \xi) = V_2 \xi$ for all $\xi \in \mathcal{H}$. Finally, if $x \in A$, then $U \pi_1(x) \pi_1(u) \varrho_1(y) V_1 \xi = U \pi_1(xu) \varrho_1(y) V_1 \xi = \pi_2(xu) \varrho_2(y) V_2 \xi = \pi_2(x) \pi_2(u) \varrho_2(y) V_2 \xi = \pi_2(x) U \pi_1(u) \varrho_1(y) V_1 \xi$, for all $u \in A$, $y \in A$, $\xi \in \mathcal{H}$, which implies that $U \pi_1(x) = \pi_2(x) U$. A similar argument shows that $U \varrho_1(y) = \varrho_2(y) U$ for all $y \in B$.

(d) Using (b) we get $\|V\xi\|^2 = \|\pi(1) \varrho(1) V\xi\|^2 = \langle V\xi | V^* \pi(1) \varrho(1) V\xi \rangle = \langle \xi | \Psi(1, 1) \xi \rangle = \|\xi\|^2$ for all $\xi \in \mathcal{H}$. $\square$

We may thus use the definite article in connection with minimal Stinespring type representations in the kind of situations we are considering in this section.

In the next theorem and the remark following it we still retain our basic setup for the mapping $\Psi : A \times B \rightarrow \mathcal{L}(\mathcal{H})$ and let $(\mathcal{K}, \pi, \varrho, V)$ be a minimal Stinespring type representation for Ψ . We get some further information under extra assumptions.

Theorem 7.3 (a) *If B is a von Neumann algebra and the map $y \mapsto \Psi(\chi_X, y)$ on B is normal for every $X \in \mathcal{A}$, then $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ is a normal $*$ -representation.*
 (b) *If $\mathcal{A}$ is a σ -algebra and the mapping $X \mapsto \Psi(\chi_X, y)$ is a positive operator measure on $\mathcal{A}$ for every positive $y \in B$, then the map $X \mapsto \pi(\chi_X)$ from $\mathcal{A}$ into $\mathcal{L}(\mathcal{K})$ is a spectral measure.*

Proof (a) Let B_* be the predual of B , i.e. the Banach space of the σ -weakly continuous linear functionals on B . If $u, v \in A$, $y, z \in B$ and $\xi, \eta \in \mathcal{H}$, then it follows from our assumptions that the linear functional

$$w \mapsto \langle \pi(u) \varrho(y) V\xi | \varrho(w) \pi(v) \varrho(z) V\eta \rangle = \langle \xi | \Psi(u^* v, y^* w^* z) \xi \rangle$$

on B belongs to B_* , since the map $w \mapsto y^* w^* z$ on B is normal. It follows that the transpose of π maps a dense subset of the predual $\mathcal{L}(\mathcal{K})_* = \mathcal{T}(\mathcal{K})$ of $\mathcal{L}(\mathcal{K})$ into B_* , and since B_* is norm closed the same is true of the whole of $\mathcal{L}(\mathcal{K})_*$. Thus π is normal.

(b) The proof of part (a) works here with obvious modifications. Indeed, instead of B_* we now consider the Banach space $\text{ca}(\Omega, \mathcal{A})$ of complex measures on $\mathcal{A}$, equipped with the total variation norm $\|\mu\| = |\mu|(\Omega)$, as a closed subspace of the dual of A . $\square$

Remark 7.4 If $(\mathcal{K}, \pi, \varrho, V)$ is a minimal Stinespring representation for $\Psi : A \times B \rightarrow \mathcal{L}(\mathcal{H})$ where the Hilbert space $\mathcal{H}$ and the unital C^* -algebras A and B are assumed to be separable, then it is straightforward to verify that the Hilbert space $\mathcal{K}$ is also separable. It is also clear that $\mathcal{K}$ is finite-dimensional if $\mathcal{H}$, A and B are finite-dimensional. For many central applications, however, even the former observation is useless. For example, the type of commutative C^* -algebra A discussed above is nonseparable even if we take $\Omega = \mathbb{N}$ and $\mathcal{A} = 2^{\mathbb{N}}$. Likewise, $B = \mathcal{L}(\mathcal{H}_0)$ is not separable if $\mathcal{H}_0$ is infinite-dimensional. We next prove a result which in many cases remedies the situation. $\triangleleft$

In the following result we assume that A is the space of bounded $\mathcal{A}$ -measurable complex functions on Ω where $\mathcal{A} \subset 2^{\Omega}$ is a σ -algebra. We say that a set $D \subset A$ is *measure separating* if zero is the only measure $\mu \in \text{ca}(\Omega, \mathcal{A})$ such that $\int_{\Omega} f d\mu = 0$ for all $f \in D$. We use without explicit formulation also the obvious one-variable versions of the next result.

Proposition 7.3 *Assume that there is a countable measure separating set $D \subset A$ and suppose that $B = \mathcal{L}(\mathcal{H}_0)$ for a separable Hilbert space $\mathcal{H}_0$. Let $\mathcal{H}$ be a separable Hilbert space and $\Psi : A \times B \rightarrow \mathcal{L}(\mathcal{H})$ a positive bilinear map such that for every $X \in A$ the map $y \mapsto \Psi(\chi_X, y)$ is a normal completely positive linear map on $\mathcal{L}(\mathcal{H}_0)$ and for every positive $y \in B$ the map $X \mapsto \Psi(\chi_X, y)$ is a positive operator measure on $\mathcal{A}$. If $(\mathcal{K}, \pi, \varrho, V)$ is a minimal Stinespring type representation for Ψ then the Hilbert space $\mathcal{K}$ is separable.*

Proof In this proof we use some functional analytic methods we have not previously introduced. The relevant notions and auxiliary results may be found, for instance, in [4]. We will use the locally convex topologies $\sigma(A, \text{ca}(\Omega, \mathcal{A}))$ and $\sigma(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_0)_*)$ on A and $\mathcal{L}(\mathcal{H}_0)$, respectively, and the norm topology on $\mathcal{H}$. (As will be explained in Remark 7.5 below that the former is Hausdorff, too.) Applying Theorem 7.3 we see that for any $\eta \in \mathcal{K}$ the map

$$(x, y, \xi) \mapsto \langle \eta | \pi(x) \varrho(y) V \xi \rangle = \theta_{\eta}(x, y, \xi)$$

is separately continuous with respect to these topologies. (For example, note that $\langle \eta | \pi(x) \varrho(y) V \xi \rangle = \int x d\mu(x)$, where $\mu(X) = \langle \eta | \pi(\chi_X) \varrho(y) V \xi \rangle$, and on the other hand $\langle \eta | \pi(x) \varrho(y) V \xi \rangle = \langle \pi(x)^* \eta | \varrho(y) \xi \rangle$.)

We consider the natural duality $\langle A, \text{ca}(\Omega, \mathcal{A}) \rangle$ (via integration). From the bipolar theorem (see [4]), it follows that the linear span of D is $\sigma(A, \text{ca}(\Omega, \mathcal{A}))$ -dense in A . Since D is countable, the set $\tilde{D}$ of the linear combinations of its elements with rational coefficients is countable, and it is clearly also $\sigma(A, \text{ca}(\Omega, \mathcal{A}))$ -dense in A . Let $\{\phi_1, \phi_2, \phi_3, \dots\}$ be a (countable) orthonormal basis of $\mathcal{H}_0$, and for $n = 1, 2, \dots$

let P_n denote the projection onto the (automatically closed) linear span of the set $\{\phi_1, \dots, \phi_n\}$. For any $y \in B = \mathcal{L}(\mathcal{H}_0)$ the sequence (yP_n) converges to y with respect to $\sigma(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_0)_*)$. Since each yP_n is a finite rank and hence compact operator, it follows that $\mathcal{C}(\mathcal{H}_0)$, the space of the compact operators on $\mathcal{H}_0$, is $\sigma(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_0)_*)$ -dense in $\mathcal{L}(\mathcal{H}_0)$. (This follows also from the bipolar theorem.) Since $\mathcal{C}(\mathcal{H}_0)$ is (even norm) separable (exercise), it follows that $\mathcal{L}(\mathcal{H}_0)$ has a countable $\sigma(\mathcal{L}(\mathcal{H}_0), \mathcal{L}(\mathcal{H}_0)_*)$ -dense subset E .

Choose a norm dense subset $F \subset \mathcal{H}$. Let S denote the closed linear span of the set $\{\pi(x)\varrho(y)V\xi \mid x \in \tilde{D}, y \in E, \xi \in F\}$. We claim that $S = \mathcal{K}$. This will prove the separability of $\mathcal{K}$, since then obviously the countable set of the linear combinations, with rational coordinates, of the vectors $\pi(x)\varrho(y)V\xi$ with $x \in \tilde{D}, y \in E, \xi \in F$ will also be dense in $\mathcal{K}$. To see that $S = \mathcal{K}$, take an arbitrary $\eta \in S^\perp$. Using the separate continuity of the map θ_η defined at the beginning of the proof it is easily seen in three stages that $\langle \eta \mid \pi(x)\varrho(y)V\xi \rangle = 0$ for all $x \in A, y \in B$. Since the linear span of the set $\{\pi(x)\varrho(y)V\xi \mid x \in \tilde{D}, y \in B, \xi \in \mathcal{H}\}$ is by definition dense in $\mathcal{K}$, we must have $\eta = 0$. Thus $S = \mathcal{K}$. $\square$

Remark 7.5 (a) A usable sufficient condition for the existence of a countable measure separating set $D \subset A$ is that the σ -algebra $\mathcal{A}$ be *countably generated*, which means that there is a countable subset $\mathcal{C} \subset \mathcal{A}$ generating $\mathcal{A}$. For then the algebra $\mathcal{F}$ generated by $\mathcal{C}$ is easily seen to be countable, and we can take $D = \{\chi_X \mid X \in \mathcal{F}\}$. Indeed if $\mu \in \text{ca}(\Omega, \mathcal{A})$ is such that $\mu(X) = 0$ for all $X \in \mathcal{F}$, then $\mu(X) = 0$ for every X in the monotone class generated by $\mathcal{F}$ (since the set of those $X \in \mathcal{A}$ for which $\mu(X) = 0$ is a monotone class). But by the monotone class theorem [5, p. 27] the monotone class generated by $\mathcal{F}$ is $\mathcal{A}$.

(b) For each $a \in \Omega$ denote by $\delta_a : \mathcal{A} \rightarrow \{0, 1\}$ the *point measure* for which $\delta_a(X) = 1$ if and only if $a \in X$. If $f : \Omega \rightarrow \mathbb{C}$ is an $\mathcal{A}$ -measurable function, then it is the norm limit of a sequence of $\mathcal{A}$ -simple functions f_n , and we can clearly assume that $f_n(a) = f(a)$ for each $n \in \mathbb{N}$. Then $\int f_n d\delta_a = f(a)$, and it follows that $\int f d\delta_a = f(a)$. Thus already the point measures suffice to separate the elements of A , and so $\sigma(A, \text{ca}(\Omega, \mathcal{A}))$ is a Hausdorff topology. $\triangleleft$

7.3 The Stinespring and Naimark Dilation Theorems

Here we present two classic dilation theorems whose proofs will be reduced to the results in the preceding section.

Theorem 7.4 *Let B be a C^* -algebra with identity and $\Phi : B \rightarrow \mathcal{L}(\mathcal{H})$ a completely positive linear map. There exist a Hilbert space $\mathcal{K}$, a bounded linear map $V : \mathcal{H} \rightarrow \mathcal{K}$ and a unital $*$ -representation $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ such that $\Phi(y) = V^*\varrho(y)V$ for all $y \in B$.*

Proof It is clearly possible to adapt the proof of Theorem 7.1 to this case, but we may also deduce our present claim from that result directly as follows. Choose $\Omega = \{1\}$ and $\mathcal{A} = \{\emptyset, \Omega\}$, so that we may identify $A = \mathbb{C}$. We now define $\Psi(x, y) = x\Phi(y)$ for $x \in A$ and $y \in B$. Applying Theorem 7.1 to this Ψ we get the claim. $\square$

Similarly, the rest of the preceding section has an obvious translation to the one variable case (essentially taking A or B to be $\mathbb{C}$); the proofs need not be repeated. With the assumptions of Theorem 7.4 we say that the triple $(\mathcal{K}, \varrho, V)$ is a *Stinespring representation* for Φ if $\mathcal{K}$ is a Hilbert space, $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ is a $*$ -representation and $V : \mathcal{H} \rightarrow \mathcal{K}$ is a bounded linear map such that $\Phi(y) = V^*\varrho(y)V$ for all $y \in B$. It is called *unital* if the $*$ -representation ϱ is unital. The Stinespring representation $(\mathcal{K}, \varrho, V,)$ for Φ is said to be *minimal* if the linear combinations of the vectors $\varrho(y)V\xi$ with $y \in B$ and $\xi \in \mathcal{H}$ are dense in $\mathcal{K}$. In this situation ϱ is also called a *minimal Stinespring dilation* of Φ . We keep the assumptions of Theorem 7.4 in the following obvious consequence of Theorems 7.2 and 7.3.

- Theorem 7.5** (a) *There is a minimal Stinespring representation for Φ .*
 (b) *If $(\mathcal{K}, \varrho, V)$ is a minimal Stinespring representation for Φ , then $\varrho(1) = I_{\mathcal{K}}$.*
 (c) *If $(\mathcal{K}_1, \varrho_1, V_1)$ and $(\mathcal{K}_2, \varrho_2, V_2,)$ are two minimal Stinespring representations for Φ , there is a unique unitary map $U : \mathcal{K}_1 \rightarrow \mathcal{K}_2$ such that $UV_1 = V_2$ and $U\varrho_1(y) = \varrho_2(y)U$ for all $y \in B$.*
 (d) *If $(\mathcal{K}, \varrho, V)$ is a minimal Stinespring representation for Φ and $\Phi(1) = I_{\mathcal{H}}$, then V is isometric.*
 (e) *If B is a von Neumann algebra and the completely positive map Φ is normal, then in the minimal Stinespring representation $(\mathcal{K}, \varrho, V)$ for Φ the $*$ -representation $\varrho : B \rightarrow \mathcal{L}(\mathcal{K})$ is normal.*

In the rest of this section we consider the special case of the two-variable Stinespring type theory of the previous section where this time the second variable is in a one-dimensional space. Again, we can list some immediate consequences without repeating the proofs. Our basic setting is part of that in the preceding section: we let $\mathcal{A}$ be an algebra of subsets of a set Ω and let A be the closed $*$ -subalgebra of the C^* -algebra of all bounded complex functions $f : \Omega \rightarrow \mathbb{C}$ (with the sup-norm) obtained as the closure of the space of the linear combinations of the characteristic functions of the sets in $\mathcal{A}$.

Theorem 7.6 *Let $F : A \rightarrow \mathcal{L}(\mathcal{H})$ be a positive linear map. There exist a Hilbert space $\mathcal{K}$, a bounded linear map $V : \mathcal{H} \rightarrow \mathcal{K}$ and a unital $*$ -representation $\pi : A \rightarrow \mathcal{L}(\mathcal{K})$ such that $F(x) = V^*\pi(x)V$ for all $x \in A$.*

A new proof is not needed as the choice $B = \mathbb{C}$ and $\Psi(x, y) = yF(x)$ will essentially reduce the theorem to Theorem 7.1. It is worth observing that the proof of Theorem 7.1 shows that the positive linear map F on the commutative C^* -algebra A is actually completely positive. (Of course this is also seen from Theorem 7.6 itself in conjunction with Proposition 7.2.) If, with the assumptions of Theorem 7.6, $\mathcal{K}$ is a Hilbert space, $\pi : A \rightarrow \mathcal{L}(\mathcal{K})$ is a $*$ -representation and $V : \mathcal{H} \rightarrow \mathcal{K}$ is a bounded

linear map, we say that the triple $(\mathcal{K}, \pi, V)$ is a *Naimark representation* for F . It is *unital* if $\pi(1) = I_{\mathcal{K}}$. The Naimark representation $(\mathcal{K}, \pi, V)$ for F is said to be *minimal*, if the linear combinations of the vectors $\pi(x)V\xi$ with $x \in A$ and $\xi \in \mathcal{H}$ are dense in $\mathcal{K}$. In this situation π is also called a *minimal Naimark dilation* of F . Again, keeping the assumptions of Theorem 7.6, we get the following consequence of Theorems 7.2 and 7.3.

- Theorem 7.7** (a) *There is a minimal Naimark representation for F .*
 (b) *If $(\mathcal{K}, \pi, V)$ is a minimal Naimark representation for F , then $\pi(1) = I_{\mathcal{K}}$.*
 (c) *If $(\mathcal{K}_1, \pi_1, V_1)$ and $(\mathcal{K}_2, \pi_2, V_2)$ are two minimal Naimark representations for F , there is a unique unitary map $U : \mathcal{K}_1 \rightarrow \mathcal{K}_2$ such that $UV_1 = V_2$ and $U\pi_1(x) = \pi_2(x)U$ for all $x \in A$.*
 (d) *If $(\mathcal{K}, \pi, V)$ is a minimal Naimark representation for F and $F(1) = I_{\mathcal{H}}$, then V is isometric.*
 (e) *If $\mathcal{A}$ is a σ -algebra and the mapping $X \mapsto F(\chi_X)$ is a positive operator measure on $\mathcal{A}$, then in the minimal Naimark representation $(\mathcal{K}, \pi, V)$ for F the map $X \mapsto \pi(\chi_X)$ from $\mathcal{A}$ into $\mathcal{L}(\mathcal{K})$ is a spectral measure.*

Remark 7.6 We will also make use of the one-variable analogues of Proposition 7.3 though we do not formulate them explicitly. Given the manifold appearances of dilations of semispectral measures in this text, we introduce the notation $(\mathcal{K}, F, V)$ for a dilation of a semispectral measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ into a spectral measure $F : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{K})$, with $E(X) = V^*F(X)V$, $X \in \mathcal{A}$, and $V : \mathcal{H} \rightarrow \mathcal{K}$ being an isometry. The minimality condition then says that the linear combinations of the vectors $F(X)V\phi$, $X \in \mathcal{A}$, $\phi \in \mathcal{H}$, are dense in $\mathcal{K}$. $\triangleleft$

For the interested reader we conclude with a somewhat different form of the Naimark dilation theorem (which we will not use later on).

Theorem 7.8 *Let Ω be a locally compact Hausdorff space and $\Phi : C_0(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ a positive linear map. Then there is a Hilbert space $\mathcal{K}$ with a bounded linear map $V : \mathcal{H} \rightarrow \mathcal{K}$ and a $*$ -representation $\pi : C_0(\Omega) \rightarrow \mathcal{L}(\mathcal{K})$ such that $\Phi(f) = V^*\pi(f)V$ for all $f \in C_0(\Omega)$. In particular, Φ is completely positive.*

Proof By Theorem 4.4, there is a unique positive operator measure $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ such that for all $\varphi, \psi \in \mathcal{H}$ the function $X \mapsto \langle \psi | E(X)\varphi \rangle = E_{\varphi, \psi}(X)$ is a regular complex Borel measure and $\langle \psi | \Phi(f)\varphi \rangle = \int_{\Omega} f dE_{\varphi, \psi}$. From the proof of Theorem 7.1 we see that the function $f \mapsto \int_{\Omega} f dE$ is completely positive on the C^* -algebra of bounded Borel functions $f : \Omega \rightarrow \mathbb{C}$, and thus on $C_0(\Omega)$. Since its restriction to $C_0(\Omega)$ is Φ , the claim follows from the Stinespring theorem. $\square$

7.4 Normal Completely Positive Operators from $\mathcal{L}(\mathcal{H})$ into $\mathcal{L}(\mathcal{K})$

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Normal completely positive operators from $\mathcal{L}(\mathcal{H})$ into $\mathcal{L}(\mathcal{K})$ have a special form which is useful in many applications. The present section is devoted to proving this so-called Kraus representation. The proof depends on a version of the Stinespring dilation theorem combined with the special formulation of a nondegenerate normal representation of $\mathcal{L}(\mathcal{H})$ as a direct sum of identity representations.

In the first stage we study a class of normal linear maps expressible in a sum form. We have discussed the convergence of a net in a topological space. It is also clear how to define a Cauchy net $(x_n)_{n \in \Lambda}$ in a metric space (X, d) : for each $\epsilon > 0$ there is $p \in \Lambda$ such that $d(x_n, x_m) < \epsilon$ whenever $n, m \geq p$. We need the following auxiliary result. For the proof, see e.g. [6, p. 28].

Lemma 7.2 *In a complete metric space (X, d) every Cauchy net converges.*

Lemma 7.3 *Suppose $A \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ and $B \in \mathcal{L}(\mathcal{K}, \mathcal{H})$. If $T \in \mathcal{T}(\mathcal{K})$, then $BT A \in \mathcal{T}(\mathcal{H})$ and for any $S \in \mathcal{L}(\mathcal{H})$ we have $\text{tr}[SBTA] = \text{tr}[ASBT]$.*

Proof We denote $\mathcal{H}_0 = \mathcal{H} \oplus \mathcal{K}$ and identify $\mathcal{H}$ and $\mathcal{K}$ with closed subspaces of $\mathcal{H}_0$. Let P be the projection of $\mathcal{H}_0$ onto $\mathcal{H}$ and Q the projection of $\mathcal{H}_0$ onto $\mathcal{K}$. We define the following elements of $\mathcal{L}(\mathcal{H}_0)$: $A_0 = AP$, $B_0 = BQ$, $S_0 = SP$, and $T_0 = TQ$. Choose an orthonormal basis K_1 for $\mathcal{H}$ and an orthonormal basis K_2 for $\mathcal{K}$. Then $K = K_1 \cup K_2$ is an orthonormal basis for $\mathcal{H}_0$, and using these bases it is easy to show that an operator $C \in \mathcal{L}(\mathcal{H})$ is in $\mathcal{T}(\mathcal{H})$ if and only if $CP \in \mathcal{T}(\mathcal{H}_0)$, and then $\text{tr}[C] = \text{tr}[CP]$. A similar statement holds for $D \in \mathcal{L}(\mathcal{K})$ and DQ . In particular, since $BTAP = BTPAP \in \mathcal{T}(\mathcal{H}_0)$, we have $BT A \in \mathcal{T}(\mathcal{H})$, and $\text{tr}[SBTA] = \text{tr}[SBTAP] = \text{tr}[SPBQTQAP] = \text{tr}[S_0B_0T_0A_0] = \text{tr}[A_0S_0B_0T_0] = \text{tr}[ASBT]$ whenever $S \in \mathcal{L}(\mathcal{H})$ and we write $S_0 = SP$. $\square$

In this section the statement “ $\sum_{\lambda \in L} A_\lambda^* A_\lambda$ exists in the sense of weak convergence” means that the net $F \mapsto \sum_{\lambda \in F} A_\lambda^* A_\lambda$ defined on the directed set of the finite subsets of the set L converges weakly, and we have a similar usage in analogous situations. In case L is countable it is easy to see that this entails the convergence of the corresponding series, and the sum is independent of the way the set L is arranged in a sequence.

Lemma 7.4 *Let $(A_\lambda)_{\lambda \in L}$ be a family in $\mathcal{L}(\mathcal{K}, \mathcal{H})$ such that $\sum_{\lambda \in L} A_\lambda^* A_\lambda$ exists in the sense of weak convergence. Then, for any $T \in \mathcal{T}(\mathcal{K})$, $\sum_{\lambda \in L} A_\lambda T A_\lambda^*$ exists in $\mathcal{T}(\mathcal{H})$, where the sum converges in the trace class norm topology.*

Proof Since $\mathcal{T}(\mathcal{H})$ is complete, in view of Lemma 7.2 it is enough to show that for the directed set $\mathcal{F}(L)$ of the finite subsets of L , $(\sum_{\lambda \in F} A_\lambda T A_\lambda^*)_{F \in \mathcal{F}(L)}$ is a Cauchy net in $\mathcal{T}(\mathcal{H})$. We may assume that $T \geq 0$. Since the net $(\sum_{\lambda \in F} A_\lambda^* A_\lambda)_{F \in \mathcal{F}(L)}$ is norm bounded it converges to some B in the σ -weak topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$. It follows that for any $\epsilon > 0$ we can find $F_0 \in \mathcal{F}(L)$ such that

$$\left| \operatorname{tr} \left[\sum_{\lambda \in F} A_\lambda T A_\lambda^* - B T \right] \right| = \left| \operatorname{tr} \left[\sum_{\lambda \in F} A_\lambda^* A_\lambda T - B T \right] \right| < \frac{\epsilon}{4}$$

for any $F \supset F_0$ in $\mathcal{F}(L)$ (see Lemma 7.3). For any $F_1 \supset F_0$ and $F_2 \supset F_0$ in $\mathcal{F}(L)$,

$$\begin{aligned} \left\| \sum_{\lambda \in F_1} A_\lambda T A_\lambda^* - \sum_{\lambda \in F_2} A_\lambda T A_\lambda^* \right\|_1 &= \left\| \sum_{\lambda \in F_1 \setminus F_2} A_\lambda T A_\lambda^* - \sum_{\lambda \in F_2 \setminus F_1} A_\lambda T A_\lambda^* \right\|_1 \\ &\leq \left\| \sum_{\lambda \in F_1 \setminus F_2} A_\lambda T A_\lambda^* \right\|_1 + \left\| \sum_{\lambda \in F_2 \setminus F_1} A_\lambda T A_\lambda^* \right\|_1 \\ &= \operatorname{tr} \left[\sum_{\lambda \in F_1 \setminus F_2} A_\lambda T A_\lambda^* \right] + \operatorname{tr} \left[\sum_{\lambda \in F_2 \setminus F_1} A_\lambda T A_\lambda^* \right]. \end{aligned}$$

Here

$$\begin{aligned} \operatorname{tr} \left[\sum_{\lambda \in F_1 \setminus F_2} A_\lambda T A_\lambda^* \right] &= \operatorname{tr} \left[\sum_{\lambda \in F_0 \cup (F_1 \setminus F_2)} A_\lambda T A_\lambda^* - \sum_{\lambda \in F_0} A_\lambda T A_\lambda^* \right] \\ &= \operatorname{tr} \left[\sum_{\lambda \in F_0 \cup (F_1 \setminus F_2)} A_\lambda T A_\lambda^* - B T \right] - \operatorname{tr} \left[\sum_{\lambda \in F_0} A_\lambda T A_\lambda^* - B T \right] \\ &\leq \left| \operatorname{tr} \left[\sum_{\lambda \in F_0 \cup (F_1 \setminus F_2)} A_\lambda T A_\lambda^* - B T \right] \right| + \left| \operatorname{tr} \left[\sum_{\lambda \in F_0} A_\lambda T A_\lambda^* - B T \right] \right| \\ &< \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{\epsilon}{2}, \end{aligned}$$

and similarly $\operatorname{tr} \left[\sum_{\lambda \in F_2 \setminus F_1} A_\lambda T A_\lambda^* \right] < \frac{\epsilon}{2}$, so that

$$\left\| \sum_{\lambda \in F_1} A_\lambda T A_\lambda^* - \sum_{\lambda \in F_2} A_\lambda T A_\lambda^* \right\|_1 < \epsilon.$$

□

Proposition 7.4 *For a family $(A_\lambda)_{\lambda \in L}$ in $\mathcal{L}(\mathcal{K}, \mathcal{H})$ the following conditions are equivalent:*

- (i) *the sum $\sum_{\lambda \in L} A_\lambda^* A_\lambda$ exists in $\mathcal{L}(\mathcal{K})$ in the sense of weak convergence;*
- (ii) *for any $T \in \mathcal{T}(\mathcal{K})$ the sum $\sum_{\lambda} A_\lambda T A_\lambda^*$ exists in $\mathcal{T}(\mathcal{H})$ with respect to the trace norm;*
- (iii) *there is a trace norm bounded linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ such that $\Psi(T) = \sum_{\lambda} A_\lambda T A_\lambda^*$ for all $T \in \mathcal{T}(\mathcal{H})$ (convergence in the trace norm topology);*
- (iv) *there is a normal linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ such that $\Phi(S) = \sum_{\lambda \in L} A_\lambda^* S A_\lambda$ for all $S \in \mathcal{L}(\mathcal{H})$ (summation in the sense of weak convergence).*

If these conditions hold, then Φ is the transpose of Ψ .

Proof (i) $\implies$ (ii): See Lemma 7.4.

(ii) $\implies$ (iii): Assume (ii) and define $\Psi(T) = \sum_{\lambda \in L} A_\lambda T A_\lambda^*$ for all $T \in \mathcal{T}(\mathcal{K})$. Then Ψ is a positive linear map from $\mathcal{T}(\mathcal{K})$ into $\mathcal{T}(\mathcal{H})$, and so it is norm bounded by Proposition 6.2.

(iii) $\implies$ (iv): Assume (iii) and let $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ be the transpose of Ψ . Then Φ is a normal linear map. We show that it can be represented as claimed in (iv). Let $S \in \mathcal{L}(\mathcal{H})$. Take any $T \in \mathcal{T}(\mathcal{K})$ and $\epsilon > 0$. From (iii) it follows that there is a finite set $F_0 \subset L$ such that

$$\left| \operatorname{tr}[S\Psi(T)] - \operatorname{tr}\left[S \sum_{\lambda \in F} A_\lambda T A_\lambda^*\right] \right| < \epsilon$$

whenever $F \subset L$ is a finite set containing F_0 . Using Lemma 7.3 and the definition of Φ we see that this inequality implies

$$\left| \operatorname{tr}\left[\left(\Phi(S) - \sum_{\lambda \in F} A_\lambda^* A_\lambda\right)T\right] \right| < \epsilon.$$

Thus $\Phi(S) = \sum_{\lambda \in L} A_\lambda^* S A_\lambda$ where the convergence takes place even in the σ -weak topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$ of $\mathcal{L}(\mathcal{H})$.

(iv) $\implies$ (i): Choose $S = I$ in (iv). □

We need another auxiliary result.

Proposition 7.5 *Let $\pi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ be a normal $*$ -representation of $\mathcal{L}(\mathcal{H})$. Then there is a Hilbert space $\mathcal{K}_0$ such that π is unitarily equivalent to the direct sum $\sum_{i \in \mathcal{I} \cup \{0\}} \{\pi_i, \mathcal{K}_i\}$, where the index set $\mathcal{I}$ does not contain 0, $\pi_0(S) = 0$ for all $S \in \mathcal{L}(\mathcal{H})$, and each $\{\pi_i, \mathcal{K}_i\}$, with $i \in \mathcal{I}$, is the identity representation of $\mathcal{L}(\mathcal{H})$, i.e. $\mathcal{K}_i = \mathcal{H}$ and $\pi_i(S) = S$ for all $S \in \mathcal{L}(\mathcal{H})$. If π is nondegenerate, then $\sum_{i \in \mathcal{I} \cup \{0\}} \{\pi_i, \mathcal{K}_i\}$ may be replaced with $\sum_{i \in \mathcal{I}} \{\pi_i, \mathcal{K}_i\}$.*

Proof Retaining the notation, we first replace the representation with its restriction to $\mathcal{C}(\mathcal{H})$, the ideal of the compact operators on $\mathcal{H}$. The representation $\{\pi, \mathcal{K}\}$ can be expressed as a direct sum of zero and a nondegenerate representation (and the zero representation may be left out if π is already nondegenerate). Any nondegenerate representation is a direct sum of cyclic representations (Proposition 6.4). It is thus enough, in our first stage dealing with $\mathcal{C}(\mathcal{H})$, to show that any cyclic representation of $\mathcal{C}(\mathcal{H})$ is unitarily equivalent to some direct sum $\sum_{i \in \mathcal{I} \cup \{0\}} \{\pi_i, \mathcal{K}_i\}$, where each $\{\pi_i, \mathcal{K}_i\}$ is the identity representation of $\mathcal{C}(\mathcal{H})$. Now let $\{\pi', \mathcal{K}'\}$ be a representation of $\mathcal{C}(\mathcal{H})$ with a cyclic vector $\xi' \in \mathcal{K}'$. There is a $T \in \mathcal{T}(\mathcal{H})$ such that

$$\langle \xi' | \pi'(S) \xi' \rangle = \operatorname{tr}[ST],$$

for all $S \in \mathcal{C}(\mathcal{H})$ (see Theorem 3.12). In particular, for all $\varphi \in \mathcal{H}$, $\langle \varphi | T \varphi \rangle = \operatorname{tr}[T|\varphi\rangle\langle\varphi|] = \langle \xi' | \pi'(|\varphi\rangle\langle\varphi|) \xi' \rangle \geq 0$, and so $T \geq 0$. Thus there is a finite or countable orthonormal set L with a positive number λ_η for each $\eta \in L$, such that

$\sum_{\eta \in L} \lambda_\eta^2 < \infty$ and $T = \sum_{\eta \in L} \lambda_\eta^2 |\eta\rangle\langle\eta|$. For each $\eta \in L$, let $\{\pi_\eta, \mathcal{H}_\eta\}$ be the identity representation of $\mathcal{C}(\mathcal{H})$, that is, $\mathcal{H}_\eta = \mathcal{H}$ and $\pi_\eta(S) = S$ for all $S \in \mathcal{C}(\mathcal{H})$. Consider the direct sum

$$\{\tilde{\pi}, \tilde{\mathcal{K}}\} = \sum_{\eta \in L} \{\pi_\eta, \mathcal{H}_\eta\}.$$

We show that $\tilde{\eta} = (\lambda_\eta \eta)_{\eta \in L}$ is a cyclic vector for $\tilde{\pi}$. Let $\tilde{x} = (x_\eta)_{\eta \in L} \in \tilde{\mathcal{K}}$ and $\epsilon > 0$. Choose a finite set $F \subset L$ such that $\sum_{\eta \in L \setminus F} \|x_\eta\|^2 < \epsilon^2$. The finite rank operator $S_F = \sum_{u \in F} \lambda_u^{-1} |x_u\rangle\langle u|$ satisfies

$$S_F(\lambda_\eta \eta) = \sum_{u \in F} \lambda_u^{-1} \langle u | \lambda_\eta \eta \rangle x_u = x_\eta$$

for $\eta \in F$, and $S_F(\lambda_\eta \eta) = 0$ if $\eta \in L \setminus F$. Thus

$$\|\tilde{\pi}(S_F)\tilde{\eta} - \tilde{x}\| = \left(\sum_{\eta \in L \setminus F} \|x_\eta\|^2 \right)^{1/2} < \epsilon,$$

and so $\tilde{\eta}$ is a cyclic vector. Moreover, for any $S \in \mathcal{C}(\mathcal{H})$

$$\begin{aligned} \text{tr}[ST] &= \sum_{\eta \in L} \lambda_\eta^2 \text{tr}[|S\eta\rangle\langle\eta|] = \sum_{\eta \in L} \lambda_\eta^2 \langle\eta|S\eta\rangle \\ &= \sum_{\eta \in L} \langle\lambda_\eta \eta|S\lambda_\eta \eta\rangle = \langle\tilde{\eta}|\tilde{\pi}(S)\tilde{\eta}\rangle. \end{aligned}$$

Since $\langle\tilde{\eta}|\tilde{\pi}(S)\tilde{\eta}\rangle = \text{tr}[ST] = \langle\xi'|\pi'(S)\xi'\rangle$ for all $S \in \mathcal{C}(\mathcal{H})$, by Proposition 6.3 the representations π' and $\tilde{\pi}$ are unitarily equivalent.

Returning to our original situation, note that the direct sum of the identity representations of $\mathcal{C}(\mathcal{H})$ obtained above yields in an obvious way a corresponding direct sum of identity representations of $\mathcal{L}(\mathcal{H})$, which can be shown to be normal by using, e.g., Proposition 7.4. Since the original representation π of $\mathcal{L}(\mathcal{H})$ was normal, we get the claim since any two normal representations of $\mathcal{L}(\mathcal{H})$ coinciding on $\mathcal{C}(\mathcal{H})$ are the same (see Lemma 6.3). $\square$

The following theorem contains the *Kraus decomposition* of Φ .

Theorem 7.9 *For a map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ the following two conditions are equivalent:*

- (i) Φ is a normal completely positive linear map;
- (ii) there is a family $(A_\lambda)_{\lambda \in L}$ in $\mathcal{L}(\mathcal{K}, \mathcal{H})$ satisfying the four equivalent conditions of Proposition 7.4, and $\Phi(S) = \sum_{\lambda \in L} A_\lambda^* S A_\lambda$ for all $S \in \mathcal{L}(\mathcal{H})$ (summation in the sense of weak convergence).

Proof (i) $\implies$ (ii): By Theorem 7.5 there is a Hilbert space $\mathcal{K}'$ with a nondegenerate normal $*$ -representation $\varrho : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K}')$ and a bounded linear map $V : \mathcal{K} \rightarrow \mathcal{K}'$ such that $\Phi(S) = V^* \varrho(S) V$ for all $S \in \mathcal{L}(\mathcal{H})$. In accordance with Proposition 7.5, we identify $\mathcal{K}'$ with a direct sum $\sum_{\lambda \in L} \mathcal{H}_\lambda$ where $\mathcal{H}_\lambda = \mathcal{H}$ for all $\lambda \in L$ and $\varrho(S)(\eta_\lambda)_{\lambda \in L} = (S\eta_\lambda)_{\lambda \in L}$ whenever $S \in \mathcal{L}(\mathcal{H})$ and $(\eta_\lambda)_{\lambda \in L} \in \mathcal{K}'$. (To simplify notation we replace unitary equivalence with identification.) We regard each $\mathcal{H}_\lambda$ as a closed subspace of $\mathcal{K}'$, let P_λ be the projection of $\mathcal{K}'$ onto $\mathcal{H}_\lambda$, and denote $A_\lambda = P_\lambda V : \mathcal{K} \rightarrow \mathcal{H}_\lambda = \mathcal{H}$. If $\xi \in \mathcal{H}$ then for any finite set $F \subset L$ we have

$$\begin{aligned} \sum_{\lambda \in F} \langle \xi | A_\lambda^* A_\lambda \xi \rangle &= \sum_{\lambda \in F} \langle \xi | V^* P_\lambda V \xi \rangle \\ &= \sum_{\lambda \in F} \langle V \xi | P_\lambda V \xi \rangle \leq \langle V \xi | V \xi \rangle \leq \|V\|^2 \langle \xi | \xi \rangle. \end{aligned}$$

Moreover, for any $S \in \mathcal{L}(\mathcal{H})$ and $\xi, \eta \in \mathcal{K}$ we have

$$\begin{aligned} \langle \eta | \Phi(S) \xi \rangle &= \langle \eta | V^* \Phi(S) V \xi \rangle = \langle V \eta | \Phi(S) V \xi \rangle \\ &= \sum_{\lambda \in L} \langle P_\lambda V \eta | S P_\lambda V \xi \rangle = \sum_{\lambda \in L} \langle A_\lambda \eta | S A_\lambda \xi \rangle \\ &= \sum_{\lambda \in L} \langle \eta | A_\lambda^* S A_\lambda \xi \rangle \end{aligned}$$

implying that $V^* \Phi(S) V = \sum_{\lambda \in L} A_\lambda^* S A_\lambda$.

(ii) $\implies$ (i): Assume (ii). It follows from Proposition 7.4 that Φ is normal. To see that it is completely positive, we use the criterion in Lemma 7.1 (iii). For any $S_1, \dots, S_n \in \mathcal{L}(\mathcal{H})$ and $\xi_1, \dots, \xi_n \in \mathcal{K}$ we get

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | \Phi(S_i^* S_j) \xi_j \rangle &= \sum_{i=1}^n \sum_{j=1}^n \left\langle \xi_i \left| \left(\sum_{\lambda \in L} A_\lambda^* S_i^* S_j A_\lambda \right) \xi_j \right. \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \left\langle \xi_i \left| \left(\sum_{\lambda \in L} A_\lambda^* S_i^* S_j A_\lambda \right) \xi_j \right. \right\rangle \\ &= \sum_{\lambda \in L} \sum_{i=1}^n \sum_{j=1}^n \langle \xi_i | (A_\lambda^* S_i^* S_j A_\lambda) \xi_j \rangle = \sum_{\lambda \in L} \left\| \sum_{i=1}^n S_i A_\lambda \xi_i \right\|^2 \geq 0. \quad \square \end{aligned}$$

In the case of separable or finite-dimensional Hilbert spaces we can get a more specific conclusion.

Proposition 7.6 *If, in the situation of Theorem 7.9, the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ are separable or finite-dimensional, the statement of the theorem applies while the index set L may be taken to be countable or finite, respectively.*

Proof Suppose $\mathcal{H}$ and $\mathcal{K}$ are separable. Then it follows from Proposition 7.3 that in the proof of the above theorem the Hilbert space $\mathcal{K}'$ may be taken to be separable.

Since $\mathcal{K}'$ is isometrically isomorphic to the direct sum $\sum_{\lambda \in L} \mathcal{H}_\lambda$ where $\mathcal{H}_\lambda = \mathcal{H}$ for all $\lambda \in L$ (and we may assume that $\mathcal{H} \neq \{0\}$), L must be countable. The finite-dimensional situation is dealt with similarly. $\square$

7.5 Naimark Projections of Operator Integrals

Consider a semispectral measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ and let $(\mathcal{K}, F, V)$ be a dilation of it into a spectral measure acting on a Hilbert space $\mathcal{K}$. For any $\mathcal{A}$ -measurable function $f : \Omega \rightarrow \mathbb{C}$ the operator $L(f, F) = \int f dF$ is densely defined with the domain $\mathcal{D}(L(f, F)) = \mathcal{D}(f, F) = \tilde{\mathcal{D}}(f, F)$. The question arises under what conditions the ‘projected’ operator $V^*L(f, F)V$ equals the operator $L(f, E)$. This will be answered next [7].

Theorem 7.10 *With the above notations, the operator $V^*L(f, F)V$ is the restriction of the operator $L(f, E)$ to $\tilde{\mathcal{D}}(f, E)$. Thus, $V^*L(f, F)V = L(f, E)$ if and only if $\tilde{\mathcal{D}}(f, E) = \mathcal{D}(f, E)$.*

Proof For any $\varphi, \psi \in \mathcal{H}$ the complex measures $E_{\psi, \varphi}$ and $F_{V\psi, V\varphi}$ are the same. Hence, for any $\varphi \in \tilde{\mathcal{D}}(f, E)$,

$$\int |f|^2 dF_{V\varphi, V\varphi} = \int |f|^2 dE_{\varphi, \varphi} < \infty,$$

showing that $V(\tilde{\mathcal{D}}(f, E)) \subset \tilde{\mathcal{D}}(f, F) = \mathcal{D}(f, F)$. But then for any $\varphi \in \tilde{\mathcal{D}}(f, E)$ and for each $\psi \in \mathcal{H}$,

$$\begin{aligned} \langle \psi | L(f, E)\varphi \rangle &= \int f dE_{\psi, \varphi} = \int f dF_{V\psi, V\varphi} \\ &= \langle V\psi | L(f, F)V\varphi \rangle = \langle \psi | V^*L(f, F)V\varphi \rangle. \end{aligned}$$

This means that

$$L(f, E)|_{\tilde{\mathcal{D}}(f, E)} = V^*L(f, F)V|_{\tilde{\mathcal{D}}(f, E)}.$$

But we have

$$\begin{aligned} \mathcal{D}(V^*L(f, F)V) &= \mathcal{D}(L(f, F)V) = \{\varphi \in \mathcal{H} \mid V\varphi \in \mathcal{D}(L(f, F))\} \\ &= \left\{ \varphi \in \mathcal{H} \mid \int |f|^2 dF_{V\varphi, V\varphi} < \infty \right\} = \left\{ \varphi \in \mathcal{H} \mid \int |f|^2 dE_{\varphi, \varphi} < \infty \right\} \\ &= \tilde{\mathcal{D}}(f, E) \subset \mathcal{D}(f, E). \end{aligned}$$

Thus $V^*L(f, F)V$ is the restriction of $L(f, E)$ to $\tilde{\mathcal{D}}(f, E)$, and $V^*L(f, F)V = L(f, E)$ if and only if $\tilde{\mathcal{D}}(f, E) = \mathcal{D}(f, E)$. $\square$

7.6 Operations and Instruments

Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert spaces. According to Proposition 6.2, a positive linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ is bounded, with

$$\|\Psi\| = \sup\{\text{tr}[\Psi(T)] \mid T \in \mathcal{T}(\mathcal{K})_1^+\}.$$

By Proposition 6.1, its transpose (or dual) $\Psi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is a normal positive linear map. Moreover,

$$0 \leq \Psi^*(I_{\mathcal{H}}) \leq I_{\mathcal{K}} \quad (7.2)$$

if and only if

$$0 \leq \text{tr}[\Psi(T)] \leq 1 \quad (7.3)$$

for any $T \in \mathcal{T}(\mathcal{K})_1^+$. Conversely, if $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is a normal positive linear map, then there is a unique positive linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ such that $\Phi = \Psi^*$. We say that Ψ is the predual map of Φ and denote $\Phi_* = \Psi$.

Positive linear maps $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ and their duals $\Psi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ with the properties (7.3) and (7.2) have many important applications in quantum mechanics. Therefore, we introduce the following terminology.

Definition 7.2 A positive linear map $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ for which

$$0 \leq \text{tr}[\Psi(T)] \leq 1 \quad (7.4)$$

for all $T \in \mathcal{T}(\mathcal{K})_1^+$ is a (*Schrödinger*) *operation* or an *S-operation*. A normal positive linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ for which

$$0 \leq \Phi(I_{\mathcal{H}}) \leq I_{\mathcal{K}} \quad (7.5)$$

is a (*Heisenberg*) *operation* or an *H-operation* or a *dual operation*.

The transpose of an *S-operation* is an *H-operation* and any *H-operation* is the transpose of a unique *S-operation*. Due to this one-to-one correspondence between the *S-operations* and the *H-operations* we may occasionally simply call them *operations*.

The positive unit bounded operator $\Phi(I_{\mathcal{H}})$ is the *effect operator* of the (dual) operation Φ . Any effect operator $0 \leq A \leq I_{\mathcal{H}}$ is the effect operator of some operation $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$. Indeed, the map $\Phi_A : B \mapsto A^{1/2}BA^{1/2}$ is an operation for which $\Phi_A(I_{\mathcal{H}}) = A$. The correspondence between operations and effect operators is many-to-one. Calling two operations $\Phi_1 : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ and $\Phi_2 : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ equivalent if their effect operators are the same, that is, $\Phi_1(I_{\mathcal{H}}) = \Phi_2(I_{\mathcal{H}})$, the effect operators correspond bijectively to the equivalence classes of operations.

The notion of complete positivity applies directly to *H-operations*, and we use it to define the complete positivity of an *S-operation*.

Definition 7.3 An operation $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ is *completely positive* if its dual $\Psi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is completely positive.

Using Proposition 7.4, Theorem 7.9, and Proposition 7.6 we get the following result.

Proposition 7.7 Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces. An operation $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ is completely positive if and only if it can be represented in terms of a trace norm convergent series as

$$\Psi(T) = \sum_n A_n T A_n^*, \quad (7.6)$$

$T \in \mathcal{T}(\mathcal{K})$, for a countable set of bounded linear operators $A_n : \mathcal{K} \rightarrow \mathcal{H}$.

The effect operators stand in a bijective correspondence with the equivalence classes of operations. Similarly, normalised positive operator measures correspond to equivalence classes of operation valued measures. This leads us to the notion of an instrument:

Definition 7.4 Let $(\Omega, \mathcal{A})$ be a measurable space. A (*Heisenberg*) *instrument* (also called *H-instrument* or *dual instrument*) is a mapping $\mathfrak{M} : \mathcal{A} \times \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ defined through the properties

- (i) $\mathfrak{M}(X) : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$, $B \mapsto \mathfrak{M}(X)(B) = \mathfrak{M}(X, B)$, is linear, positive, and normal for all $X \in \mathcal{A}$;
- (ii) $\mathfrak{M}(\Omega, I_{\mathcal{H}}) = I_{\mathcal{K}}$;
- (iii) $\text{tr}[T \mathfrak{M}(\cup_{i=1}^{\infty} X_i, B)] = \sum_{i=1}^{\infty} \text{tr}[T \mathfrak{M}(X_i, B)]$ for any disjoint sequence $(X_i)_{i=1}^{\infty}$ in $\mathcal{A}$ and for all $T \in \mathcal{T}(\mathcal{K})$, $B \in \mathcal{L}(\mathcal{H})$.

For any *H-instrument* $\mathfrak{M}$ properties (i) and (ii) entail that $0 \leq \mathfrak{M}(X)(I_{\mathcal{H}}) \leq I_{\mathcal{K}}$, that is, $\mathfrak{M}(X)$ is an *H-operation* for all $X \in \mathcal{A}$. Hence, we may define an *S-operation* $\mathfrak{J}(X)$ as a predual map $\mathfrak{M}(X)_*$ and an *S-instrument* as a mapping $\mathfrak{M}_*$ from $\mathcal{A} \times \mathcal{T}(\mathcal{K})$ to $\mathcal{T}(\mathcal{H})$ by $\mathfrak{M}_*(X, T) = \mathfrak{J}(X)(T) = \mathfrak{M}(X)_*(T)$. We summarise this observation in the following definition.

Definition 7.5 Let $(\Omega, \mathcal{A})$ be a measurable space. A (*Schrödinger*) *instrument* or an (*S*-) *instrument* is an *S-operation* valued measure $\mathfrak{J} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{K}), \mathcal{T}(\mathcal{H}))$ defined through the properties

- (i) $\mathfrak{J}(X)$ is an *S-operation* for all $X \in \mathcal{A}$;
- (ii) $\text{tr}[\mathfrak{J}(\Omega)(T)] = \text{tr}[T]$ for all $T \in \mathcal{T}(\mathcal{K})$;
- (iii) $\mathfrak{J}(\cup_{i=1}^{\infty} X_i)(T) = \sum_{i=1}^{\infty} \mathfrak{J}(X_i)(T)$ for all disjoint sequences $(X_i)_{i=1}^{\infty}$ in $\mathcal{A}$ and all $T \in \mathcal{T}(\mathcal{K})$, with the series converging in the trace norm.

Remark 7.7 (a) Every *S-instrument* $\mathfrak{J}$ determines a unique dual instrument $\mathfrak{J}^*$ through the duality $\mathfrak{J}^*(X) = \mathfrak{J}(X)^*$, $X \in \mathcal{A}$, and each *H-instrument* can be obtained from a unique instrument in this way. Therefore, we may again omit the distinction

between instruments and dual instruments when there is no need to emphasise the formulation being used.

(b) Every instrument defines a unique semispectral measure $X \mapsto \mathbf{E}(X)$ via

$$\mathrm{tr}[\mathcal{T}\mathbf{E}(X)] = \mathrm{tr}[\mathcal{I}(X)(T)], \quad (7.7)$$

$X \in \mathcal{A}$, $T \in \mathcal{T}(\mathcal{K})$, or, equivalently, via

$$\mathbf{E}(X) = \mathcal{I}(X)^*(I_{\mathcal{H}}), \quad X \in \mathcal{A}.$$

We will call $\mathbf{E}$ the *associate semispectral measure (observable)* of $\mathcal{I}$.

(c) For every semispectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{K})$ there are infinitely many instruments having $\mathbf{E}$ as the associate semispectral measure. Such instruments will be called *$\mathbf{E}$ -compatible* instruments. As an example, let $(X_i)_{i \in I}$ be any countable partition of Ω into disjoint ($\mathcal{A}$ -measurable) sets and $(T_i)_{i \in I} \subset \mathcal{T}(\mathcal{H})$ a collection of positive operators of trace 1. Then the following defines an $\mathbf{E}$ -compatible instrument $\mathcal{I} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{K}), \mathcal{T}(\mathcal{H}))$:

$$\mathcal{I}(X)(T) = \sum_{i \in I} \mathrm{tr}[\mathbf{E}(X \cap X_i)]T_i, \quad (7.8)$$

where $X \in \mathcal{F}$ and $T \in \mathcal{T}(\mathcal{K})$. In Theorem 7.12 below we determine all completely positive $\mathbf{E}$ -compatible instruments. $\triangleleft$

Completely Positive Instruments

The notion of complete positivity can be applied to instruments, as follows. An instrument $\mathfrak{M} : \mathcal{A} \times \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is *completely positive* if all the operations $\mathfrak{M}(X)$, $X \in \mathcal{A}$, in its range are completely positive. The set of *completely positive* instruments $\mathfrak{M} : \mathcal{A} \times \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$, denoted by $\mathrm{Ins}(\mathcal{A}, \mathcal{H}, \mathcal{K})$, is convex, and the set of its extreme points will be denoted by $\mathrm{ex} \, \mathrm{Ins}(\mathcal{A}, \mathcal{H}, \mathcal{K})$. We formulate next a minimal Stinespring type representation for completely positive instruments.

Theorem 7.11 *Let $\mathfrak{M} \in \mathrm{Ins}(\mathcal{A}, \mathcal{H}, \mathcal{K})$. There is a Hilbert space $\mathcal{H}_1$, a spectral measure $\mathbf{P}_1 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_1)$, and a linear isometry $V : \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{H}_1$ such that*

$$\mathfrak{M}(X, B) = V^*(B \otimes \mathbf{P}_1(X))V$$

for all $X \in \mathcal{A}$, $B \in \mathcal{L}(\mathcal{H})$, and the linear span of the vectors $(B \otimes \mathbf{P}_1(X))V\xi$, $B \in \mathcal{L}(\mathcal{H})$, $X \in \mathcal{A}$, $\xi \in \mathcal{K}$ is dense in $\mathcal{H} \otimes \mathcal{H}_1$.

Proof By Theorems 7.2 and 7.3, there is a Hilbert space $\mathcal{H}_2$ with a spectral measure $\mathbf{P}_2 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_2)$, a normal *-representation $\pi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}_2)$, and a linear isometry $V : \mathcal{K} \rightarrow \mathcal{H}_2$ satisfying

$$\begin{aligned}\mathfrak{M}(X, B) &= V^* P_2(X) \pi(B) V, \\ P_2(X) \pi(B) &= \pi(B) P_2(X)\end{aligned}$$

for all $X \in \mathcal{A}$ and $B \in \mathcal{L}(\mathcal{H})$. In addition, the linear span of the vectors $P_2(X) \pi(B) V \xi$, $X \in \mathcal{A}$, $B \in \mathcal{L}(\mathcal{H})$, $\xi \in \mathcal{K}$, is dense in $\mathcal{H}_2$.

By Proposition 7.5, π is unitarily equivalent to the direct sum of identity representations. Therefore, using Remark 2.6 we see that there is a Hilbert space $\mathcal{H}_1$ such that $\mathcal{H}_2 = \mathcal{H} \otimes \mathcal{H}_1$ and $\pi(B) = B \otimes I_{\mathcal{H}_1}$ for all $B \in \mathcal{L}(\mathcal{H})$. Since $P_2(X) \pi(B) = \pi(B) P_2(X)$ for all $X \in \mathcal{A}$ and $B \in \mathcal{L}(\mathcal{H})$, there is a spectral measure $P_1 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_1)$ such that $P_2(X) = I_{\mathcal{H}} \otimes P_1(X)$ (exercise). Then for any $X \in \mathcal{A}$, $B \in \mathcal{L}(\mathcal{H})$, one gets $\mathfrak{M}(X, B) = V^*(B \otimes P_1(X))V$. The last density claim is obvious. $\square$

By using the above theorem, it is easy to characterise all *completely positive* E -compatible instruments.

Definition 7.6 A linear map $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is a *channel* if it is normal, completely positive, and unital (i.e. $\Phi(I_{\mathcal{H}}) = I_{\mathcal{K}}$).

Let $\mathcal{H}'$ be a Hilbert space, $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ a semispectral measure with a minimal Naimark dilation $(\mathcal{K}, F, V)$, and $\mathfrak{M} \in \text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$. The next result is from [8].

Theorem 7.12 $\mathfrak{M}(X, I_{\mathcal{H}'}) = E(X)$ for all $X \in \mathcal{A}$ if and only if there exists a (unique) channel $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{K})$ such that, for all $B \in \mathcal{L}(\mathcal{H}')$, $X \in \mathcal{A}$, $\Phi(B)F(X) = F(X)\Phi(B)$ and $\mathfrak{M}(X, B) = V^*F(X)\Phi(B)V$.

Proof Assume that $\mathfrak{M}(X, I_{\mathcal{H}'}) = E(X)$, $X \in \mathcal{A}$, and let $(\mathcal{H}_1, P_1, V')$ be a minimal Stinespring type representation for $\mathfrak{M}$ as in Theorem 7.11 so that

$$\mathfrak{M}(X, B) = V'^*(B \otimes P_1(X))V', \quad X \in \mathcal{A}, B \in \mathcal{L}(\mathcal{H}'),$$

where $P_1 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_1)$ is a spectral measure and $V' : \mathcal{H} \rightarrow \mathcal{H}' \otimes \mathcal{H}_1$ is an isometry. Thus,

$$E(X) = V^*F(X)V = V'^*(I_{\mathcal{H}'} \otimes P_1(X))V' = \mathfrak{M}(X, I_{\mathcal{H}'}), \quad X \in \mathcal{A},$$

implying that $(\mathcal{H}' \otimes \mathcal{H}_1, I_{\mathcal{H}'} \otimes P_1, V')$ is a (not necessarily minimal) Naimark representation for E . Hence, there exists a linear isometry $C : \mathcal{K} \rightarrow \mathcal{H}' \otimes \mathcal{H}_1$ such that, for all $X \in \mathcal{A}$,

$$CF(X)V = (I_{\mathcal{H}'} \otimes P_1(X))V'$$

and $CF(X) = (I_{\mathcal{H}'} \otimes P_1(X))C$. Define a channel $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{K})$ by

$$\Phi(B) = C^*(B \otimes I_{\mathcal{H}_1})C, \quad B \in \mathcal{L}(\mathcal{H}').$$

It is easy to see that, for all $B \in \mathcal{L}(\mathcal{H}')$, $X \in \mathcal{A}$, $\Phi(B)F(X) = F(X)\Phi(B)$ and $\mathfrak{M}(X, B) = V^*F(X)\Phi(B)V$. The converse claim is trivial. Note that Φ is unique due to the minimality of the Naimark dilation $(\mathcal{K}, F, V)$. $\square$

Remark 7.8 If $\mathbf{E}$ is projection valued in the context of the preceding theorem, then one may choose $\mathcal{K} = \mathcal{H}$, $\mathbf{F} = \mathbf{E}$, and $V = I_{\mathcal{H}}$ so that

$$\mathfrak{M}(X, B) = \mathbf{E}(X)\Phi(B) = \mathbf{E}(X)\Phi(B)\mathbf{E}(X)$$

where $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H})$ is a unique channel commuting with $\mathbf{E}$. This special case (with $\mathcal{H}' = \mathcal{H}$) was first obtained in [9]. Note that Φ is determined by $\Phi(X) = \mathfrak{M}(X, I_{\mathcal{H}'})$. One can view $\mathfrak{M}$ as a combination of Φ and Lüders operations $A \mapsto \mathbf{E}(X)A\mathbf{E}(X)$. $\triangleleft$

Extreme Completely Positive Instruments

Let $\mathfrak{M} \in \text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ and let $(\mathcal{H}_1, \mathbf{P}_1, V)$ be its minimal Stinespring type representation as in Theorem 7.11, that is, $\mathcal{H}_1$ is a Hilbert space, $\mathbf{P}_1 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_1)$ a spectral measure, and $V : \mathcal{H} \rightarrow \mathcal{H}' \otimes \mathcal{H}_1$ an isometry such that

$$\mathfrak{M}(X, B) = V^*(B \otimes \mathbf{P}_1(X))V, \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}'),$$

and the linear space

$$\mathcal{V} = \text{lin}\{(B \otimes \mathbf{P}_1(X))V\xi \mid B \in \mathcal{L}(\mathcal{H}'), X \in \mathcal{A}, \xi \in \mathcal{H}\}$$

is dense in $\mathcal{H}' \otimes \mathcal{H}_1$. The next theorem characterises the extreme points of the convex set $\text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$.

Theorem 7.13 ([10]) $\mathfrak{M} \in \text{ex Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ if and only if, for any $D \in \mathcal{L}(\mathcal{H}_1)$ such that $D\mathbf{P}_1(X) = \mathbf{P}_1(X)D$ for all $X \in \mathcal{A}$, the condition

$$V^*(I_{\mathcal{H}'} \otimes D)V = 0 \tag{7.9}$$

implies $D = 0$.

Proof Assume that there exists a nonzero $D \in \mathcal{L}(\mathcal{H}_1)$ such that $D\mathbf{P}_1(X) = \mathbf{P}_1(X)D$ for all $X \in \mathcal{A}$ and (7.9) holds. We may assume that $D^* = D$ and $\|D\| \leq 1$ since we may redefine D to be $(2i\|D\|)^{-1}(D - D^*)$ (if $D \neq D^*$). Now one can define positive operators $D_{\pm} = I_{\mathcal{H}_1} \pm D$ and maps $\mathfrak{M}_{\pm} : \mathcal{A} \times \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H})$ by

$$\mathfrak{M}_{\pm}(X, B) = V^*(B \otimes D_{\pm}\mathbf{P}_1(X))V, \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}').$$

It is left as an exercise to show that $\mathfrak{M}_{\pm} \in \text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ and $\mathfrak{M} = \frac{1}{2}(\mathfrak{M}_+ + \mathfrak{M}_-)$. Since $D_+ \neq D_-$ and $\mathcal{V}$ is dense in $\mathcal{H}' \otimes \mathcal{H}_1$ one sees that $\mathfrak{M}_+ \neq \mathfrak{M}_-$ and $\mathfrak{M}$ is not an extreme point.

Suppose then that $\mathfrak{M}$ is not extreme, that is, there exist instruments $\mathfrak{M}_{\pm} \in \text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ such that $\mathfrak{M}_+ \neq \mathfrak{M}_-$ and $\mathfrak{M} = \frac{1}{2}(\mathfrak{M}_+ + \mathfrak{M}_-)$. Let $(\mathcal{H}_{\pm}, \mathbf{P}_{\pm}, V_{\pm})$ be minimal Stinespring type representations for $\mathfrak{M}_{\pm}$ so that, for all $X \in \mathcal{A}$, $B \in \mathcal{L}(\mathcal{H}')$, $\mathfrak{M}_{\pm}(X, B^*B) \leq 2\mathfrak{M}(X, B^*B)$ or

$$[(B \otimes P_{\pm}(X))V_{\pm}]^*(B \otimes P_{\pm}(X))V_{\pm} \leq 2[(B \otimes P_1(X))V]^*(B \otimes P_1(X))V.$$

It is an easy exercise to show that one gets (well-defined) bounded linear operators $G_{\pm} : \mathcal{H}' \otimes \mathcal{H}_1 \rightarrow \mathcal{H}' \otimes \mathcal{H}_{\pm}$ by setting

$$G_{\pm}(B \otimes P_1(X))V = (B \otimes P_{\pm}(X))V_{\pm}, \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}').$$

Define then positive operators $F_{\pm} = G_{\pm}^* G_{\pm}$. Clearly,

$$F_{\pm}(B \otimes P_1(X)) = (B \otimes P_1(X))F_{\pm}, \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}'),$$

so that $F_{\pm}$ must be of the form $F_{\pm} = I_{\mathcal{H}'} \otimes D_{\pm}$ where $D_{\pm} \in \mathcal{L}(\mathcal{H}_1)$ and $D_{\pm}P_1(X) = P_1(X)D_{\pm}$ for all $X \in \mathcal{A}$. Since

$$\mathfrak{M}_{\pm}(X, B) = V^*(B \otimes D_{\pm}P_1(X))V, \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}'),$$

and $\mathfrak{M}_+ \neq \mathfrak{M}_-$ we have $D = D_+ - D_- \neq 0$ and $DP_1(X) = P_1(X)D$ for all $X \in \mathcal{A}$. From $\mathfrak{M}_+(\Omega, I_{\mathcal{H}'} - \mathfrak{M}_-(\Omega, I_{\mathcal{H}'})) = I_{\mathcal{H}'} - I_{\mathcal{H}'} = 0$ one sees that (7.9) holds and the proof is complete. $\square$

Remark 7.9 Let $\mathcal{A} = \mathcal{B}(\Omega)$, the Borel σ -algebra of a second countable Hausdorff space Ω , and assume that the dimension of $\mathcal{H}$ is finite. Then it can be shown [10] that if $\mathfrak{M} \in \text{ex Ins}(\mathcal{B}(\Omega), \mathcal{H}', \mathcal{H})$ then $\mathfrak{M}$ is concentrated on a finite set $\{\omega_i\}_{i=1}^N \subset \Omega$ where $\omega_i \neq \omega_j$, $i \neq j$, and $N \leq (\dim \mathcal{H})^2$, that is, $\mathfrak{M}$ is of the form

$$\mathfrak{M}(X, B) = \sum_{i=1}^N \chi_X(\omega_i) T_i(B), \quad X \in \mathcal{A}, \quad B \in \mathcal{L}(\mathcal{H}'),$$

where the linear maps $T_i : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H})$ are completely positive (dual) operations. Moreover, $\sum_{i=1}^N T_i(I_{\mathcal{H}'})) = I_{\mathcal{H}}$ and, if $T_i(B) = \sum_{k=1}^{n(i)} A_k(i)^* B A_k(i)$ is a minimal Kraus decomposition of T_i , then $\sum_{i=1}^N \sum_{k,l=1}^{n(i)} D(i)_{kl} A_k(i)^* A_l(i) = 0$ implies $D(i)_{kl} \equiv 0$ where $(D(i)_{kl})_{k,l=1}^{n(i)}$ is a complex matrix (with the finite operator norm if $n(i) = \infty$). $\triangleleft$

Example 7.1 Choose $\Omega = \{0\}$ and $\mathcal{A} = \{\emptyset, \{0\}\}$. Then $\text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ can be identified with the convex set of channels $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H})$. Let $(\mathcal{H}_1, P_1, V)$, $P_1(\{0\}) = I_{\mathcal{H}_1}$, be a minimal Stinespring type representation for Φ . Especially, $\Phi(B) = V^*(B \otimes I_{\mathcal{H}_1})V$, $B \in \mathcal{L}(\mathcal{H}')$. Then, by Theorem 7.13, Φ is extreme if and only if $V^*(I_{\mathcal{H}'} \otimes D)V = 0$ implies $D = 0$ (where $D \in \mathcal{L}(\mathcal{H}_1)$). As in Remark 7.9, one can rephrase this extremality condition in terms of minimal Kraus operators $A_k(1)$ of Φ (put $N = 1$ and drop the sum symbol $\sum_{i=1}^N$).

Suppose further that $\mathcal{H} = \mathbb{C}$ so that $\mathcal{L}(\mathbb{C}) \cong \mathbb{C}$. Then any channel $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathbb{C}$ can be identified with a positive trace-1 operator (i.e. a state) $\varrho \in \mathcal{S}(\mathcal{H}') = \{\varrho \in \mathcal{T}(\mathcal{H}') \mid \varrho \geq 0, \text{tr}[\varrho] = 1\}$ by $\Phi(B) = \text{tr}[\varrho B]$, $B \in \mathcal{L}(\mathcal{H}')$. Now the minimal dilation

of Φ or ϱ is just $\Phi(B) = V^*(B \otimes I_{\mathcal{H}_1})V$ where $V : \mathbb{C} \rightarrow \mathcal{H}' \otimes \mathcal{H}_1$ can be identified with the unit vector $\Psi = V1 \in \mathcal{H}' \otimes \mathcal{H}_1$. Hence, we have

$$\mathrm{tr}[\varrho B] = \Phi(B) = \langle \Psi | (B \otimes I_{\mathcal{H}_1}) \Psi \rangle = \mathrm{tr}[|\Psi\rangle\langle\Psi| (B \otimes I_{\mathcal{H}_1})]$$

or $\varrho = \mathrm{tr}_{\mathcal{H}_1} |\Psi\rangle\langle\Psi|$ (the partial state over $\mathcal{H}_1$). In the context of quantum information theory, Ψ is called a *purification of the state* ϱ . Now ϱ is an extreme point in the convex set of states if and only if $\langle \Psi | (I_{\mathcal{H}'} \otimes D) \Psi \rangle$ implies $D = 0$ (where $D \in \mathcal{L}(\mathcal{H}_1)$). Clearly, this can happen if and only if the dimension of $\mathcal{H}_1$ is 1. But then ϱ can be identified with its purification $|\Psi\rangle\langle\Psi|$, that is, the *extreme states are the pure states*.

Example 7.2 Assume that $\mathcal{H}' = \mathbb{C}$ and $\mathcal{L}(\mathcal{H}') = \mathbb{C}$. Then $\mathrm{Ins}(\mathcal{A}, \mathbb{C}, \mathcal{H})$ is the convex set of semispectral measures $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$. Let $(\mathcal{H}_1, \mathbf{P}_1, V)$ be a minimal Stinespring type (or Naimark) representation for $\mathbf{E}$ so that

$$\mathbf{E}(X) = V^* \mathbf{P}_1(X) V$$

(where we have identified $\mathbb{C} \otimes \mathcal{H}_1$ with $\mathcal{H}_1$). From Theorem 7.13 follows that $\mathbf{E}$ is an extreme semispectral measure if and only if, for any $D \in \mathcal{L}(\mathcal{H}_1)$ such that $D\mathbf{P}_1(X) = \mathbf{P}_1(X)D$, $X \in \mathcal{A}$, the condition $V^*DV = 0$ implies $D = 0$. Especially, if $\dim \mathcal{H} < \infty$ (and $\mathcal{A} = \mathcal{B}(\Omega)$ as in Remark 7.9), then any extreme semispectral measure is concentrated on a finite set.

If we also set $\mathcal{H} = \mathbb{C}$, we see that $\mathrm{Ins}(\mathcal{A}, \mathbb{C}, \mathbb{C})$ is the convex set of probability measures $\mu : \mathcal{A} \rightarrow [0, 1]$ which can be identified with the *classical states*. Similarly as in the preceding example, we see that the minimal dilation of μ is of the form

$$\mu(X) = \langle \Psi | \mathbf{P}_1(X) \Psi \rangle$$

where $\Psi \in \mathcal{H}_1$ is a unit vector and $\mathbf{P}_1 : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_1)$ is a spectral measure. Now μ is an extreme probability measure if and only if $\langle \Psi | D \Psi \rangle = 0$ implies $D = 0$ (where D commutes with $\mathbf{P}_1$). Obviously, this can hold if and only if the dimension of $\mathcal{H}_1$ is 1. But this implies that μ can be identified with the spectral measure $\mathbf{P}_1$, that is, $\mu(X)^2 = \mu(X)$ for all $X \in \mathcal{A}$. In conclusion, μ is an extreme probability measure (extreme classical state) if and only if $\mu(X) \in \{0, 1\}$ for all $X \in \mathcal{A}$. Especially, if $\mathcal{A} = \mathcal{B}(\Omega)$ as in Remark 7.9, then μ is extreme if and only if μ is a Dirac (point) measure. $\triangleleft$

7.7 Measurement Dilation

Let $\mathcal{H}$ and $\mathcal{H}_o$ be two complex separable Hilbert spaces. Fix a unit vector $\phi \in \mathcal{H}_o$, a spectral measure $\mathbf{P} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_o)$, and a unitary operator $U \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H}_o)$. The quadruple $(\mathcal{H}_o, \phi, \mathbf{P}, U)$ defines a completely positive instrument. Indeed, for each $X \in \mathcal{A}$ the map

$$\begin{aligned} T &\mapsto T \otimes P[\phi] \mapsto U(T \otimes P[\phi])U^* \mapsto (I \otimes \mathbf{P}(X))U(T \otimes P[\phi])U^*(I \otimes \mathbf{P}(X)) \\ &\mapsto \text{tr}_{\mathcal{H}_o}[U(T \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))] \end{aligned}$$

defines a completely positive operation, and the map $X \mapsto \mathfrak{J}(X)$ for which

$$\mathfrak{J}(X)(T) = \text{tr}_{\mathcal{H}_o}[U(T \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))]$$

for all $T \in \mathcal{T}(\mathcal{H})$ constitutes a completely positive instrument, that is, $\mathfrak{J}^* \in \text{Ins}(\mathcal{A}, \mathcal{H}, \mathcal{H})$.

On the basis of Theorem 7.11 we will show that each completely positive instrument $\mathfrak{J}^* \in \text{Ins}(\mathcal{A}, \mathcal{H}, \mathcal{H})$ has such a realisation.

Theorem 7.14 *Let $(\Omega, \mathcal{A})$ be a measurable space. For any completely positive instrument $\mathfrak{J} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{H}))$ there exist a Hilbert space $\mathcal{H}_o$, a unit vector $\phi \in \mathcal{H}_o$, a unitary operator $U : \mathcal{H} \otimes \mathcal{H}_o \rightarrow \mathcal{H} \otimes \mathcal{H}_o$, and a spectral measure $\mathbf{P} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_o)$ such that for all $X \in \mathcal{A}$, $T \in \mathcal{T}(\mathcal{H})$,*

$$\mathfrak{J}(X)(T) = \text{tr}_{\mathcal{H}_o}[U(T \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))].$$

Proof Let $(\mathcal{H}_1, \mathbf{P}_1, V)$ be as in Theorem 7.11 for the dual instrument $\mathfrak{M}(X, B) = \mathfrak{J}(X)^*(B)$. Note that now $\mathcal{K} = \mathcal{H}$. Let $\eta \in \mathcal{H}_2 = \mathcal{H} \otimes \mathcal{H}_1$ and $\xi \in \mathcal{H}_1$ be any two unit vectors and denote by $[\eta]$ and $[\xi]$ the one-dimensional subspaces they generate. Define a linear isometry $V_o : \mathcal{H} \otimes [\xi] \otimes [\eta] \rightarrow \mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$V_o(\varphi \otimes \xi \otimes \eta) = (V\varphi) \otimes \eta$$

for all $\varphi \in \mathcal{H}$. Since $\mathcal{H}_2 = \mathcal{H} \otimes \mathcal{H}_1$ and V is a linear isometry $\mathcal{H} \rightarrow \mathcal{H}_2$, the Hilbert codimensions of $\mathcal{H} \otimes [\xi] \otimes [\eta]$ and $V_o(\mathcal{H} \otimes [\xi] \otimes [\eta])$ are the same. (Indeed, if $\mathcal{H}$ is finite-dimensional, then $\mathcal{H} \otimes [\xi]$ and $V(\mathcal{H})$ have the same Hilbert codimensions in $\mathcal{H} \otimes \mathcal{H}_1 = \mathcal{H}_2$. If $\mathcal{H}$ is infinite-dimensional, then so is $\mathcal{H}_2$ and the Hilbert codimensions will be the same as the Hilbert dimension of $\mathcal{H}_2$.) Therefore, V_o can be extended to a unitary operator U on $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $\mathcal{H}_o = \mathcal{H}_1 \otimes \mathcal{H}_2$, $\phi = \eta \otimes \xi$ and let $\mathbf{P} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_o)$ be the spectral measure defined by $\mathbf{P}(X) = \mathbf{P}_1(X) \otimes I_{\mathcal{H}_2}$. Then for any $\varphi \in \mathcal{H}$, $B \in \mathcal{L}(\mathcal{H})$ and $X \in \mathcal{A}$,

$$\begin{aligned} \text{tr}[\mathfrak{J}(X)^*(B)P[\varphi]] &= \langle \mathfrak{J}(X)^*(B)\varphi | \varphi \rangle \\ &= \langle V^*(B \otimes \mathbf{P}_1(X))V\varphi | \varphi \rangle \\ &= \langle (B \otimes \mathbf{P}_1(X))V\varphi | V\varphi \rangle \\ &= \langle (B \otimes \mathbf{P}_1(X))V\varphi \otimes \eta | V\varphi \otimes \eta \rangle \\ &= \langle (B \otimes \mathbf{P}_1(X) \otimes I)U(\varphi \otimes \xi \otimes \eta) | U(\varphi \otimes \xi \otimes \eta) \rangle \\ &= \langle (B \otimes \mathbf{P}(X))U(\varphi \otimes \phi) | U(\varphi \otimes \phi) \rangle \\ &= \text{tr}[B \otimes \mathbf{P}(X)U(P[\varphi] \otimes P[\phi])U^*] \end{aligned}$$

$$\begin{aligned}
&= \operatorname{tr}[B \operatorname{tr}_{\mathcal{H}_o}[U(P[\varphi] \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))]] \\
&= \operatorname{tr}[B\mathfrak{I}(X)(P[\varphi])],
\end{aligned}$$

which shows that

$$\mathfrak{I}(X)(P[\varphi]) = \operatorname{tr}_{\mathcal{H}_o}[U(P[\varphi] \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))]$$

for all $X \in \mathcal{A}$ and for any unit vector $\varphi \in \mathcal{H}$. $\square$

To close this section, we observe that for each semispectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ there exist completely positive instruments $\mathfrak{I}$ such that $\mathbf{E}(X) = \mathfrak{I}(X)^*(I)$ for all $X \in \mathcal{A}$. For instance, the instrument introduced in Remark 7.7 (b) is completely positive; for simplicity, we set $\mathcal{K} = \mathcal{H}$. Let $(\varphi_j)_{j=1}^{\dim \mathcal{H}}$ be an orthonormal basis of $\mathcal{H}$ and put

$$A_{ijk} = T_i^{1/2} |\varphi_j\rangle \langle \varphi_k| \mathbf{E}(X \cap X_i)^{1/2}.$$

Then the operation (7.8) can be written in the Kraus form

$$\mathfrak{I}(X)(T) = \sum_{i \in I} \operatorname{tr}[T \mathbf{E}(X \cap X_i)] T_i = \sum_{i,j,k} A_{ijk} T A_{ijk}^*,$$

showing the complete positivity of the instrument $X \mapsto \mathfrak{I}(X)$.

Corollary 7.1 *For any semispectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ there is a quadruple $(\mathcal{H}_o, \phi, \mathbf{P}, U)$ such that $\mathbf{E}(X) = \mathfrak{I}(X)^*(I)$, with*

$$\mathfrak{I}(X)(T) = \operatorname{tr}_{\mathcal{H}_o}[U(T \otimes P[\phi])U^*(I \otimes \mathbf{P}(X))].$$

We call a quadruple $(\mathcal{H}_o, \phi, \mathbf{P}, U)$ with the properties mentioned in Corollary 7.1 a *measurement dilation* of $\mathbf{E}$. If $V_\phi : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}_o$ denotes the embedding $V_\phi(\varphi) = \varphi \otimes \phi$, then

$$\mathbf{E}(X) = V_\phi^* U^* (I \otimes \mathbf{P}(X)) U V_\phi$$

for all $X \in \mathcal{A}$, and the triple $(\mathcal{H} \otimes \mathcal{H}_o, I \otimes \mathbf{P}, U V_\phi)$ constitutes a Naimark dilation of $\mathbf{E}$ into a spectral measure $X \mapsto I \otimes \mathbf{P}(X)$. Usually, the dilation is not minimal.

7.8 Exercises

1. Show that the formula for $\langle u | v \rangle_0$ in the proof of Theorem 7.1 yields a well-defined sesquilinear form on $K_0 \times K_0$. (Hint: Sesquilinearity may be reduced to bilinearity by redefining the scalar multiplication in one of the spaces.)
2. Show that the space $\mathcal{C}(\mathcal{H})$ of the compact operators on a separable Hilbert space $\mathcal{H}$ is separable.

3. Let $\Psi : \mathcal{T}(\mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H})$ be a positive linear map and $\Psi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ its transpose. Show that

$$\Psi^*(I_{\mathcal{H}}) = I_{\mathcal{K}} \quad (7.10)$$

if and only if

$$\mathrm{tr}[\Psi(T)] = 1 \quad (7.11)$$

for any $T \in \mathcal{T}(\mathcal{K})_1^+$. Show that the last condition implies that $\mathrm{tr}[\Psi(T)] = \mathrm{tr}[T]$ for all $T \in \mathcal{T}(\mathcal{K})$.

4. Let $A \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $C \in \mathcal{L}(\mathcal{H}_2)$ be such that $A(B \otimes C) = (B \otimes C)A$ for all $B \in \mathcal{L}(\mathcal{H}_1)$. Show that $A = I_{\mathcal{H}_1} \otimes D$ for some $D \in \mathcal{L}(\mathcal{H}_2)$ for which $CD = DC$ holds.
5. Complete the proof of Theorem 7.11.
6. Complete the proof of Theorem 7.13.
7. Determine the extreme points of the convex set $\mathrm{Ins}(\mathcal{A}, \mathcal{H}', \mathbb{C})$.
8. Let Ω be a second countable Hausdorff space. Show that a classical state $\mu : \mathcal{B}(\Omega) \rightarrow [0, 1]$ is extreme if and only if $\mu = \delta_\omega$, the point measure at $\omega \in \Omega$. Show that any convex combination of the extreme (or pure) classical states to a classical state is unique and compare this situation to quantum mechanics.
9. Show that any $\mathfrak{M} \in \mathrm{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ has a unique predual map $\mathfrak{J} = \mathfrak{M}_*$, i.e. the dual map $\mathfrak{J}^* = \mathfrak{M}$, and that the predual maps also form a convex set. Show that $\mathfrak{M}$ is extreme if and only if $\mathfrak{M}_*$ is extreme. What is the physical interpretation of the predual map of a channel?

References

1. Stinespring, W.F.: Positive functions on C^* -algebras. *Proc. Am. Math. Soc.* **6**, 211–216 (1955)
2. Ylinen, K.: Positive operator bimeasures and a noncommutative generalization. *Studia Math.* **118**(2), 157–168 (1996)
3. Takesaki, M.: *Theory of operator algebras. I.* *Encyclopaedia of Mathematical Sciences*, vol. 124. Springer, Berlin (2002) (Reprint of the first (1979) edition, *Operator Algebras and Noncommutative Geometry*, 5)
4. Schaefer, H.H.: *Topological Vector Spaces. Graduate Texts in Mathematics*, vol. 3. Springer, New York (1971) (Corrected third printing)
5. Halmos, P.R.: *Measure Theory*, 4th edn. Springer, New York (1988) (Reprint of the 1950 edition)
6. Dunford, N., Schwartz, J.T.: *Linear Operators. Part I.* *Wiley Classics Library*. Wiley, New York (1988) (Reprint of the 1958 original, *A Wiley-Interscience Publication*)
7. Lahti, P., Pellonpää, J.-P., Ylinen, K.: Operator integrals and phase space observables. *J. Math. Phys.* **40**(4), 2181–2189 (1999)
8. Pellonpää, J.P.: Quantum instruments: II. Measurement theory. *J. Phys. A* **46**(2), 025303, 15 (2013)
9. Ozawa, M.: Quantum measuring processes of continuous observables. *J. Math. Phys.* **25**(1), 79–87 (1984)
10. Pellonpää, J.P.: Quantum instruments: I. Extreme instruments. *J. Phys. A* **46**(2), 025302, 16 (2013)

Chapter 8

Positive Operator Measures: Examples

According to the Naimark dilation theorem (Theorem 7.7), any semispectral measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ can be obtained from a spectral measure $F : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{K})$ through a suitable isometric embedding of $\mathcal{H}$ into a larger Hilbert space $\mathcal{K}$. In this chapter we discuss some common examples of semispectral measures and their natural dilations.

8.1 The Canonical Spectral Measure and Its Fourier-Plancherel Transform

In this section $\mathcal{H} = L^2(\mathbb{R})$. For any (Borel) measurable $f : \mathbb{R} \rightarrow \mathbb{C}$ we define the multiplication operator M_f by the formula

$$(M_f \varphi)(x) = (f\varphi)(x) = f(x)\varphi(x),$$

$x \in \mathbb{R}$, $\varphi \in \mathcal{D}(M_f) = \{\varphi \in L^2(\mathbb{R}) \mid f\varphi \in L^2(\mathbb{R})\}$. In particular, the characteristic function χ_X of the set $X \in \mathcal{B}(\mathbb{R})$ defines a projection operator M_{χ_X} and the assignment $X \mapsto M_{\chi_X}$ constitutes a spectral measure, the *canonical spectral measure* in $L^2(\mathbb{R})$. We denote it by $\mathbf{M}$, that is, $\mathbf{M}(X) = M_{\chi_X}$ for each $X \in \mathcal{B}(\mathbb{R})$.

For any measurable function f the multiplication operator M_f is just the operator integral $L(f, \mathbf{M})$, with the dense domain $\mathcal{D}(M_f) = \tilde{\mathcal{D}}(f, \mathbf{M}) = \mathcal{D}(f, \mathbf{M})$. Its adjoint is $M_f^* = M_{\bar{f}}$. If $f \in L^\infty(\mathbb{R})$, then M_f is a bounded operator with the operator norm $\|M_f\| = \text{ess sup}_{x \in \mathbb{R}} |f(x)|$. Especially, we may define the selfadjoint operator M_x where we have denoted the identity function of $\mathbb{R}$ simply by x , i.e. $(M_x \varphi)(y) = y\varphi(y)$ for all $y \in \mathbb{R}$ and for $\varphi \in \mathcal{D}(M_x) = \left\{ \psi \in L^2(\mathbb{R}) \mid \int y^2 |\varphi(y)|^2 dy < \infty \right\}$.

Proposition 8.1 *If a bounded operator $S \in \mathcal{L}(L^2(\mathbb{R}))$ commutes with every M_f , $f \in L^\infty(\mathbb{R})$, then $S = M_g$ for some $g \in L^\infty(\mathbb{R})$.*

Proof Denote by μ the positive Borel measure $X \mapsto \int_X \psi(t)dt$, where $\psi(t) = \exp(-t^2)$. The map $\phi \mapsto \sqrt{\psi}\phi$ is an isometric isomorphism from $L^2(\mathbb{R}, \mu)$ onto $L^2(\mathbb{R})$. Since everything in the claim remains invariant under this isomorphism, we may replace $L^2(\mathbb{R})$ with $L^2(\mathbb{R}, \mu)$ for the finite measure μ . Now denote $g = S1 \in L^2(\mathbb{R}, \mu)$. If $f \in L^\infty(\mathbb{R}, \mu) \subset L^2(\mathbb{R}, \mu)$, then $gf = M_f g = M_f S1 = SM_f 1 = Sf$. The function g belongs to $L^\infty(\mathbb{R}, \mu)$, for otherwise for each $n \in \mathbb{N}$ there is a Borel set X_n with $\mu(X_n) > 0$ such that $|g(t)| \geq n$ on X_n , contradicting the fact that S is continuous in the L^2 -norm. Since $L^\infty(\mathbb{R}, \mu)$ is dense in $L^2(\mathbb{R}, \mu)$, it is clear that $S = M_g$. $\square$

Since every $M_f, f \in L^\infty(\mathbb{R})$, is a norm limit of a sequence of linear combinations of projections of the type $\mathbf{M}(X)$, any $S \in \mathcal{L}(L^2(\mathbb{R}))$ commuting with every $\mathbf{M}(X)$ also commutes with every $M_f, f \in L^\infty(\mathbb{R})$. We thus get the following corollary showing that the canonical spectral measure $\mathbf{M}$ has the important property of being *maximal*.

Corollary 8.1 *If P is a projection in $\mathcal{L}(L^2(\mathbb{R}))$ commuting with every $\mathbf{M}(X)$, $X \in \mathcal{B}(\mathbb{R})$, then $P = \mathbf{M}(Y)$ for some $Y \in \mathcal{B}(\mathbb{R})$.*

In view of the fact that the range of any spectral measure is a Boolean sub- σ -algebra of the projection lattice $\mathcal{P}(\mathcal{H})$, Corollary 8.1 states that the range $\{\mathbf{M}(X) \mid X \in \mathcal{B}(\mathbb{R})\}$ of $\mathbf{M}$ is a maximal Boolean sub- σ -algebra of $\mathcal{P}(\mathcal{H})$. Applying Sikorski's theorem in conjunction with the fact that $\mathcal{B}(\mathbb{R})$ is separable [1] one then obtains another well-known result.

Corollary 8.2 *If a selfadjoint operator A commutes with the selfadjoint operator M_x , that is, all the spectral projections of A commute with all the spectral projections $\mathbf{M}(X)$ of M_x , then there is a Borel function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $A = L(f, \mathbf{M})$.*

Let F be the Fourier-Plancherel operator on $L^2(\mathbb{R})$, that is, the unitary operator that extends the Fourier transform on $C^\infty_0(\mathbb{R})$, the Schwartz space of C^∞ functions of rapid decrease: for all $f \in C^\infty_0(\mathbb{R})$,

$$(Ff)(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ixy} f(x) dx \quad (8.1)$$

$$(F^{-1}f)(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixy} f(x) dx. \quad (8.2)$$

Since $F : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ is unitary, the map $X \mapsto F^{-1}\mathbf{M}(X)F = \mathbf{M}_F(X)$ is a spectral measure. Due to the unitary equivalence of $\mathbf{M}$ and $\mathbf{M}_F$, the spectral measure $\mathbf{M}_F$ is maximal, too.

The selfadjoint operator $L(x, \mathbf{M}_F)$ defined by the spectral measure $\mathbf{M}_F$ is known to be the differential operator $-iD$, with the domain (see, e.g. [2, Sect. 49])

$$\begin{aligned} \mathcal{D}(x, \mathbf{M}_F) &= \tilde{\mathcal{D}}(x, \mathbf{M}_F) = \mathcal{D}(-iD) \\ &= \{\varphi \in L^2(\mathbb{R}) \mid \varphi \text{ absolutely continuous, } \varphi' \in L^2(\mathbb{R})\}. \end{aligned}$$

The unitary equivalence of the spectral measures M and M_F gives also the unitary equivalence of the operators, $-iD = F^{-1}M_xF$ on a dense domain.

The selfadjoint operators M_x and $-iD$ and their spectral measures M and M_F have a number of important coupling properties, their noncommutativity

$$M_x(-iD) - (-iD)M_x = iI$$

on a dense domain (containing $C^\infty_{\downarrow}(\mathbb{R})$) being immediate.

Proposition 8.2 *For any bounded sets $X, Y \in \mathcal{B}(\mathbb{R})$,*

$$M(X) \wedge M_F(Y) = M(X) \wedge M_F(Y^c) = M(X^c) \wedge M_F(Y) = 0.$$

Proof For any $X, Y \in \mathcal{B}(\mathbb{R})$, the condition $M(X) \wedge M_F(Y) \neq 0$ implies that $M(X)\varphi = \varphi$ and $M_F(Y)\varphi = \varphi$, for some $\varphi \in L^2(\mathbb{R})$, $\varphi \neq 0$, that is, $\varphi = 0$ a.e. in X^c and $F\varphi = 0$ a.e. in Y^c . If, for instance, X is bounded, then $\varphi \in L^2(\mathbb{R})$ is compactly supported and its Fourier transform $z \mapsto \frac{1}{2\pi} \int_{\text{supp}(\varphi)} e^{-itz} \varphi(t) dt$ is an analytic function in the whole complex plane $\mathbb{C}$. The identity theorem of analytic functions then says that the Fourier transform of φ cannot vanish on a subset of $\mathbb{R}$ of positive measure unless it vanishes identically. $\square$

The following important result is an immediate corollary.

Corollary 8.3 *The operators M_x and $-iD$ are totally noncommutative, that is their commutativity domain*

$$\text{dom}(M_x, -iD) = \{\varphi \in \mathcal{H} \mid M(X)M_F(Y)\varphi = M_F(Y)M(X)\varphi \text{ for all } X, Y \in \mathcal{B}(\mathbb{R})\}$$

contains only the null vector: $\text{dom}(M_x, -iD) = \{0\}$.

Remark 8.1 The selfadjoint operators $L(x, M) = M_x$ and $L(x, M_F) = -iD$ are the standard position and momentum operators Q and P of a nonrelativistic spin-0 particle moving on the line $\mathbb{R}$ if $L^2(\mathbb{R})$ is the Hilbert space of the position representation, and $\hbar = 1$. In this interpretation it is helpful to keep track of the physical dimensions. If $L^2(\mathbb{R})$ is the position space, the points $x \in \mathbb{R}$ have the dimension of *length*; to highlight this we write temporarily $L^2(\mathbb{R}_Q)$ and $q \in \mathbb{R}_Q$. In order to get the appropriate physical dimension for the momentum, the Fourier-Plancherel transform should go from the position representation space $L^2(\mathbb{R}_Q)$ to the momentum representation space $L^2(\mathbb{R}_P)$, with the points $p \in \mathbb{R}_P$ having the dimension of *momentum*. In that case the Fourier transform needs to be adjusted accordingly since the standard physical dimension of the product qp , $q \in \mathbb{R}_Q$, $p \in \mathbb{R}_P$, is that of action, $[qp] = \text{kg m}^2/\text{s} = \text{Js}$. Using the Dirac constant $\hbar$, which has the right dimension, the transforms (8.1) and (8.2) must be replaced by

$$(F\psi)(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}_Q} e^{-iqp/\hbar} \psi(q) dq$$

$$(F^{-1}\varphi)(q) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}_P} e^{iqp/\hbar} \varphi(p) dp$$

and P by $P = -i\hbar D$. Usually, we use units where $\hbar = 1$ to simplify our notations but it is easy to display $\hbar$ by replacing P by $\hbar P$. In the context of uncertainty relations, $\hbar$ often occurs explicitly.

We recall that the deeper meaning of Planck's constant h and Dirac's constant $\hbar = h/(2\pi)$ as universal conversion factors is obtained from the idea of Galilei invariance, where $\hbar$ converts the multiplier of a projective representation of the Galilei group to the physical mass of the quantum object described by the representation, see, e.g. [3]. Its present numerical value is $\hbar = 1.054572 \times 10^{-34}$ Js [4]. $\triangleleft$

8.2 Restrictions of Spectral Measures

Semispectral measures arise from spectral measures as Naimark projections. In many applications, particularly in the measurement context, the semispectral measure $\mathbf{E}$ at hand is actually a spectral measure. The following result characterises such cases.

Lemma 8.1 *For any two projection operators P and R on a Hilbert space $\mathcal{H}$, the product PRP is a projection if and only if $PR = RP$.*

Proof If $PR = RP$, then PRP is a projection. Assume now that PRP is a projection. Then $PRPRP = PRP$ and thus $((I - P)RP)^*((I - P)RP) = PR(I - P)RP = 0$. It follows that $(I - P)RP = 0$, and so $RP = PRP$ which in turn means that RP is selfadjoint and thus $PR = P^*R^* = (RP)^* = RP$. $\square$

Corollary 8.4 *Let $(\mathcal{K}, \mathbf{F}, V)$ be a Naimark dilation of $\mathbf{E}$. For any $X \in \mathcal{A}$, $\mathbf{E}(X) = V^*\mathbf{F}(X)V$ is a projection if and only if the Naimark projection VV^* commutes with the projection $\mathbf{F}(X)$.*

Proof If the projection operators VV^* and $\mathbf{F}(X)$ commute with each other, then $\mathbf{E}(X)^2 = V^*\mathbf{F}(X)V V^*\mathbf{F}(X)V = V^*V V^*\mathbf{F}(X)V = V^*\mathbf{F}(X)V = \mathbf{E}(X)$, since $V^*V = I_{\mathcal{H}}$. Assume next that $\mathbf{E}(X)^2 = \mathbf{E}(X)$ so that $V^*\mathbf{F}(X)V V^*\mathbf{F}(X)V = V^*\mathbf{F}(X)V$. But then also $V V^*\mathbf{F}(X)V V^*\mathbf{F}(X)V V^* = V V^*\mathbf{F}(X)V V^*$, showing that $V V^*\mathbf{F}(X)V V^*$ is a projection. By Lemma 8.1 this means that the projections $V V^*$ and $\mathbf{F}(X)$ commute with each other. $\square$

Corollary 8.5 *A dilation $(\mathcal{K}, \mathbf{F}, V)$ of a spectral measure $\mathbf{E}$ is minimal if and only if $\mathcal{K} = V(\mathcal{H})$.*

Proof For any $X \in \mathcal{A}$ and $\varphi \in \mathcal{H}$, $V V^*(\mathbf{F}(X)V\varphi) = \mathbf{F}(X)V V^*V\varphi = \mathbf{F}(X)V\varphi$, showing that $\mathbf{F}(X)V\varphi \in V(\mathcal{H})$. Therefore, $(\mathcal{K}, \mathbf{F}, V)$ can be minimal only if $\mathcal{K} = V(\mathcal{H})$. If $\mathcal{K} = V(\mathcal{H})$, then clearly $(\mathcal{K}, \mathbf{F}, V)$ is minimal. $\square$

We consider next the restrictions of the canonical spectral measure $\mathbf{M}$ of $\mathcal{H} = L^2(\mathbb{R})$. Due to the maximality of $\mathbf{M}$, any projection P which commutes with $\mathbf{M}$ is of the form $P = \mathbf{M}(Y)$, $Y \in \mathcal{B}(\mathbb{R})$. For such projections the space $P(\mathcal{H})$ can be identified with the space $L^2(Y)$ and the map $X \mapsto P\mathbf{M}(X)|_{P(\mathcal{H})}$ can be identified with the canonical spectral measure on $L^2(Y)$. Thus, to obtain a semispectral measure $X \mapsto \mathbf{E}(X) = P\mathbf{M}(X)|_{P(\mathcal{H})}$, which is not a spectral measure, it is necessary and sufficient that P is not in the range of $\mathbf{M}$, that is, $P(\mathcal{H})$ is not invariant under the action of the projections $\mathbf{M}(X)$, $X \in \mathcal{B}(\mathbb{R})$.

Let K be any orthonormal basis of $L^2(\mathbb{R})$, $K' \subset K$, $K' \neq K$, and $P_{K'} = \sum_{f \in K'} |f\rangle\langle f|$. Clearly, $P_{K'}(\mathcal{H})$ is not invariant under $\mathbf{M}$. The Naimark projection of $\mathbf{M}$ onto $P_{K'}(\mathcal{H})$ is the semispectral measure $X \mapsto \mathbf{E}(X)$ which has the form

$$\mathbf{E}(X) = \sum_{f, g \in K'} \int_X \overline{f(x)} g(x) dx |f\rangle\langle g|.$$

As a further illustration, consider the canonical spectral measure $\mathbf{M}$ of the Hilbert space $\mathcal{H} = L^2([0, 2\pi])$, and let K be the Fourier basis of $\mathcal{H}$, that is, $K = \{e_k \mid k \in \mathbb{Z}\}$, with $e_k(x) = \frac{1}{\sqrt{2\pi}} e^{-ikx}$, $x \in [0, 2\pi]$. Let P be the projection $\sum_{n=0}^{\infty} |e_n\rangle\langle e_n| \in \mathcal{L}(\mathcal{H})$. Clearly, it is not of the form $\mathbf{M}(X)$ for any Borel set $X \subset [0, 2\pi]$ so that the Naimark projection of $\mathbf{M}$ onto $P(\mathcal{H})$ is a semispectral measure of the above type,

$$\Phi_{\text{can}}(X) = \sum_{n, m=0}^{\infty} \frac{1}{2\pi} \int_X e^{i(n-m)x} dx |e_n\rangle\langle e_m| \in \mathcal{L}(P(\mathcal{H})). \quad (8.3)$$

This operator measure has important applications in the theory of covariant phase observables, which is the subject of Chap. 16. Here we consider some basic properties of (8.3).

The moment operators $\Phi_{\text{can}}[n] \equiv L(x^n, \Phi_{\text{can}}) = \int_0^{2\pi} x^n d\Phi_{\text{can}}(x)$, $n \in \mathbb{N}_0$, are bounded selfadjoint operators. The same is true of the sine and cosine operators $L(\sin, \Phi_{\text{can}})$ and $L(\cos, \Phi_{\text{can}})$, respectively. If $(\mathcal{K}, \mathbf{F}, V)$ is any Naimark dilation of Φ_{can} into a projection measure $\mathbf{F}$, then, for all $n \in \mathbb{N}_0$, $\Phi_{\text{can}}[n] = V^* A^n V$, as well as $L(\sin, \Phi_{\text{can}}) = V^* \sin(A) V$ and $L(\cos, \Phi_{\text{can}}) = V^* \cos(A) V$, where $A = \int_0^{2\pi} x d\mathbf{F} = L(x, \mathbf{F}) = \mathbf{F}[1]$.

The minimal Naimark dilation of Φ_{can} is also easily constructed. By construction, the canonical spectral measure $\mathbf{M}$ and the isometry $V : P(\mathcal{H}) \rightarrow \mathcal{H}$, for which $V(e_n) = e_n$, for all $n \in \mathbb{N}_0$, constitute a dilation of Φ_{can} . According to the next theorem, this is the minimal dilation.

Theorem 8.1 *The closed linear span D of the vectors $\mathbf{M}(X)Vg$, $g \in P(\mathcal{H})$, $X \in \mathcal{B}([0, 2\pi])$, is the whole of $\mathcal{H} = L^2([0, 2\pi])$.*

Proof Since the constant function $1 = \sqrt{2\pi}e_0$ belongs to $P(\mathcal{H})$ and $\mathbf{M}(X)V1$ is the characteristic function of $X \in \mathcal{B}([0, 2\pi])$, by approximating any $f \in L^2([0, 2\pi])$ uniformly on $[0, 2\pi]$ by linear combinations of characteristic functions of sets in $\mathcal{B}([0, 2\pi])$, we see in the limit that $f \in D$. Hence, $D = L^2([0, 2\pi])$. $\square$

The operator measure (8.3) is unitarily equivalent to the Toeplitz measure on the Hardy subspace H^2 of $L^2([0, 2\pi])$. Indeed, recalling that $H^2 = V(P(\mathcal{H}))$ and $P = VV^*$, the Naimark projection of $L^2([0, 2\pi])$ onto H^2 , we may write

$$\Phi_{\text{can}}(X) = V^*M(X)V = V^*PM(X)V$$

where $X \mapsto PM(X)$ is the Toeplitz measure. For any bounded Borel function $f : [0, 2\pi] \rightarrow \mathbb{R}$, the selfadjoint operators $L(f, \Phi_{\text{can}})$ and $P(L(f, M)) = P(M_f)$ are thus unitarily equivalent, too. In particular, this means that their spectra $\sigma(L(f, \Phi_{\text{can}}))$ and $\sigma(P(M_f))$ are the same.

Theorem 8.2 *Let $f : [0, 2\pi] \rightarrow \mathbb{R}$ be a bounded Borel function. The spectrum of the operator $L(f, \Phi_{\text{can}})$ is the closed interval $[\text{ess inf } f, \text{ess sup } f]$. In addition, if f is not constant (almost everywhere) then $L(f, \Phi_{\text{can}})$ has no eigenvalues.*

Proof The operator $L(f, \Phi_{\text{can}})$ is unitarily equivalent to the Toeplitz operator $P(M_f)$. By the Hartman-Wintner theorem [5, p. 179], the spectrum of the Toeplitz operator $P(M_f)$ is $[\text{ess inf } f, \text{ess sup } f]$. Suppose that $L(f, \Phi_{\text{can}})g = \lambda g$, i.e. $L(f - \lambda, \Phi_{\text{can}})g = 0$, for some nonzero $g \in P(\mathcal{H})$ and $\lambda \in \mathbb{R}$. Then, for any nonnegative n , $ge_n = \sum_{m=0}^{\infty} \langle e_m | g \rangle e_m e_n \in P(\mathcal{H})$ and

$$0 = \langle L(f - \lambda, \Phi_{\text{can}})g | ge_n \rangle = \int_0^{2\pi} [f(x) - \lambda] |g(x)|^2 e_n(x) dx$$

which, by complex conjugation, holds for all $n \in \mathbb{Z}$. Hence, the function $x \mapsto [f(x) - \lambda] |g(x)|^2 = 0$ for almost all $x \in [0, 2\pi]$ due to the uniqueness of the Fourier coefficients. But $x \mapsto f(x) - \lambda$ is not constant (especially zero) almost everywhere so that $g(x) = 0$ on a set of positive measure. It is a standard result that if a function of H^2 vanishes on a set of positive measure then it is zero almost everywhere. Thus, $g = 0$ yielding a contradiction. $\square$

As an immediate application of the above theorem one may determine, for instance, the spectra of the operators $\Phi_{\text{can}}(X)$, $\Phi_{\text{can}}[n]$, $L(\sin, \Phi_{\text{can}})$, and $L(\cos, \Phi_{\text{can}})$. In particular, for any $X \in \mathcal{B}([0, 2\pi])$ for which both X and X^c have nonzero Lebesgue measure, $\sigma(\Phi_{\text{can}}(X)) = [0, 1]$ with no eigenvalues. Similarly, for any $n \in \mathbb{N}$, $\sigma(\Phi_{\text{can}}[n]) = [0, (2\pi)^n]$, and also $\sigma(L(\sin, \Phi_{\text{can}})) = [-1, 1]$ and $\sigma(L(\cos, \Phi_{\text{can}})) = [-1, 1]$.

8.3 Smearings and Convolutions

Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be two measurable spaces. We say that a map $p : \mathcal{A}_2 \times \Omega_1 \mapsto [0, 1]$ is a *Markov kernel* if

- (i) for each $X \in \mathcal{A}_2$, $p(X, \cdot) : \Omega_1 \rightarrow [0, 1]$, $\omega \mapsto p(X, \omega)$ is a measurable function,
- (ii) for every $\omega \in \Omega_1$, $p(\cdot, \omega)$ is a probability measure on $\mathcal{A}_2$.

Consider a semispectral measure $E_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$. For any $X \in \mathcal{A}_2$, the operator

$$L(p(X, \cdot), E_1) = \int_{\Omega_1} p(X, \omega) dE_1(\omega)$$

is a unit bounded positive operator, and the assignment

$$X \mapsto L(p(X, \cdot), E_1) = E_2(X) \quad (8.4)$$

defines a semispectral measure E_2 on $\mathcal{A}_2$. We refer to p as a (conditional) confidence measure. Such a confidence measure p on $\mathcal{A}_2 \times \Omega_1$ induces a map from the semispectral measures on $\mathcal{A}_1$ to those on $\mathcal{A}_2$ with the same Hilbert space. We call E_2 a *smearing* of E_1 by p and denote $E_2 = p \star E_1$.

If E_1 is a spectral measure, then due to its multiplicativity, the semispectral measure E_2 is commutative. It can be shown that a semispectral measure derives from a spectral measure according to (8.4) if and only if it is commutative [6, p. 43].

Forming convolutions of complex measures and semispectral measures is another important method of constructing new semispectral measures. We shall describe this construction next.

The *convolution* of two complex Borel measures $\mu, \nu : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{C}$ is the measure $\mu * \nu : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{C}$, defined by

$$(\mu * \nu)(X) = (\mu \times \nu)(\{(x, y) \mid x + y \in X\}), \quad X \in \mathcal{B}(\mathbb{R}),$$

where $\mu \times \nu$ is the product measure defined on $\mathcal{B}(\mathbb{R}^2)$. The following lemma is crucial for our applications.

Lemma 8.2 *Let $\mu, \nu : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{C}$ be two complex measures, and let $k \in \mathbb{N}_0$.*

- (a) *A measurable function $f : \mathbb{R} \rightarrow \mathbb{C}$ is $\mu * \nu$ -integrable if $(x, y) \mapsto f(x + y)$ is integrable with respect to the product measure $\mu \times \nu : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathbb{C}$. In that case,*

$$\int f(x) d(\mu * \nu)(x) = \int f(x + y) d(\mu \times \nu)(x, y).$$

- (b) *The function $(x, y) \mapsto (x + y)^k$ is $\mu \times \nu$ -integrable if and only if $x \mapsto x^k$ is both μ - and ν -integrable. In that case,*

$$\int (x + y)^k d(\mu \times \nu)(x, y) = \sum_{n=0}^k \binom{k}{n} \left(\int x^{k-n} d\mu(x) \right) \left(\int y^n d\nu(y) \right).$$

Proof Let $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ denote addition, i.e. $\phi(x, y) = x + y$. Write the product measure $\mu \times \nu$ in terms of the positive and negative parts of its real and imaginary parts:

$$\mu \times \nu = \nu_1 + i\nu_2 = \nu_1^+ - \nu_1^- + i(\nu_2^+ - \nu_2^-),$$

where $\nu_i^\pm = \frac{1}{2}(|\nu_i| \pm \nu_i)$. Then for all $X \in \mathcal{B}(\mathbb{R})$

$$(\mu * \nu)(X) = \nu_1^+(\phi^{-1}(X)) - \nu_1^-(\phi^{-1}(X)) + i(\nu_2^+(\phi^{-1}(X)) - \nu_2^-(\phi^{-1}(X))). \quad (8.5)$$

Assume now that $f \circ \phi$ is $\mu \times \nu$ -integrable. Then $f \circ \phi$ is integrable with respect to each $\nu_i^\pm$, and

$$\int f \circ \phi d(\mu \times \nu) = \int f \circ \phi d\nu_1^+ - \int f \circ \phi d\nu_1^- + i \left(\int f \circ \phi d\nu_2^+ - \int f \circ \phi d\nu_2^- \right). \quad (8.6)$$

Since the measures $\nu_i^\pm$ are positive, it follows that f is integrable with respect to each induced measure $X \mapsto \nu_i^\pm(\phi^{-1}(X))$, and the corresponding integrals are equal (see e.g. [7, p. 163]). Now (8.5) and (8.6) imply that f is $\mu * \nu$ -integrable, with

$$\int f d(\mu * \nu) = \int f \circ \phi d(\mu \times \nu).$$

This proves (a).

To prove (b), suppose first that $(x, y) \mapsto \phi(x, y)^k$ is $\mu \times \nu$ -integrable. Since both measures μ and ν are finite, it follows from [8, Theorem 13] that $x \mapsto \phi(x, y)^k = (x + y)^k$ is μ -integrable for ν -almost all $y \in \mathbb{R}$. Take any such $y \in \mathbb{R}$. Now $x \mapsto |x + y|^k$ is also $|\mu|$ -integrable, where $|\mu|$ denotes the total variation measure of μ . There are positive constants M and K satisfying

$$|x^k| \leq K|x + y|^k + M, \quad x \in \mathbb{R}.$$

This implies that $x \mapsto |x|^k$ is $|\mu|$ -integrable, and hence also μ -integrable. It is similarly seen that $x \mapsto |x|^k$ is ν -integrable.

Suppose now that $x \mapsto x^k$ is both μ - and ν -integrable. Since $x \mapsto |x|^k$ is now $|\mu|$ - and $|\nu|$ -integrable, and these are finite positive measures, it follows that $x \mapsto |x|^l$ is $|\mu|$ - and $|\nu|$ -integrable for all $l \in \mathbb{N}_0$, $l \leq k$. Hence, $(x, y) \mapsto |x^l y^m|$ is $|\mu| \times |\nu|$ -integrable for all $l, m \in \mathbb{N}_0$, $l \leq k$, $m \leq k$. Since $|x + y|^k \leq \sum_{n=0}^k \binom{k}{n} |x^{k-n} y^n|$, this implies that $(x, y) \mapsto |\phi(x, y)^k| = |(x + y)^k|$ is $|\mu| \times |\nu|$ -integrable. But $|\mu| \times |\nu| = |\mu \times \nu|$ by [8, p. 192, Lemma 11], so $x \mapsto \phi(x, y)^k$ is $\mu \times \nu$ -integrable.

The claimed formula now follows easily, since we have shown above that the equivalent integrability conditions imply that $(x, y) \mapsto x^l y^m$ is $\mu \times \nu$ -integrable for all $l, m \in \mathbb{N}_0$, $l \leq k$, $m \leq k$. $\square$

Remark 8.2 The converse implication in part (a) of the above lemma does not hold if the measures μ and ν are not assumed to be positive. A discussion of this aspect together with a counterexample is given in [9]. $\triangleleft$

Let again $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ denote addition. For any $X \in \mathcal{B}(\mathbb{R})$, the function $\chi_X \circ \phi$ is bounded and thereby integrable with respect to the product measure. The above result together with Fubini's theorem then implies that the function

$$y \mapsto \mu(X - y) = \int \chi_{X-y}(x) d\mu(x) = \int \chi_X(x + y) d\mu(x)$$

coincides almost everywhere with a Borel function, and

$$(\mu * \nu)(X) = \int \left(\int \chi_X(x + y) d\mu(x) \right) d\nu(y) = \int \mu(X - y) d\nu(y), \quad X \in \mathcal{B}(\mathbb{R}).$$

Let now $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure, and let $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be a probability measure. Since the sesquilinear form

$$(\varphi, \psi) \mapsto \int_{\mathbb{R}} \mu(X - y) d\mathbf{E}_{\varphi, \psi}(y)$$

is clearly bounded, one can define $\mu * \mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ via $\langle \varphi | (\mu * \mathbf{E})(X) \psi \rangle = \mu * \mathbf{E}_{\varphi, \psi}(X)$, $\varphi, \psi \in \mathcal{H}$. It follows from the monotone convergence theorem that $\mu * \mathbf{E}$ is a semispectral measure. This semispectral measure $\mu * \mathbf{E}$ is the *convolution* of $\mathbf{E}$ by μ and it is thus defined through the complex measures

$$(\mu * \mathbf{E})_{\varphi, \psi} = \mu * \mathbf{E}_{\varphi, \psi}, \quad \varphi, \psi \in \mathcal{H}. \quad (8.7)$$

Note that $\mu * \mathbf{E}$ is a smearing of $\mathbf{E}$ by a Markov kernel $p(X, y) = \mu(X - y)$, that is, $\mu * \mathbf{E} = p \star \mathbf{E}$. If $\mathbf{E}$ is a spectral measure then $\mu * \mathbf{E}$ is a commutative semispectral measure.

As an example, consider the canonical spectral measure $\mathbf{M}$ on $\mathcal{H} = L^2(\mathbb{R})$ and let $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be a probability measure. Their convolution $\mu * \mathbf{M}$ is a commutative semispectral measure. Due to their characteristic symmetry properties these measures have important applications. To describe those properties, consider the strongly continuous one parameter unitary groups $q \mapsto U(q)$ and $p \mapsto V(p)$ defined as

$$(U(q)\varphi)(x) = \varphi(x - q), \quad (8.8)$$

$$(V(p)\varphi)(x) = e^{ipx} \varphi(x), \quad (8.9)$$

for all $x \in \mathbb{R}$ and $\varphi \in \mathcal{H}$. As is well known, the selfadjoint generators of these groups are the operators iD and M_x . By a direct computation one may check that the semispectral measure $\mu * \mathbf{M}$ is *covariant* under U and *invariant* under V , that is, for all $q, p \in \mathbb{R}$, $X \in \mathcal{B}(\mathbb{R})$,

$$U(q)(\mu * \mathbf{M})(X)U(q)^* = (\mu * \mathbf{M})(X + q), \quad (8.10)$$

$$V(p)(\mu * \mathbf{M})(X)V(p)^* = (\mu * \mathbf{M})(X). \quad (8.11)$$

The following theorem, proved in [10], shows that the conditions (8.10) and (8.11) characterise the operator measure.

Theorem 8.3 *If $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$, $\mathcal{H} = L^2(\mathbb{R})$, is a semispectral measure satisfying the covariance and invariance conditions (8.10) and (8.11), then it is of the form $\mu * \mathbf{M}$ for some probability measure μ .*

Lemma 8.2 allows one to compute the moment operators of any semispectral measure $\mu * E$. If, in particular, $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure so that its first moment operator is a selfadjoint operator, then these moment operators have a particularly simple form as shown below. Detailed computations are given in [9].

Proposition 8.3 *Let $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be a spectral measure, let $k \in \mathbb{N}_0$, and let $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be a probability measure such that its k th moment $\mu[k]$ exists. Denote $A = L(x, E)$. Then*

$$L(x^k, \mu * E) = \sum_{n=0}^k \binom{k}{n} \mu[k-n] A^n, \quad \mathcal{D}(x^k, \mu * E) = D(A^k).$$

Moreover, $\widetilde{D}(x^k, \mu * E)$ equals either $D(A^k) = \mathcal{D}(x^k, \mu * E)$ or $\{0\}$, depending on whether $\mu[2k]$ exists or not.

8.4 Phase Space Operator Measures

Let $\mathcal{H} = L^2(\mathbb{R})$ and define the Weyl operators $W(q, p)$ as

$$W(q, p) = e^{i\frac{qp}{2}} U(q) V(p), \quad q, p \in \mathbb{R}, \quad (8.12)$$

where the strongly continuous unitary representations U and V of the additive group $\mathbb{R}$ are the ones introduced in (8.8) and (8.9).

Proposition 8.4 *Let $A \in \mathcal{L}(\mathcal{H})$ and assume that $AW(q, p) = W(q, p)A$ for all $(q, p) \in \mathbb{R}^2$. Then $A = cI$ for some $c \in \mathbb{C}$.*

Proof Since $AW(q, p) = W(q, p)A$ for all $(q, p) \in \mathbb{R}^2$ implies that $A^*W(q, p) = W(q, p)A^*$ for all $(q, p) \in \mathbb{R}^2$, it suffices to assume that A is selfadjoint. The commutativity assumption includes the assumption that $AW(q, 0) = W(q, 0)A$ and $AW(0, p) = W(0, p)A$ for all $q, p \in \mathbb{R}$. Since the ranges of the spectral measures of M_x and $-iD$ are maximal Boolean sub- σ -algebras of $\mathcal{P}(\mathcal{H})$, any spectral projection of A is also a spectral projection of both M_x and $-iD$. But a spectral projection of M_x which commutes with all the spectral projections of $-iD$ is either 0 or I (since the only nonempty Borel subset of the real line $\mathbb{R}$ which is invariant under all the translations (8.8) is $\mathbb{R}$). Hence A is constant. $\square$

Recalling that an operator $A \in \mathcal{L}(\mathcal{H})$ commutes with a one-parameter unitary group exactly when it commutes with the spectral projections of the selfadjoint generator of the group we get the following well-known result.

Corollary 8.6 Any $A \in \mathcal{L}(\mathcal{H})$ which commutes with the operators M_x and $-iD$ is a constant.

The map $(q, p) \mapsto W(q, p)$ thus constitutes an irreducible projective representation of $\mathbb{R}^2$ with the multiplication law

$$W(q, p)W(q', p') = e^{-i\frac{qp' - q'p}{2}} W(q + q', p + p'). \quad (8.13)$$

In particular, $W(q, p)^* = W(-q, -p)$.

For each $x = (q, p) \in \mathbb{R}^2$, define the map $\gamma(x) : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ by $\gamma(x)(T) = W(x)TW(-x)$. It has the following well-known properties.

- Proposition 8.5** (a) $\gamma(x + y) = \gamma(x) \circ \gamma(y)$ for all $x, y \in \mathbb{R}^2$.
 (b) $\gamma(x)^*(A) = W(-x)AW(x)$ for all $A \in \mathcal{L}(\mathcal{H})$ and $x \in \mathbb{R}^2$.
 (c) $\gamma(x)$ is a positive trace-norm isometry for all $x \in \mathbb{R}^2$.
 (d) For each $A \in \mathcal{L}(\mathcal{H})$ and $S \in \mathcal{T}(\mathcal{H})$, the function $x \mapsto \text{tr}[A\gamma(x)(S)]$ is continuous.
 (e) $\int \text{tr}[P_1\gamma(x)(P_2)]dx = 2\pi$ for all one-dimensional projections P_1 and P_2 on $\mathcal{H}$.

Proof (Adapted from [11]) Part (a) is a direct consequence of the relation (8.13), and (b) follows from a basic property of the trace. If U is a unitary operator, $|USU^*| = U|S|U^*$ for each $S \in \mathcal{L}(\mathcal{H})$. Therefore, since $W(x)$ is unitary and $W(x)^* = W(-x)$, $\|\gamma(x)(S)\|_1 = \text{tr}[|W(x)SW(-x)|] = \|S\|_1$ for each $S \in \mathcal{T}(\mathcal{H})$. This proves (c), as it is clear that $\gamma(x)$ is positive.

To prove (d), take $A \in \mathcal{L}(\mathcal{H})$ and $S \in \mathcal{T}(\mathcal{H})$. Let $x \in \mathbb{R}^2$, and (x_n) be a sequence converging to x . Since $x \mapsto W(x)$ is strongly continuous, $\gamma(x_n)^*(A) = W(-x_n)AW(x_n) \rightarrow W(-x)AW(x) = \gamma(x)^*(A)$ weakly. Since all $W(x)$ are unitary, the sequence $(\gamma(x_n)^*(A))$ is norm bounded, from which it follows that it converges to $\gamma(x)^*(A)$ also ultraweakly. Thus we get

$$\text{tr}[A\gamma(x_n)(S)] = \text{tr}[\gamma(x_n)^*(A)S] \rightarrow \text{tr}[\gamma(x)^*(A)S] = \text{tr}[A\gamma(x)(S)],$$

which proves (d).

To prove (e), assume that $P_1 = |\psi\rangle\langle\psi|$ and $P_2 = |\varphi\rangle\langle\varphi|$, where $\psi, \varphi \in \mathcal{H}$ are unit vectors. Define the function ϕ_q for each $q \in \mathbb{R}$ by $\phi_q(t) = \psi(t)\varphi(t + q)$. Then

$$1 = \|\psi\|^2\|\varphi\|^2 = \int \left(\int |\psi(t)|^2 |\varphi(q)|^2 dq \right) dt = \int \left(\int |\psi(t)|^2 |\varphi(t + q)|^2 dt \right) dq$$

by the Fubini–Tonelli theorem, so that $\phi_q \in \mathcal{H}$ for almost all q . By the unitarity of the inverse Fourier–Plancherel operator F , we have now

$$1 = \int \int |(F\phi_q)(p)|^2 dp dq.$$

But since ψ and $\varphi(\cdot + q)$ are in $\mathcal{H}$, ϕ_q is also integrable, so

$$(F\phi_q)(p) = \frac{1}{\sqrt{2\pi}} \int e^{ip \cdot t} \phi_q(t) dt = \frac{1}{\sqrt{2\pi}} e^{-i\frac{1}{2}p \cdot q} \langle \psi | W(q, p) \varphi \rangle,$$

from which it follows that

$$2\pi = \int |\langle \psi | W(x) \varphi \rangle|^2 dx = \int \text{tr}[P_1 \gamma(x) (P_2)] dx$$

and the proposition is proved. $\square$

Let $T \in \mathcal{T}(\mathcal{H})$ be positive, $T \geq 0$, and of trace one, $\text{tr}[T] = 1$, and consider the function $(q, p) \mapsto \frac{1}{2\pi} W(q, p) T W(q, p)^*$. By the above proposition one may now easily check that the weakly defined integral

$$\mathbf{G}^T(Z) = \frac{1}{2\pi} \int_{\mathbb{Z}} W(q, p) T W(q, p)^* dq dp \quad (8.14)$$

(using the duality $\mathcal{C}(\mathcal{H})^* = \mathcal{T}(\mathcal{H})$) defines a semispectral measure $\mathbf{G}^T : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$, called a *phase space semispectral measure* (generated by T). They will be extensively used in Chaps. 15 and 18.

To close this section we give some basic properties of the phase space semispectral measures (8.14). A direct computation gives the following.

Proposition 8.6 *Any phase space semispectral measure $\mathbf{G}^T$ is covariant under the action of the representation $(q, p) \mapsto W(q, p)$, that is, for any $Z \in \mathcal{B}(\mathbb{R}^2)$ and $(q, p) \in \mathbb{R}^2$,*

$$W(q, p) \mathbf{G}^T(Z) W(q, p)^* = \mathbf{G}^T(Z + (q, p)).$$

This covariance property is a characteristic property of the semispectral measures (8.14). The theorem below goes back to [12, 13]. A direct measure theoretical proof is given in [11] whereas an alternative group theoretical proof can be found in [14].

Theorem 8.4 *If a semispectral measure $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ satisfies the condition $W(q, p) \mathbf{E}(Z) W(q, p)^* = \mathbf{E}(Z + (q, p))$ for all $Z \in \mathcal{B}(\mathbb{R}^2)$, $(q, p) \in \mathbb{R}^2$, then $\mathbf{E}$ is of the form (8.14) for some positive operator T of trace one, i.e. $\mathbf{E} = \mathbf{G}^T$.*

The irreducibility of the representation $(q, p) \mapsto W(q, p)$ implies the following proposition. A detailed proof is given in [15, Proposition 3].

Proposition 8.7 *The only projections in the range of a phase space semispectral measure $\mathbf{G}^T$ are 0 and I .*

Assume that the generating operator T is a one-dimensional projection, that is, $T = P[\varphi]$ for some unit vector $\varphi \in \mathcal{H}$. In that case we write $\mathbf{G}^\varphi$ instead of $\mathbf{G}^{P[\varphi]}$. Let $\mathcal{K} =$

$L^2(\mathbb{R}^2, (2\pi)^{-1} dq dp)$ and let $\mathbf{M} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{K})$ be the canonical spectral measure $Z \mapsto \mathbf{M}(Z)$, $\mathbf{M}(Z)\Psi = \chi_Z\Psi$. Consider the mapping $V_\varphi : \mathcal{H} \rightarrow \mathcal{K}$ defined by

$$(V_\varphi\psi)(q, p) = \langle \varphi | W(q, p)^*\psi \rangle_{\mathcal{H}}, \quad (q, p) \in \mathbb{R}^2.$$

The mapping is linear and isometric. Clearly, for each ψ , $\xi \in \mathcal{H}$ and $Z \in \mathcal{B}(\mathbb{R}^2)$

$$\langle \xi | \mathbf{G}^\varphi(Z)\psi \rangle = \langle V_\varphi\xi | \mathbf{M}(Z)V_\varphi\psi \rangle,$$

which shows that $(\mathcal{K}, V_\varphi, \mathbf{M})$ is a dilation of $\mathbf{G}^\varphi$. The minimality of this dilation follows from the irreducibility of the representation $(q, p) \mapsto W(q, p)$. We formulate this result as a proposition.

Proposition 8.8 *With the above notations, the triple $(\mathcal{K}, V_\varphi, \mathbf{M})$ is the minimal dilation of $\mathbf{G}^\varphi$.*

Consider a phase space semispectral measure $\mathbf{G}^T$ and let $T = \sum_i t_i |\varphi_i\rangle\langle\varphi_i|$ be the spectral decomposition of T . Let μ and ν be the probability measures with the densities $e(q) = \sum_i t_i |\varphi_i(-q)|^2$ and $f(p) = \sum_i t_i |(F\varphi_i)(-p)|^2$, that is, for instance, $\mu(X) = \int_X e(q) dq$. Let $\mathbf{G}_1^T$ and $\mathbf{G}_2^T$ be the marginal measures of $\mathbf{G}^T$, that is, $\mathbf{G}_1^T(X) = \mathbf{G}^T(X \times \mathbb{R})$ and $\mathbf{G}_2^T(Y) = \mathbf{G}^T(\mathbb{R} \times Y)$ for all $X, Y \in \mathcal{B}(\mathbb{R})$. The next result follows by a direct computation.

Proposition 8.9 *Let $\mathbf{G}^T$ be a phase space observable generated by the operator $T = \sum_i t_i |\varphi_i\rangle\langle\varphi_i|$, where $t_i \geq 0$, $\sum t_i = 1$, and the φ_i are mutually orthogonal unit vectors. Then $\mathbf{G}_1^T = \mu * \mathbf{M}$ and $\mathbf{G}_2^T = \nu * \mathbf{M}_F$, where $\mathbf{M}$ is the canonical spectral measure in $L^2(\mathbb{R})$ and $\mathbf{M}_F$ its Fourier-Plancherel transform.*

8.5 Moment Operators and Spectral Measures

An important class of operators of the form $L(f, \mathbf{E})$ are the moment operators of a semispectral measure $\mathbf{E}$ defined on (a Borel subspace of) the real value space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, or on some of its Cartesian products $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Some examples of these operators have already been discussed in the previous sections. In this section we show that a spectral measure $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is distinguished by the fact that its second moment operator is just the square of its first moment operator.

Consider a semispectral measure $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ and denote the function $\mathbb{R} \ni x \mapsto x^k \in \mathbb{R}$ by x^k . The operator integrals $L(x^k, \mathbf{E})$ are well-defined symmetric but not necessarily densely defined operators for each $k = 0, 1, 2, \dots$; they are the moment operators of $\mathbf{E}$. Occasionally, we use for them also a shorter notation $\mathbf{E}[k]$. Clearly, if $\mathbf{E}$ is bounded, that is, $\text{supp } \mathbf{E}$ is a bounded subset of $\mathbb{R}$, then the operators $L(x^k, \mathbf{E})$ are bounded selfadjoint operators.

If $\mathbf{E}$ is a spectral measure, the first moment $L(x, \mathbf{E})$ is always selfadjoint on the domain $\tilde{\mathcal{D}}(x, \mathbf{E})$, and $\int x^2 d\mathbf{E}_{\varphi, \varphi} = \|L(x, \mathbf{E})\varphi\|^2$ for all $\varphi \in \tilde{\mathcal{D}}(x, \mathbf{E})$. In the case of a general positive operator measure, this need not be the case. In fact, the following proposition shows that this condition is sufficient for a semispectral measure $\mathbf{E}$ to be a spectral measure.

For each $k \in \mathbb{N}_0$, we let $\tilde{L}(x^k, \mathbf{E})$ denote the restriction of $L(x^k, \mathbf{E})$ to $\tilde{\mathcal{D}}(x^k, \mathbf{E})$.

Proposition 8.10 *Let $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure such that*

$$\int x^2 d\mathbf{E}_{\varphi, \varphi} = \|L(x, \mathbf{E})\varphi\|^2 \quad (8.15)$$

for all $\varphi \in \tilde{\mathcal{D}}(x, \mathbf{E})$.

- (a) $\tilde{L}(x^n, \mathbf{E}) = \tilde{L}(x, \mathbf{E})^n$ for all $n \in \mathbb{N}_0$.
- (b) If $\tilde{L}(x, \mathbf{E})$ is selfadjoint, then $\mathbf{E}$ is projection valued.

Proof (Adapted from [15]) Let $\mathbf{F} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{K}$ be a Naimark dilation of $\mathbf{E}$ into a spectral measure acting on a Hilbert space $\mathcal{K}$. Let $V : \mathcal{H} \rightarrow \mathcal{K}$ be the associated isometric map, so that $\mathbf{E}(X) = V^*\mathbf{F}(X)V$ for all $X \in \mathcal{B}(\mathbb{R})$. Denote by $P_{\mathcal{H}}$ the projection VV^* , acting on $\mathcal{K}$ with $V\mathcal{H}$ as its range. Now $\tilde{L}(x^k, \mathbf{E}) = V^*L(x^k, \mathbf{F})V$ for each $k \in \mathbb{N}_0$. Since $\mathbf{F}$ is a spectral measure, we thus have

$$\tilde{L}(x^k, \mathbf{E}) = V^*A^kV \quad (8.16)$$

for all $k \in \mathbb{N}_0$, where $A = L(x, \mathbf{F})$.

Denote $E_1 = \tilde{L}(x, \mathbf{E})$. We prove by induction that for each $n \in \mathbb{N}_0$,

$$\tilde{\mathcal{D}}(x^n, \mathbf{E}) = \mathcal{D}(E_1^n), \text{ and } A^n V\varphi = VE_1^n\varphi \text{ for all } \varphi \in \tilde{\mathcal{D}}(x^n, \mathbf{E}). \quad (8.17)$$

Take first $n = 1$, and let $\varphi \in \tilde{\mathcal{D}}(x, \mathbf{E}) = \mathcal{D}(E_1) = \mathcal{D}(AV)$. Since the measures $\mathbf{E}_{\varphi, \varphi}$ and $\mathbf{F}_{V\varphi, V\varphi}$ are the same, and $\mathbf{F}$ is a spectral measure, the assumption implies that

$$\|AV\varphi\|^2 = \int x^2 d\mathbf{F}_{V\varphi, V\varphi} = \int x^2 d\mathbf{E}_{\varphi, \varphi} = \|E_1\varphi\|^2. \quad (8.18)$$

Using (8.16) and the fact that V is isometric, we thus get

$$\|AV\varphi\|^2 = \|E_1\varphi\|^2 = \|V^*AV\varphi\|^2 = \|P_{\mathcal{H}}AV\varphi\|^2. \quad (8.19)$$

Since $P_{\mathcal{H}}$ is a projection, this means that

$$AV\varphi = P_{\mathcal{H}}AV\varphi = VE_1\varphi \quad (8.20)$$

for all $\varphi \in \tilde{\mathcal{D}}(x, \mathbf{E}) = \mathcal{D}(E_1)$, i.e. (8.17) holds for $n = 1$. Now let $k \in \mathbb{N}$ and assume that (8.17) holds for $n = k - 1$. Let $\varphi \in \tilde{\mathcal{D}}(x^k, \mathbf{E})$. By (8.16), this implies that $V\varphi \in$

$\mathcal{D}(A^k)$, so that $AV\varphi \in \mathcal{D}(A^{k-1})$. Since $\varphi \in \tilde{\mathcal{D}}(x^k, E) \subset \tilde{\mathcal{D}}(x, E)$, it thus follows from (8.20) that $V(E_1\varphi) = AV\varphi \in \mathcal{D}(A^{k-1})$, so (8.16) and the induction assumption give $E_1\varphi \in \tilde{\mathcal{D}}(x^{k-1}, E) = \mathcal{D}(E_1^{k-1})$. Hence, $\varphi \in \mathcal{D}(E_1^k)$. Conversely, if $\varphi \in \mathcal{D}(E_1^k)$, then $\varphi \in \tilde{\mathcal{D}}(x, E) = \mathcal{D}(AV)$ and $E_1\varphi \in \mathcal{D}(E_1^{k-1}) = \tilde{\mathcal{D}}(x^{k-1}, E)$, so $AV\varphi = V(E_1\varphi) \in \mathcal{D}(A^{k-1})$ by (8.20) and (8.16), implying that $V\varphi \in \mathcal{D}(A^k)$, i.e. $\varphi \in \tilde{\mathcal{D}}(x^k, E)$. Thus, $\tilde{\mathcal{D}}(x^k, E) = \mathcal{D}(E_1^k)$. Let φ be in this set. Since now $E_1\varphi \in \mathcal{D}(E_1^{k-1})$, the induction assumption (along with the fact that $AV\varphi = V(E_1\varphi)$) gives

$$A^k V\varphi = A^{k-1}(AV\varphi) = A^{k-1}V(E_1\varphi) = VE_1^{k-1}(E_1\varphi) = VE_1^k\varphi, \quad (8.21)$$

completing the induction proof of (8.17).

Let $n \in \mathbb{N}_0$. Now (8.16) and (8.17) give $\tilde{L}(x^n, E)\varphi = V^*A^nV\varphi = V^*VE_1^n\varphi = E_1^n\varphi$ for all $\varphi \in \tilde{\mathcal{D}}(x^n, E) = \mathcal{D}(E_1^n)$, so $\tilde{L}(x^n, E) = \tilde{L}(x, E)^n$. This proves (a).

If we assume that $\tilde{L}(x, E)$ is selfadjoint, (8.20) entails that $P_{\mathcal{H}}\mathcal{D}(A) \subset \mathcal{D}(A)$. Indeed, let $\psi \in \mathcal{D}(A)$, and let $\varphi \in \mathcal{D}(E_1)$ be arbitrary. Using (8.20), we get

$$\langle E_1\varphi | V^*\psi \rangle = \langle VE_1\varphi | \psi \rangle = \langle AV\varphi | \psi \rangle = \langle \varphi | V^*A\psi \rangle, \quad (8.22)$$

which implies that $V^*\psi \in \mathcal{D}(E_1^*)$. Since E_1 is selfadjoint, $V^*\psi \in \mathcal{D}(E_1)$, so $P_{\mathcal{H}}\psi = V(V^*\psi) \in V\mathcal{D}(E_1)$. But $V\mathcal{D}(E_1)$ is contained in $\mathcal{D}(A)$, because $\mathcal{D}(E_1) = \mathcal{D}(AV)$. Thus $P_{\mathcal{H}}\psi \in \mathcal{D}(A)$, proving the fact that $P_{\mathcal{H}}\mathcal{D}(A) \subset \mathcal{D}(A)$. In addition, the above calculation shows that $V^*\mathcal{D}(A) \subset \mathcal{D}(E_1)$, and $E_1V^*\psi = E_1^*(V^*\psi) = V^*A\psi$ for all $\psi \in \mathcal{D}(A)$. Combining this with (8.20), we get

$$P_{\mathcal{H}}A\psi = VV^*A\psi = V(E_1V^*\psi) = VE_1(V^*\psi) = AV(V^*\psi) = AP_{\mathcal{H}}\psi \quad (8.23)$$

for all $\psi \in \mathcal{D}(A)$. Consequently, $P_{\mathcal{H}}A \subset AP_{\mathcal{H}}$. Since A is selfadjoint, this implies that $P_{\mathcal{H}}$ commutes with all the spectral projections $F(X)$ [16, pp. 320, 301]. It follows from Corollary 8.4 that each $E(X)$ is a projection. $\square$

Together with Theorem 5.2 the previous proposition leads to the following characterisation of spectral measures.

Theorem 8.5 *Let $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure, such that $\tilde{L}(x, E)$ is selfadjoint. Then the following conditions are equivalent.*

- (i) E is a spectral measure;
- (ii) $L(x^2, E) = L(x, E)^2$;
- (iii) $\int x^2 dE_{\varphi, \varphi} = \|L(x, E)\varphi\|^2$ for all $\varphi \in \tilde{\mathcal{D}}(x, E)$.

Note that the (intrinsic) noise operator $L(x^2, E) - L(x, E)^2$ related to a semispectral measure E is zero exactly when E is a spectral measure. Indeed, it is always positive as the following lemma shows.

Lemma 8.3 *Let $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure. Then $L(x^2, E) \geq L(x, E)^2$.*

Proof Let $(\mathcal{K}, \mathbf{F}, V)$ be a Naimark dilation of $\mathbf{E}$. Then for any $\varphi \in \mathcal{H}$ we have

$$\begin{aligned} \langle \varphi | L(x^2, \mathbf{E}) \varphi \rangle &= \int x^2 \langle V\varphi | \mathbf{F}(dx) V\varphi \rangle = \left\langle V\varphi \left| \left[\int x \mathbf{F}(dx) \right]^2 V\varphi \right\rangle \right. \\ &= \left\langle \int x \mathbf{F}(dx) V\varphi \left| \int x \mathbf{F}(dx) V\varphi \right\rangle \right. \\ &\geq \left\langle \int x \mathbf{F}(dx) V\varphi \left| V V^* \int x \mathbf{F}(dx) V\varphi \right\rangle = \langle \varphi | L(x, \mathbf{E})^2 \varphi \rangle, \end{aligned}$$

which shows that $L(x^2, \mathbf{E}) \geq L(x, \mathbf{E})^2$. $\square$

8.6 Semispectral Measures and Direct Integral Hilbert Spaces

Throughout this section $\mathcal{H}$ is a separable Hilbert space, $\mathcal{A}$ a σ -algebra of subsets of a nonempty set Ω , and $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ a semispectral measure. Let $S \in \mathcal{T}(\mathcal{H})$ be a positive trace-1 operator with all eigenvalues strictly positive, i.e. $S = \sum_{n=1}^{\dim \mathcal{H}} s_n |\eta_n\rangle\langle \eta_n|$ where (η_n) is an orthonormal basis of $\mathcal{H}$, $s_n > 0$ for all n , and $\sum_n s_n = 1$. Define a probability measure $\mu : \mathcal{A} \rightarrow [0, 1]$ by $\mu(X) = \text{tr}[SE(X)]$. For any $T \in \mathcal{T}(\mathcal{H})$ the complex measure $X \mapsto \text{tr}[TE(X)]$ is absolutely continuous with respect to μ so that $\mu(X) = 0$ implies $\mathbf{E}(X) = 0$, i.e. $\mathbf{E}$ is *absolutely continuous with respect to* μ . Obviously, also $\mathbf{E}(X) = 0$ yields $\mu(X) = 0$ showing that $\mathbf{E}$ and μ are *mutually absolutely continuous*. In the following, μ can be replaced by any σ -finite positive measure such that μ (and $\mathbf{E}$) is absolutely continuous with respect to it.

Let $(\varphi_n)_{n=1}^{\dim \mathcal{H}}$ be an orthonormal basis of $\mathcal{H}$ and choose $T = |\varphi_m\rangle\langle \varphi_n|$, so that

$$\langle \varphi_n | \mathbf{E}(X) \varphi_m \rangle = \int_X c_{nm}(\omega) d\mu(\omega)$$

and one may write

$$\langle \varphi | \mathbf{E}(X) \psi \rangle = \sum_{n,m=1}^{\dim \mathcal{H}} \int_X c_{nm}(\omega) d\mu(\omega) \langle \varphi | \varphi_n \rangle \langle \varphi_m | \psi \rangle$$

for all $\varphi, \psi \in \mathcal{H}$. The functions (Radon-Nikodým derivatives) $\Omega \ni \omega \mapsto c_{nm}(\omega) \in \mathbb{C}$ are μ -integrable and, for all $\omega \in \Omega$, the matrix $(c_{nm}(\omega))_{n,m=1}^{\dim \mathcal{H}}$ can be chosen to be positive semidefinite due to the positivity of $\mathbf{E}$ (see [17]). Recall that a complex matrix $(c_{nm})_{n,m \in \mathbb{Z}}$, $\mathbb{Z} \subset \mathbb{Z}$, is *positive semidefinite* if $\sum_{n,m \in F} \overline{d_n} c_{nm} d_m \geq 0$ for any finite subset $F \subset \mathbb{Z}$ and any function $n \mapsto d_n \in \mathbb{C}$ on F . It is known (exercise) that any such matrix has a *Kolmogorov decomposition*, i.e., there exist vectors η_n in some Hilbert space $\mathcal{K}$ such that $c_{nm} = \langle \eta_n | \eta_m \rangle$. Clearly, this decomposition can be chosen

to be *minimal* in the sense that the vectors η_n span $\mathcal{K}$ and $\dim \mathcal{K} \leq \#Z$. In particular, we get

$$c_{nm}(\omega) = \langle \eta_n(\omega) | \eta_m(\omega) \rangle$$

for all n, m and $\omega \in \Omega$, where $\eta_n(\omega) \in \mathcal{H}$.

Let $\mathcal{H}_0 = \{0\}$, $\mathcal{H}_r = \text{lin}\{\varphi_k \mid 1 \leq k \leq r\}$ for all $1 \leq r < \dim \mathcal{H} + 1$, and $\mathcal{H}_\infty = \mathcal{H}$ if $\dim \mathcal{H} = \infty$. It can be shown [17, 18] that the maps $\eta_n : \omega \mapsto \eta_n(\omega)$ can be chosen in the following way:

1. for all $1 \leq n < \dim \mathcal{H} + 1$, the mapping $\eta_n : \Omega \rightarrow \mathcal{H}$ is weakly μ -measurable, i.e., $\omega \mapsto \langle \psi | \eta_n(\omega) \rangle$ is μ -measurable for all $\psi \in \mathcal{H}$;
2. there exists a μ -measurable map $m : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ such that $m(\omega) \leq \dim \mathcal{H}$ and $\eta_n(\omega) \in \mathcal{H}_{m(\omega)}$ for all $1 \leq n < \dim \mathcal{H} + 1$, $\omega \in \Omega$;
3. the set of linear combinations of the vectors $\chi_X \eta_n$, $X \in \mathcal{A}$, $n < \dim \mathcal{H} + 1$, is dense in $\int_\Omega^\oplus \mathcal{H}_{m(\omega)} d\mu(\omega)$.

We let $L^2(\Omega, \mu, \mathcal{H})$ denote the space of the (equivalence classes of) weakly measurable functions $f : \Omega \rightarrow \mathcal{H}$ for which the function $\|f(\cdot)\|^2$ is integrable (it follows from the separability of $\mathcal{H}$ that it is measurable, see e.g. [8]). Equipped with the obvious inner product, this space is (by essentially the same proof as in the scalar case) a Hilbert space. Analogously to Remark 2.6, there is a natural isometric isomorphism from $L^2(\Omega, \mu, \mathcal{H})$ onto the Hilbert tensor product $L^2(\Omega, \mu) \otimes \mathcal{H}$. This isomorphism can be built by starting from simple functions and utilising σ -finiteness. We leave the details as an exercise. Using this notation, in the preceding discussion $\int_\Omega^\oplus \mathcal{H}_{m(\omega)} d\mu(\omega) \subset L^2(\Omega, \mu, \mathcal{H}) \cong L^2(\Omega, \mu) \otimes \mathcal{H}$ is the direct integral of the Hilbert spaces $\mathcal{H}_{m(\omega)}$ consisting of elements $\psi \in L^2(\Omega, \mu, \mathcal{H})$ such that $\psi(\omega) \in \mathcal{H}_{m(\omega)}$ for (μ -almost) all $\omega \in \Omega$. Note that $\int_\Omega^\oplus \mathcal{H}_{m(\omega)} d\mu(\omega)$ is a closed subspace of $L^2(\Omega, \mu, \mathcal{H})$ since any L^2 -convergent sequence has an almost everywhere convergent subsequence.

Denote $\mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}_{m(\omega)} d\mu(\omega)$ and define an isometry $V_\oplus : \mathcal{H} \rightarrow \mathcal{H}_\oplus$ by $V_\oplus = \sum_{n=1}^{\dim \mathcal{H}} |\eta_n\rangle \langle \varphi_n|$. Let $\mathbf{M} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H}_\oplus)$ be the *canonical spectral measure* of $\mathcal{H}_\oplus$, that is, for all $\psi \in \mathcal{H}_\oplus$, $(\mathbf{M}(X)\psi)(\omega) = \chi_X(\omega)\psi(\omega)$ for μ -almost all $\omega \in \Omega$. Now $\mathbf{E}(X) = V_\oplus^* \mathbf{M}(X) V_\oplus$. From item 3 above, we see that $(\mathcal{H}_\oplus, \mathbf{M}, V_\oplus)$ is a minimal Naimark dilation for $\mathbf{E}$.

Definition 8.1 We say that $(\mathcal{H}_\oplus, \mathbf{M}, V_\oplus)$ is a *minimal diagonal Naimark dilation* for $\mathbf{E}$ and the expression

$$\mathbf{E}(X) = \sum_{n,m=1}^{\dim \mathcal{H}} \int_X \langle \eta_n(\omega) | \eta_m(\omega) \rangle d\mu(\omega) |\varphi_n\rangle \langle \varphi_m|$$

is a *minimal (diagonal) Kolmogorov decomposition* of $\mathbf{E}$. If for some $r \in \mathbb{N} \cup \{\infty\}$, $r \leq \dim \mathcal{H}$, $m(\omega) = r$ for μ -almost all $\omega \in \Omega$, the semispectral measure $\mathbf{E}$ is of *rank* r .

Note that if $\mathbf{E}$ is of rank r then $\mathcal{H}_\oplus$ is isomorphich to $L^2(\Omega_+, \mu_+, \mathcal{H}_r) \cong L^2(\Omega_+, \mu_+) \otimes \mathcal{H}_r$ where $\Omega_+ = \{\omega \in \Omega \mid m(\omega) > 0\}$ is μ -measurable and μ_+ is the restriction of μ to Ω_+ , i.e. $\mu_+(X) = \mu(X \cap \Omega_+)$ for all $X \in \mathcal{A}$. If $\mathbf{E}$ and μ are mutually absolutely continuous, $\mu(\Omega \setminus \Omega_+) = 0$ and thus $L^2(\Omega_+, \mu_+) \cong L^2(\Omega, \mu)$.

Let then $\mu : \mathcal{A} \rightarrow [0, \infty]$ be a σ -finite measure and $\mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}_{m(\omega)} d\mu(\omega) \subset L^2(\Omega, \mu, \mathcal{H})$ a direct integral Hilbert space. Let $D(\omega) \in \mathcal{L}(\mathcal{H}_{m(\omega)})$ for μ -almost all $\omega \in \Omega$ and assume that the maps

$$\Omega \ni \omega \mapsto \langle \varphi | D(\omega) \psi \rangle \in \mathbb{C}$$

are μ -measurable for all $\varphi, \psi \in \mathcal{H}$. We say that $\{D(\omega)\}_{\omega \in \Omega}$ is a μ -measurable family of operators. Any such family defines a *decomposable operator*

$$D = \int_\Omega^\oplus D(\omega) d\mu(\omega)$$

which acts as

$$(D\varphi)(\omega) = D(\omega)\varphi(\omega)$$

for all $\varphi \in \mathcal{D}(D) = \{\varphi \in \mathcal{H}_\oplus \mid D\varphi \in \mathcal{H}_\oplus\}$. It is easy to show that a decomposable operator D is bounded if and only if its norm $\|D\| = \text{ess sup}_{\omega \in \Omega} \|D(\omega)\|$ is finite. Especially, for any μ -measurable $f : \Omega \rightarrow \mathbb{C}$ we define the multiplication operator M_f by the formula $(M_f\varphi)(\omega) = f(\omega)\varphi(\omega)$, $\omega \in \Omega$, $\varphi \in \mathcal{D}(M_f) = \{\varphi \in \mathcal{H}_\oplus \mid f\varphi \in \mathcal{H}_\oplus\}$. The operator M_f is clearly decomposable. Similarly as before, the characteristic function χ_X of the set $X \in \mathcal{A}$ defines a projection operator $\mathbf{M}(X) = M_{\chi_X}$ and the canonical spectral measure $\mathbf{M}$ of $\mathcal{H}_\oplus$. Moreover, M_f is the operator integral $L(f, \mathbf{M}) = \int f d\mathbf{M}$, with the dense domain $\mathcal{D}(M_f) = \tilde{\mathcal{D}}(f, \mathbf{M}) = \mathcal{D}(f, \mathbf{M})$ and its adjoint is $M_f^* = M_{\bar{f}}$. If $f \in L^\infty(\mu)$, then M_f is a bounded operator with the operator norm $\|M_f\| = \text{ess sup}_{\omega \in \Omega} |f(\omega)|$.

If $m(\omega) \in \{0, 1\}$ for μ -almost all $\omega \in \Omega$ any decomposable operator is of the form M_f for some measurable function f . Indeed, then $\mathcal{H}_{m(\omega)} \cong \mathbb{C}$ (or $\{0\}$) and any decomposable D consists of operators $D(\omega)$ which can be identified with complex numbers $f(\omega)$. Proposition 8.1 has the following generalisation (for the proof, see [19, Theorem 1, p. 187]):

Proposition 8.11 *A bounded operator $S \in \mathcal{L}(\mathcal{H}_\oplus)$ commutes with every M_f , $f \in L^\infty(\Omega, \mu)$, if and only if S is decomposable.*

Note that if S above commutes with every $\mathbf{M}(X)$, $X \in \mathcal{A}$, then it commutes with any $M_f = \int f d\mathbf{M}$ and is thus decomposable.

Let $\mathbf{E}$ be a rank-1 semispectral measure and $(\mathcal{H}_\oplus, \mathbf{M}, V_\oplus)$ its minimal diagonal Naimark dilation (i.e. $\mathcal{H}_\oplus \cong L^2(\Omega, \mu)$). Now we can generalise Corollary 8.1 and show that $\mathbf{M}$ is *maximal* in the following sense:

Corollary 8.7 *If P is a projection in $\mathcal{L}(L^2(\Omega, \mu))$ commuting with every $\mathbf{M}(X)$, $X \in \mathcal{A}$, then $P = \mathbf{M}(Y)$ for some $Y \in \mathcal{A}$.*

Proof Since P commutes with $\mathbf{M}$ it is decomposable and determined by the projections $P(\omega)$ on one-dimensional Hilbert spaces. Hence, we can identify any $P(\omega)$ with 0 or 1. Now the μ -measurable set $Y' = \{\omega \in \Omega \mid P(\omega) = 1\}$ coincides (almost everywhere) with a $Y \in \mathcal{A}$ so that $\mathbf{M}(Y) = P$ and the corollary follows. $\square$

Remark 8.3 Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure with the minimal diagonal Naimark dilation $(\mathcal{H}_\oplus, \mathbf{M}, V_\oplus)$. Let $\mathfrak{M} \in \text{Ins}(\mathcal{A}, \mathcal{H}', \mathcal{H})$ be a completely positive instrument.

From Theorem 7.12 one sees that $\mathfrak{M}(X, I_{\mathcal{H}'}) = \mathbf{E}(X)$ for all $X \in \mathcal{A}$ if and only if there exists a (unique) channel $\Phi : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H}_\oplus)$ such that, for all $B \in \mathcal{L}(\mathcal{H}')$, $X \in \mathcal{A}$, $\mathfrak{M}(X, B) = V_\oplus^* \mathbf{M}(X) \Phi(B) V_\oplus$ and $\Phi(B) \mathbf{M}(X) = \mathbf{M}(X) \Phi(B)$. The last condition and Proposition 8.11 imply that, for any fixed B , the operator $\Phi(B)$ is decomposable. Indeed, one can show [20] that

$$\Phi(B) = \int_{\Omega}^{\oplus} [\Phi(\omega)](B) d\mu(\omega)$$

for all $B \in \mathcal{L}(\mathcal{H}')$ where $\Phi(\omega) : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H}_{m(\omega)})$ is a channel for μ -almost all $\omega \in \Omega$. Hence,

$$\begin{aligned} \mathfrak{M}(X, B) &= V_\oplus^* \mathbf{M}(X) \Phi(B) V_\oplus \\ &= \sum_{n,m=1}^{\dim \mathcal{H}} \int_X \langle \eta_n(\omega) | [\Phi(\omega)](B) \eta_m(\omega) \rangle d\mu(x) |\varphi_n\rangle \langle \varphi_m|. \end{aligned}$$

Especially, if $\mathbf{E}$ is of rank 1 (i.e. $m(\omega) \in \{0, 1\}$) then $\Phi(\omega)$ can be identified with a trace-1 positive operator $\rho(\omega) \in \mathcal{T}(\mathcal{H}')$ such that $[\Phi(\omega)](B) = \text{tr}[\rho(\omega)B]$ and (weakly) $\mathfrak{M}(X, B) = \int_X \text{tr}[\rho(\omega)B] d\mathbf{E}(\omega)$, i.e. $\mathfrak{M}$ is called *nuclear* or *preparative*. Clearly, any semispectral measure $\mathbf{E}$ has an infinite number of $\mathbf{E}$ -compatible nuclear instruments but if $\mathbf{E}$ is rank-1 then all $\mathbf{E}$ -compatible (completely positive) instruments are nuclear.

Suppose then that $\mathcal{H}' = \mathcal{H}$. Since any $\mathcal{H}_{m(\omega)}$ is a subspace of $\mathcal{H}$ one can define a μ -measurable family of isometries $C(\omega) : \mathcal{H}_{m(\omega)} \rightarrow \mathcal{H}$ and (rank-1) channels $\Phi(\omega) : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}_{m(\omega)})$ by

$$[\Phi(\omega)](B) = C(\omega)^* B C(\omega), \quad B \in \mathcal{L}(\mathcal{H}).$$

Now the corresponding instrument $\mathfrak{M}_C$ is

$$\mathfrak{M}_C(X, B) = \sum_{n,m=1}^{\dim \mathcal{H}} \int_X \langle \eta_n(\omega) | C(\omega)^* B C(\omega) \eta_m(\omega) \rangle d\mu(x) |\varphi_n\rangle \langle \varphi_m|.$$

Let $\mathbf{M} : \mathcal{A} \mapsto \mathcal{L}(L^2(\mu))$ be the canonical spectral measure of $L^2(\mu) = L^2(\Omega, \mu)$ and fix some unit vector $\phi \in L^2(\mu)$. Define a unitary operator $U_\phi \in \mathcal{L}(\mathcal{H} \otimes L^2(\mu))$ by extending

$$[U_\phi(\varphi_m \otimes \phi)](\omega) = C(\omega)\eta_m(\omega) \otimes 1.$$

It can be shown that such a U_ϕ exists if $\mathbf{E}$ is not of rank ∞ [21]. If $\mathbf{E}$ is rank- ∞ (i.e. $m(\omega) = \infty = \dim \mathcal{H}$ for μ -a.a. ω) then the isometries $C(\omega)$ can be chosen in such a way that U_ϕ exists (or alternatively one can add an extra dimension to $L^2(\mu)$).

If $V_\phi : \mathcal{H} \rightarrow \mathcal{H} \otimes L^2(\mu)$ denotes the embedding $V_\phi(\varphi) = \varphi \otimes \phi$ one gets

$$\begin{aligned} \langle \varphi_n | V_\phi^* U_\phi^* [B \otimes \mathbf{M}(X)] U_\phi V_\phi \varphi_m \rangle &= \langle U_\phi(\varphi_n \otimes \phi) | [B \otimes \mathbf{M}(X)] U_\phi(\varphi_m \otimes \phi) \rangle \\ &= \int_{\Omega} \langle C(\omega)\eta_n(\omega) \otimes 1 | [B \otimes \chi_X(\omega)] C(\omega)\eta_m(\omega) \otimes 1 \rangle d\mu(\omega) \\ &= \int_X \langle \eta_n(\omega) | C(\omega)^* B C(\omega) \eta_m(\omega) \rangle d\mu(x) = \langle \varphi_n | \mathfrak{M}_C(X, B) \varphi_m \rangle, \end{aligned}$$

so that

$$\mathfrak{M}_C(X, B) = V_\phi^* U_\phi^* [B \otimes \mathbf{M}(X)] U_\phi V_\phi,$$

and the quadruple $(L^2(\mu), \phi, \mathbf{M}, U_\phi)$ is a measurement dilation of $\mathbf{E}$. We call it a *minimal measurement dilation* of $\mathbf{E}$. Indeed, it can be shown [20] that for any measurement dilation $(\mathcal{H}_o, \phi', \mathbf{P}, U)$ of $\mathbf{E}$ there exists an isometry $W : L^2(\mu) \rightarrow \mathcal{H}_o$ such that $W^* \mathbf{P}(X) W = \mathbf{M}(X)$, $X \in \mathcal{A}$, that is, $L^2(\mu)$ is the smallest ancillary or probe Hilbert space which can be used to measure $\mathbf{E}$. Especially, if $L^2(\mu)$ is not separable (i.e. the measure space $(\Omega, \mathcal{A}, \mu)$ has no countable basis) there does not exist a separable ancillary Hilbert space for $\mathbf{E}$. However, this is a quite rare situation in physics, since if $\mathcal{A}$ is countably generated (e.g. $\mathcal{B}(\mathbb{R}^n)$) then $L^2(\mu)$ is separable. $\triangleleft$

8.7 A Dirac Type Formalism: An Elementary Approach

Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a semispectral measure and

$$\mathbf{E}(X) = \sum_{n,m=1}^{\dim \mathcal{H}} \int_X \langle \eta_n(\omega) | \eta_m(\omega) \rangle d\mu(\omega) |\varphi_n\rangle \langle \varphi_m|$$

its minimal Kolmogorov decomposition associated with $\mathcal{H}_\oplus = \int_{\Omega}^{\oplus} \mathcal{H}_{m(\omega)} d\mu(\omega)$. Let V be the (dense) linear space of finite linear combinations of the basis vectors φ_n and let $V^\times$ be its algebraic antidual, i.e. the set of conjugate linear maps $V \rightarrow \mathbb{C}$. One can identify $V^\times$ with the space of the formal series $v = \sum_{n=1}^{\dim \mathcal{H}} v_n \varphi_n$ where v_n 's are arbitrary complex numbers, and then the dual pairing is

$$\langle \psi | v \rangle = \sum_{n=1}^{\dim \mathcal{H}} \langle \psi | \varphi_n \rangle v_n = \overline{\langle v | \psi \rangle}$$

the sum being finite for all $\psi \in V$ and $v \in V^\times$. Hence, $v_n = \langle \varphi_n | v \rangle$.

Define (weakly) μ -measurable maps $v_k : \Omega \rightarrow V^\times$ by

$$\Omega \ni \omega \mapsto \langle \varphi_n | v_k(\omega) \rangle = \langle \eta_n(\omega) | \varphi_k \rangle \in \mathbb{C}$$

for all $1 \leq n < \dim \mathcal{H} + 1$ and $1 \leq k < m(\omega) + 1$. Thus,

$$\langle \varphi | \mathbf{E}(X) \psi \rangle = \int_X \sum_{k=1}^{m(\omega)} \langle \varphi | v_k(\omega) \rangle \langle v_k(\omega) | \psi \rangle d\mu(\omega)$$

for all $\varphi, \psi \in V$, which we write briefly as follows

$$\mathbf{E}(X) = \int_X \sum_{k=1}^{m(\omega)} |v_k(\omega)\rangle \langle v_k(\omega)| d\mu(\omega) \quad (8.24)$$

when considering both sides as sesquilinear forms on V . It can be shown that, for μ -almost all $\omega \in \Omega$, the vectors $v_k(\omega) \in V^\times$, $1 \leq k < m(\omega) + 1$, are linearly independent [17]. We say that (8.24) is a *minimal diagonalisation of a semispectral measure* $\mathbf{E}$, and that $m(\omega)$ is the *multiplicity of the outcome ω of $\mathbf{E}$* . Especially, if $\mathbf{E}$ is of rank-1 then $\mathbf{E}(X) = \int_X |v_1(\omega)\rangle \langle v_1(\omega)| d\mu(\omega)$. Examples of rank-1 semispectral measures are position and momentum spectral measures of a nonrelativistic spin-0 particle confined to move on a line, the canonical phase Φ_{can} , and a phase space semispectral measure $\mathbf{G}^\varphi$ generated by a one-dimensional projection $P[\varphi]$. Indeed, it is an easy exercise to show that $\mathbf{G}^T$ is a rank- r semispectral measure where r is the rank of the generating trace-1 positive operator T .

If $\mathbf{E}$ is a spectral measure then the vectors η_n form an orthonormal basis of $\mathcal{H}_\oplus$; in the Dirac formalism this is written *formally* in the form $\langle v_k(\omega) | v_l(\omega') \rangle = \delta_{kl} \delta_\omega(\omega')$ where $\delta_\omega(\omega')$ is ‘Dirac’s delta’ concentrated at ω . In the general case of an arbitrary outcome set Ω , Dirac’s delta is just a formal notation.

If $\mathbf{E} : \mathcal{B}(\mathbb{C}) \rightarrow \mathcal{L}(\mathcal{H})$ is the spectral measure of a normal operator $A : \mathcal{D}(A) \rightarrow \mathcal{L}(\mathcal{H})$, one can always choose the basis $\{\varphi_n\}$ and hence the corresponding V in such a way that $V \subset \mathcal{D}(A)$, $V \subset \mathcal{D}(A^*)$, $AV \subset V$, and $A^*V \subset V$. This means that the matrices $(\langle \varphi_n | A \varphi_m \rangle)$ and $(\langle \varphi_n | A^* \varphi_m \rangle)$ are both row and column finite and they can be uniquely extended to linear maps $V^\times \rightarrow V^\times$; we use the same symbols A and A^* for these extensions.

Definition 8.2 If $v \in V^\times \setminus \{0\}$ and $x \in \mathbb{C}$ satisfy $Av = xv$ then v is called a *generalised eigenvector* of A belonging to the *generalised eigenvalue* x of A (relative to V).

The following theorem is proved in [17].

Theorem 8.6 *Let A be a normal operator with the definitions as above. Then there exist a Borel measure $\mu : \mathcal{B}(\sigma(A)) \rightarrow [0, \infty)$ and (μ -measurable) maps $m : \sigma(A) \rightarrow \mathbb{N}_0 \cup \{\infty\}$ and $v_k : \sigma(A) \rightarrow V^\times$, $k \in \mathbb{N}$, such that, for μ -almost all*

$x \in \sigma(A)$ and $1 \leq k < m(x) + 1$ the vector $v_k(x) \in V^\times$ is either zero or a simultaneous generalised eigenvector of A and A^* relative to V belonging to their generalised eigenvalues x and $\bar{x}$, respectively (i.e. $Av_k(x) = xv_k(x)$ and $A^*v_k(x) = \bar{x}v_k(x)$). Moreover, (as sesquilinear forms $V \times V \rightarrow \mathbb{C}$)

$$A = \int_{\sigma(A)} x \sum_{k=1}^{m(x)} |v_k(x)\rangle \langle v_k(x)| d\mu(x).$$

The above theorem forms the core of the Dirac formalism and can effectively be used to determine the spectral measure of a normal operator. Note that a generalised eigenvalue does not necessarily belong to the spectrum $\sigma(A)$ of A [17, Sect. 9].

Example 8.1 Let C be the cosine operator, i.e. $C = L(\cos, \Phi_{\text{can}})$ where Φ_{can} is the canonical phase. Now

$$\begin{aligned} C &= \sum_{n,m=0}^{\infty} \int_0^{2\pi} (\cos x) e^{i(n-m)x} \frac{dx}{2\pi} |e_n\rangle \langle e_m| \\ &= \sum_{n,m=0}^{\infty} \frac{1}{2} (\delta_{m,n+1} + \delta_{n,m+1}) |e_n\rangle \langle e_m| \\ &= \frac{1}{2} \sum_{n=0}^{\infty} (|e_n\rangle \langle e_{n+1}| + |e_{n+1}\rangle \langle e_n|). \end{aligned}$$

Immediately one finds that

$$Cv(x) = \cos x v(x)$$

where

$$v(x) = c \sum_{n=0}^{\infty} [\sin(n+1)x] e_n \in V^\times$$

and $c \in \mathbb{R}$. Since (as a sesquilinear form $V \times V \rightarrow \mathbb{C}$)

$$\begin{aligned} \int_0^\pi |v(x)\rangle \langle v(x)| dx &= c^2 \sum_{n,m=0}^{\infty} \int_0^\pi \sin(n+1)x \sin(m+1)x dx |e_n\rangle \langle e_m| \\ &= c^2 \frac{\pi}{2} \sum_{n,m=0}^{\infty} \delta_{nm} |e_n\rangle \langle e_m| \end{aligned}$$

one must choose $c = \sqrt{2/\pi}$ to get a (rank-1) semispectral measure

$$E'_C : \mathcal{B}([0, \pi]) \rightarrow \mathcal{L}(\mathcal{H}),$$

$$X \mapsto E'_C(X) = \frac{2}{\pi} \sum_{n,m=0}^{\infty} \int_X \sin(n+1)x \sin(m+1)x dx |e_n\rangle \langle e_m|$$

for which

$$\int_0^\pi \cos x \, d\mathbf{E}'_C(x) = C.$$

Since

$$\int_0^\pi \cos^2 x \, d\mathbf{E}'_C(x) = C^2,$$

by Theorem 8.5, we have obtained the spectral measure $\mathbf{E}_C : \mathcal{B}([-1, 1]) \rightarrow \mathcal{L}(\mathcal{H})$ of C :

$$\mathbf{E}_C(X) = \mathbf{E}'_C(\arccos X).$$

(Recall that the spectrum of C is $[-1, 1]$ by Theorem 8.2.) Similarly, it is easy to find the spectral measure of the sine operator $L(\sin, \Phi_{\text{can}})$. $\triangleleft$

8.8 Exercises

1. Show that the range $\{\mathbf{E}(X) \mid X \in \mathcal{A}\}$ of any spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a Boolean sub- σ -algebra of the projection lattice $\mathcal{P}(\mathcal{H})$. Hint: Use Proposition 4.6 together with Theorem 2.11 to show that $\mathbf{E}(X \cup Y) = \mathbf{E}(X) \vee \mathbf{E}(Y)$ for any $X, Y \in \mathcal{A}$ and then, for instance, $\mathbf{E}(X) \wedge (\mathbf{E}(Y) \vee \mathbf{E}(Z)) = (\mathbf{E}(X) \wedge \mathbf{E}(Y)) \vee (\mathbf{E}(X) \wedge \mathbf{E}(Z))$ for any $X, Y, Z \in \mathcal{A}$, and similarly the other distributive laws. Recall also that $\mathbf{E}(X^c) = \mathbf{E}(X)^\perp$. The σ -property of the lattice follows from the σ -additivity of $\mathbf{E}$.
2. Consider the restriction of the spectral measure $\mathbf{M}_F : X \mapsto F^{-1}\mathbf{M}(X)F$ in $L^2(\mathbb{R})$ by a projection $P = \mathbf{M}(\mathcal{I})$ for some interval $\mathcal{I} = [a, b] \subset \mathbb{R}$. Show that the Naimark projection of $\mathbf{M}_F$ to $\mathbf{M}(\mathcal{I})(L^2(\mathbb{R})) = L^2(\mathcal{I})$ is not a spectral measure. Denote it by $\widetilde{\mathbf{M}}_F^\mathcal{I}$. Use Theorem 7.10 to study the moment operators of $\widetilde{\mathbf{M}}_F^\mathcal{I}$. Show, in particular, that the operator $\widetilde{\mathbf{M}}_F^\mathcal{I}[1]$ equals the operator $-iD$ on the domain

$$\widetilde{\mathcal{D}}(x, \mathbf{M}_F^\mathcal{I}) = \{\psi \in L^2(\mathcal{I}) \mid \psi \text{ abs. cont. on } \mathcal{I}, \psi' \in L^2(\mathcal{I}), \psi(a) = \psi(b) = 0\}.$$

Show that the operator $\widetilde{\mathbf{M}}_F^\mathcal{I}[1]$ is densely defined and symmetric but not selfadjoint. Show also that the operator $-iD$, with the domain

$$\{\psi \in L^2(\mathcal{I}) \mid \psi \text{ abs. cont. on } \mathcal{I}, \psi' \in L^2(\mathcal{I}), \psi(a) = \psi(b)\}$$

is a selfadjoint extension of $\widetilde{\mathbf{M}}_F^\mathcal{I}[1]$ in $L^2(\mathcal{I})$. Does this operator have other selfadjoint extensions?

3. Calculate the matrix elements of the moment operators $\Phi_{\text{can}}[n]$, $n \in \mathbb{N}_0$, of the canonical phase Φ_{can} in the Fourier basis $(e_n)_{n \geq 0}$.
4. Let Ω_1 be a finite or countably infinite set, $\mathcal{A}_1$ the σ -algebra containing all subsets of Ω_1 , and $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ a semispectral measure. Let $(\Omega_2, \mathcal{A}_2)$ be an arbitrary

- trary measurable space and $p : \mathcal{A}_2 \times \Omega_1 \rightarrow [0, 1]$ a Markov kernel. Calculate the smearing of $\mathbf{E}_1$ by p .
5. Prove Proposition 8.9.
 6. Let $\mathcal{H}$ be separable, $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ a semispectral measure and $S \in \mathcal{T}(\mathcal{H})$ a positive operator with the eigenvalues nonzero. Show then that the complex measures $X \mapsto \text{tr}[TE(X)]$, $T \in \mathcal{T}(\mathcal{H})$, are absolutely continuous with respect to the positive finite measure $\mu(X) = \text{tr}[SE(X)]$. Can one assume that the measure μ related to $\mathbf{E}$ is only σ -finite? (Recall that any σ -finite positive measure is absolutely continuous with respect to a probability measure.)
 7. Show that any positive semidefinite matrix $(c_{nm})_{n,m \in \mathbb{Z}}$, $\mathbb{Z} \subset \mathbb{Z}$, has a minimal Kolmogorov decomposition, i.e., there exist vectors η_n in some Hilbert space $\mathcal{K}$ such that $c_{nm} = \langle \eta_n | \eta_m \rangle$ and η_n -vectors span $\mathcal{K}$. Show that vectors η_n are unique up to a unitary transformation. *Hint:* find positive numbers p_n such that $(\tilde{c}_{nm})$, $\tilde{c}_{nm} = p_n c_{nm} p_m$, becomes a bounded operator and use its square root operator.
 8. Let $\mathbf{G}^T$ be a covariant phase space observable generated by a positive trace-1 operator T . Determine a minimal diagonal Kolmogorov decomposition and the associated minimal Naimark dilation and diagonalisation of $\mathbf{G}^T$. Show that $\mathbf{G}^T$ is of rank r where r is the rank of T .
 9. Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a *discrete* semispectral measure, that is, there exists $\{\omega_i\}_{i=1}^N \subset \Omega$ (where N can be infinite) such that $\omega_i \neq \omega_j$ when $i \neq j$, $\{\omega_i\} \in \mathcal{A}$, $\mathbf{E}(\{\omega_i\}) > 0$, and $\sum_{i=1}^N \mathbf{E}(\{\omega_i\}) = I$. Let μ be related to $\mathbf{E}$ as before. By using the fact that $L^2(\mu)$ is the smallest ancillary Hilbert space for measuring $\mathbf{E}$, show that, to measure $\mathbf{E}$ one needs at least an N dimensional ancillary Hilbert space. What is the pointer observable $\mathbf{M}$ then? Construct also some measurement interaction U_ϕ .
 10. Calculate $L(\sin, \Phi_{\text{can}})$ and determine its spectral measure by using generalised eigenvectors. Does $L(\sin, \Phi_{\text{can}})$ have ordinary eigenvectors (belonging to the Hilbert space) or eigenvalues?

References

1. Sikorski, R.: On the inducing of homomorphisms by mappings. *Fundam. Math.* **36**, 7–22 (1949)
2. Akhiezer, N.I., Glazman, I.M.: *Theory of Linear Operators in Hilbert Space*. Dover Publications Inc., New York (1993) (Translated from the Russian and with a preface by Merlynd Nestell, Reprint of the 1961 and 1963 translations, Two volumes bound as one)
3. Varadarajan, V.S.: *Geometry of Quantum Theory*, vol. 2. Springer, New York (1985)
4. Taylor, B., Mohr, P.: *The NIST Reference on Constants, Units, and Uncertainty*. <http://physics.nist.gov/cuu/index.html>
5. Douglas, R.G.: *Banach Algebra Techniques in Operator Theory*. Pure and Applied Mathematics, vol. 49, Academic Press, New York (1972)
6. Holevo, A.S.: *Statistical Structure of Quantum Theory*. Lecture Notes in Physics, vol. 67. Monographs. Springer, Berlin (2001)
7. Halmos, P.R.: *Measure Theory*, 4th edn. Springer, Berlin (1988) (Reprint of the 1950 edition)
8. Dunford, N., Schwartz, J.T.: *Linear Operators*. Part I. Wiley Classics Library. Wiley, New York (1988) (Reprint of the 1958 original, A Wiley-Interscience Publication)

9. Kiukas, J., Lahti, P., Ylinen, K.: Semispectral measures as convolutions and their moment operators. *J. Math. Phys.* **49**(11), 112103, 6 (2008)
10. Carmeli, C., Heinonen, T., Toigo, A.: Position and momentum observables on $\mathbb{R}$ and on $\mathbb{R}^3$. *J. Math. Phys.* **45**(6), 2526–2539 (2004)
11. Kiukas, J., Lahti, P., Ylinen, K.: Normal covariant quantization maps. *J. Math. Anal. Appl.* **319**(2), 783–801 (2006)
12. Holevo, A.S.: Covariant measurements and uncertainty relations. *Rep. Math. Phys.* **16**(3), 385–400 (1979)
13. Werner, R.: Quantum harmonic analysis on phase space. *J. Math. Phys.* **25**(5), 1404–1411 (1984)
14. Cassinelli, G., De Vito, E., Toigo, A.: Positive operator valued measures covariant with respect to an irreducible representation. *J. Math. Phys.* **44**(10), 4768–4775 (2003)
15. Kiukas, J., Lahti, P., Ylinen, K.: Phase space quantization and the operator moment problem. *J. Math. Phys.* **47**(7), 072104, 18 (2006)
16. Riesz, F., Nagy, B.S.: *Functional Analysis*. Dover Books on Advanced Mathematics. Dover Publications Inc., New York (1990) (Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original)
17. Hytönen, T., Pellonpää, J.-P., Ylinen, K.: Positive sesquilinear form measures and generalized eigenvalue expansions. *J. Math. Anal. Appl.* **336**(2), 1287–1304 (2007)
18. Pellonpää, J.-P.: Complete characterization of extreme quantum observables in infinite dimensions. *J. Phys. A* **44**(8), 085304, 12 (2011)
19. Dixmier, J.: *Von Neumann Algebras*. North-Holland Mathematical Library, vol. 27. North-Holland Publishing Co., Amsterdam (1981) (Translated from the second French edition by F. Jellet)
20. Pellonpää, J.-P.: Quantum instruments: II. Measurement theory. *J. Phys. A* **46**(2), 025303, 15 (2013)
21. Pellonpää, J.-P., Tukiainen, M.: Minimal normal measurement models of quantum instruments. [arXiv:1509.08886](https://arxiv.org/abs/1509.08886)

Part II

Elements

Chapter 9

States, Effects and Observables

In this chapter we review the basic elements and structures of Hilbert space quantum mechanics. We build on the idea of a statistical duality arising from the analysis of an experiment as a preparation-measurement-registration scheme, as sketched in Sect. 1.2. The description of a physical system $\mathcal{S}$ is thus based on the notions of *states* as equivalence classes of preparations, *observables* as equivalence classes of measurements, and on the *probability measures* for the possible measurement outcomes.

Let $\mathcal{S}$ and $\mathcal{O}$ be the sets of all states and all observables of the system $\mathcal{S}$. The representation of the measurement outcome probabilities $p_s^O(X)$ for states $s \in \mathcal{S}$ and observables $(O, \Omega, \mathcal{A}) \in \mathcal{O}$ leads to the notion of an *effect* as an equivalence class $[(O, X)]$ of measurement registrations, or experimental statements, (O, X) , as described in Sect. 1.2. The set of all effects will be denoted by $\mathcal{E}$.

The statistical completeness built into the notions of states, effects and observables and the statistical duality between them entail that states separate effects, effects separate states, and observables can be identified with the totality of their measurement outcome probabilities.

In the Hilbert space realisation of this description, a complex separable Hilbert space $\mathcal{H}$ is associated with the system $\mathcal{S}$ such that the set of states will be identified with the set $\mathcal{S}(\mathcal{H})$ of positive trace one operators $\varrho : \mathcal{H} \rightarrow \mathcal{H}$, the set of effects as the set $\mathcal{E}(\mathcal{H})$ of positive unit bounded operators $E : \mathcal{H} \rightarrow \mathcal{H}$, and observables as normalised positive operator measures $\mathcal{A} \ni X \mapsto \mathbf{E}(X) \in \mathcal{E}(\mathcal{H})$. The structure of the probability measures $\mathbf{E}_\varrho$ is thereby determined to be of the form $\mathbf{E}_\varrho(X) = \text{tr}[\varrho \mathbf{E}(X)]$.

The identification of the set $\mathcal{S}$ with the set $\mathcal{S}(\mathcal{H})$ and, equivalently, $\mathcal{E}$ with the set $\mathcal{E}(\mathcal{H})$, refers to the fact that the system $\mathcal{S}$ is a proper quantum system in the sense that it has no (nontrivial) classical properties. In Chaps. 22 and 23 we shall briefly return to this assumption but will otherwise use the full identifications $\mathcal{S} = \mathcal{S}(\mathcal{H})$ and $\mathcal{E} = \mathcal{E}(\mathcal{H})$ throughout.

We have thus identified the Hilbert space formulation of the elements of the general probabilistic description $(\mathcal{S}, \mathcal{E}, \mathcal{O})$ of a physical system. In the final Chap. 23 of this

monograph we return to the fundamental question of the justification of this Hilbert space realisation, that is, its derivation from a set of physically motivated postulates.

Section 9.1 describes the basic properties of states together with the main physically relevant structures of the set $\mathcal{S}(\mathcal{H})$, its convex structure, with some characterisations of the extremal states—the pure states, the notion of superposition of pure states, and the nonuniqueness of the decomposition of mixed states.

Section 9.2 explains the main structures of the set of effects $\mathcal{E}(\mathcal{H})$, including its order, convex, and partial sum structures. This enables a characterisation of states as mappings on the set of effects.

Generalising von Neumann’s original argument [1], in Sect. 9.3 it is shown that, under some natural assumptions, an observable, defined as a mapping $\varrho \mapsto \mathbf{E}_\varrho$, is canonically represented as a semispectral measure, that is, a normalised positive operator measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$. This argument constitutes a derivation of the trace formula for the basic probabilities, $\mathbf{E}_\varrho(X) = \text{tr}[\varrho \mathbf{E}(X)]$, on the basis of the notions of state and measurement statistics. The recovery of the special case of observables $(\mathbf{E}, \mathbb{R}, \mathcal{B}(\mathbb{R}))$ as selfadjoint operators acting in $\mathcal{H}$ is also reviewed here. Further we formulate the convex structure of the set of observables $\mathbf{E}$ with a fixed value space $(\Omega, \mathcal{A})$, discussing its significance and providing a characterisation of the extremal observables.

Section 9.4 provides the general tools needed to describe the changes a physical system $\mathcal{S}$ may experience in the course of its time evolution or due to an intervention, such as a measurement. The notion of an operation is introduced there, in the Schrödinger representation, as a positive linear contracting mapping on the state space $\mathcal{T}_s(\mathcal{H})$ of the selfadjoint trace class operators. The important classes of the Wigner and Lüders operations are characterised.

The quantum theory of measurement is an application of the theory of compound systems in quantum mechanics, which is reviewed in Sect. 9.5. We recall the composition rules leading to the Hilbert tensor product structure and conclude the chapter with a brief discussion of the important concepts of subsystem states, dynamics, correlations, and entanglement.

9.1 States

Let $\mathcal{H}$ be the Hilbert space associated with a given quantum system $\mathcal{S}$. A *state* of $\mathcal{S}$ is represented by, and identified with, an element ϱ of $\mathcal{T}(\mathcal{H})$ such that ϱ is positive and of trace one. The set of all states of $\mathcal{S}$ is

$$\mathcal{S}(\mathcal{H}) = \{\varrho \in \mathcal{T}(\mathcal{H}) \mid \varrho \geq 0, \text{tr}[\varrho] = 1\}. \quad (9.1)$$

Occasionally, states are referred to as *density operators*.

If $\varrho_1, \varrho_2 \in \mathcal{S}(\mathcal{H})$ and $0 \leq w \leq 1$, then the convex combination $w\varrho_1 + (1 - w)\varrho_2 \in \mathcal{S}(\mathcal{H})$. Thus $\mathcal{S}(\mathcal{H})$ is a *convex* subset of $\mathcal{T}(\mathcal{H})$. The set of states $\mathcal{S}(\mathcal{H})$ is even *σ -convex*. Indeed, if $(\varrho_i)_{i=1}^\infty$ is a sequence of states and $(w_i)_{i=1}^\infty$ a sequence of weights

($0 \leq w_i \leq 1$, $\sum_{i=1}^{\infty} w_i = 1$), then $(\sum_{i=1}^n w_i \varrho_i) \subset \mathcal{T}(\mathcal{H})^+$ is a Cauchy sequence in $\mathcal{T}(\mathcal{H})$; hence $\lim_{n \rightarrow \infty} \sum_{i=1}^n w_i \varrho_i = \sum_{i=1}^{\infty} w_i \varrho_i \in \mathcal{T}(\mathcal{H})$, with convergence in trace norm. Since the operator $\sum_{i=1}^{\infty} w_i \varrho_i$ is positive, one gets $\|\sum_{i=1}^{\infty} w_i \varrho_i\|_1 = \text{tr}[\sum_{i=1}^{\infty} w_i \varrho_i] = \sum_{i=1}^{\infty} w_i = 1$.

The convex structure of $\mathcal{S}(\mathcal{H})$ reflects the physical possibility of combining states into new states by mixing them with given weights. Thus we say that $\varrho = w\varrho_1 + (1-w)\varrho_2$ is a *mixture of the states* ϱ_1 and ϱ_2 with the weights w , $1-w$. The extreme elements of $\mathcal{S}(\mathcal{H})$ are those states ϱ for which the condition $\varrho = w\varrho_1 + (1-w)\varrho_2$, with $\varrho_1, \varrho_2 \in \mathcal{S}(\mathcal{H})$, $0 < w < 1$, is fulfilled only for $\varrho_1 = \varrho_2 = \varrho$. These are the states, called *pure states*, that cannot be expressed as mixtures of other states. We denote by $\text{ex}(\mathcal{S}(\mathcal{H}))$ the set of pure states.

Theorem 9.1 below gives a characterisation of the pure states as one-dimensional (or rank-1) projections. These are defined as projections onto a one-dimensional subspace. If φ is any nonzero vector in the range of a one-dimensional projection P , then P can be defined as the linear operator $P[\varphi]$ that acts on $\psi \in \mathcal{H}$ as

$$P[\varphi]\psi = \frac{\langle \varphi | \psi \rangle}{\langle \varphi | \varphi \rangle} \varphi.$$

We denote by $\mathbf{P}_1(\mathcal{H})$ the set of one-dimensional projections on $\mathcal{H}$.

By Theorem 3.5 any state $\varrho \in \mathcal{S}(\mathcal{H})$ has a decomposition

$$\varrho = \sum_{i=0}^{\infty} w_i P_i, \quad (9.2)$$

where (P_i) is an orthogonal sequence ($P_i P_j = \delta_{ij} P_i$) of one-dimensional projections, $w_i \in [0, 1]$, $\sum_i w_i = 1$, with the series converging in the operator norm of $\mathcal{L}(\mathcal{H})$ —but also in the trace norm of $\mathcal{T}(\mathcal{H})$. The numbers w_i , $w_i \neq 0$, are the nonzero eigenvalues of ϱ , and each of them occurs in the decomposition as many times as given by the dimension of the corresponding eigenspace.

Theorem 9.1 *For any state $\varrho \in \mathcal{S}(\mathcal{H})$ the following are equivalent:*

- (i) $\varrho \in \text{ex}(\mathcal{S}(\mathcal{H}))$;
- (ii) $\varrho \in \mathbf{P}_1(\mathcal{H})$;
- (iii) $\varrho = \varrho^2$;
- (iv) $\text{tr}[\varrho^2] = 1$;
- (v) $\|\varrho\| = 1$;
- (vi) $S(\varrho) = -\text{tr}[\varrho \ln(\varrho)] = -\sum_i w_i \ln w_i = 0$.

Proof Assume that (i) holds. Then, in the decomposition (9.2), one of the $w_i = 1$ and all others equal zero, hence (ii) holds, $\varrho \in \mathbf{P}_1(\mathcal{H})$; for otherwise, one would have $0 < w_j < 1$ for some j , and so $\varrho = w_j P_j + w \varrho_o$, where $w = \sum_{i \neq j} w_i$ and $\varrho_o = w^{-1} \sum_{i \neq j} w_i P_i$, in contradiction to ϱ being pure.

If (ii), $\varrho \in \mathbf{P}_1(\mathcal{H})$, then clearly $\varrho = \varrho^2$. If (iii) holds, then $\text{tr}[\varrho^2] = \text{tr}[\varrho] = 1$, i.e., (iv) holds. If $\text{tr}[\varrho^2] = \sum_i w_i^2 = 1$, then again, since also $\text{tr}[\varrho] = \sum_i w_i = 1$ and so

$\sum_i w_i(1 - w_i) = 0$, one $w_i = 1$ and the others equal zero; thus ϱ is a one-dimensional projection and $\|\varrho\| = 1$. On the other hand, if $\|\varrho\| = 1$, then 1 is an eigenvalue of ϱ and hence, again, ϱ is a one-dimensional projection (since the only other eigenvalue must be 0 if $\dim \mathcal{H} \neq 1$). Clearly, $S(\varrho) = 0$ if and only if ϱ is a one-dimensional projection. Assume that $\varrho \in \mathbf{P}_1(\mathcal{H})$ and $\varrho = w\varrho_1 + (1 - w)\varrho_2$, where $\varrho_1, \varrho_2 \in \mathbf{S}(\mathcal{H})$ and $0 < w < 1$. Then, for instance, $w\varrho_1 \leq \varrho$, and thus, since $\dim \varrho(\mathcal{H}) = 1$, $\varrho w\varrho_1 = w\varrho_1 = (w\varrho_1)^* = \varrho_1 w\varrho$. This shows that $w\varrho_1$ is a composition of ϱ and a linear map on the one-dimensional space $\varrho(\mathcal{H})$, that is, $w\varrho_1 = \alpha\varrho$ for some $\alpha > 0$ (since $w\varrho_1 \geq 0$). Since $\text{tr}[\varrho_1] = \text{tr}[\varrho] = 1$, we have $\alpha = w$, so that $\varrho_1 = \varrho$. Similarly one shows that $\varrho_2 = \varrho$, so that ϱ is pure. $\square$

By the above theorem, the pure states of $\mathcal{S}$ are exactly the one-dimensional projection operators on $\mathcal{H}$, that is,

$$\text{ex}(\mathcal{S}(\mathcal{H})) = \mathbf{P}_1(\mathcal{H}).$$

Any unit vector $\varphi \in \mathcal{H}$ thus defines a pure state $|\varphi\rangle\langle\varphi| = P[\varphi]$ and any pure state is of this form for some unit vector. Therefore, with a slight abuse of language, we call these states as well as the accompanying unit vectors also the *vector states*.

The set of pure states is nonempty. Moreover, the decomposition (9.2) shows that the vector states generate the whole set of states $\mathcal{S}(\mathcal{H})$ in the sense that any state can be expressed as a finite or countably infinite convex combination of some vector states. In other words, the σ -convex hull of $\mathbf{P}_1(\mathcal{H})$ is the whole set of states $\mathcal{S}(\mathcal{H})$,

$$\sigma\text{-co}(\mathbf{P}_1(\mathcal{H})) = \mathcal{S}(\mathcal{H}), \quad (9.3)$$

where the σ -convex hull can be taken either with respect to the operator norm topology or the trace norm topology.

We will discuss next two important nonclassical features of quantum states: the superpositions of vector states and the nonunique decomposability of mixed states.

Superpositions

It is a basic feature of quantum mechanics that any two (or any finite number of) pure states can be combined into a new pure state by superposing them.¹ To describe this notion, consider the least upper bound $P_1 \vee P_2$ of two distinct pure states P_1 and P_2 . Then any $P \in \mathbf{P}_1(\mathcal{H})$ contained in $P_1 \vee P_2$, that is, $P \leq P_1 \vee P_2$, is a *superposition* of P_1 and P_2 . Conversely, every pure state P can be expressed as a superposition of any other pure state P_1 and another vector state P_2 exactly when P_1 is not orthogonal to P , $P_1 \not\leq P^\perp$, that is, if and only if $1 \neq \text{tr}[PP_1] \neq 0$.

¹The notion of superposition of states goes back to Dirac [2]. Ever since, this notion has been considered as one of the basic principles of quantum mechanics. The superposition of states is generally formulated in terms of the linear combination of state vectors, or wave functions. Here we adopt a slightly more abstract point of view, formulating it in terms of the pure states, i.e., one-dimensional projections. This perspective is needed in the axiomatic context of Chap. 23.

As is well known, the idea of superposition of pure states is most directly expressed using the linear structure of the underlying Hilbert space. Indeed, if $P_1 = P[\varphi_1]$ and $P_2 = P[\varphi_2]$, then the superpositions of P_1 and P_2 are exactly those vector states which are of the form $P = P[c_1\varphi_1 + c_2\varphi_2]$, with $c_1, c_2 \in \mathbb{C}$. If $P = P[\varphi]$ is any vector state and $P_1 = P[\varphi_1]$ is such that $P_1 \not\leq P^\perp$, then $\langle \varphi | \varphi_1 \rangle \neq 0$, and P is a superposition of P_1 and, for instance, $P[\varphi - \langle \varphi_1 | \varphi \rangle \varphi_1]$.

Nonunique Decomposability of Mixed States

By Eq. (9.3) any state ϱ can be obtained as a mixture of at most countably many vector states $(P_i)_{i \in I} \subset \mathbf{P}_1(\mathcal{H})$, $P_i \neq P_j$, for $i \neq j$, with some appropriate weights $(w_i)_{i \in I} \subset [0, 1]$, that is, $\varrho = \sum_{i \in I} w_i P_i$, with the series converging, in particular, in the trace norm. In this case, we say that the family $((P_i, w_i))_{i \in I}$ is a *decomposition* of ϱ , and that $(P_i)_{i \in I}$ *decomposes* ϱ *with the weights* $(w_i)_{i \in I}$.² As the title of this section suggests, a genuine mixed state ϱ , $\varrho^2 \neq \varrho$, admits uncountably many decompositions $((P_i, w_i))_{i \in I}$. These are described in Theorem 9.2.

Remark 9.1 Let I be an arbitrary index set. We say that an indexed family $(w_i)_{i \in I} \subset [0, 1]$ is a family of weights if it is summable with $\sum_{i \in I} w_i = 1$. We recall that this implies that $\{i \in I \mid w_i \neq 0\}$ is finite or countably infinite so that without any loss in generality we may assume that either $I = \{1, \dots, n\}$ for some $n \in \mathbb{N}$, or $I = \mathbb{N}$. $\triangleleft$

Example 9.1 To illustrate the nonunique decomposability, consider first the case where $\varrho \in \mathcal{S}(\mathcal{H})$ is of finite rank, that is $\dim \varrho(\mathcal{H}) = n$ for some $n > 1$. Then $\varrho = \sum_{i=1}^n w_i P[\varphi_i]$, with the (not necessarily different) eigenvalues $w_1, \dots, w_n$, and with some mutually orthogonal unit vectors $\varphi_1, \dots, \varphi_n$. Let $\tilde{\varphi}_i = \sqrt{w_i} \varphi_i$, so that $\varrho = \sum_{i=1}^n |\tilde{\varphi}_i\rangle\langle \tilde{\varphi}_i|$. Take $m \geq n$, and let (c_{ki}) be a complex $m \times n$ -matrix such that all its columns $(c_{1i}, \dots, c_{mi})$, $i = 1, \dots, n$, are mutually orthogonal unit vectors. Define $\tilde{\psi}_k = \sum_{i=1}^n c_{ki} \tilde{\varphi}_i$ for each $k = 1, \dots, m$. Then for each $k = 1, \dots, m$, $\|\tilde{\psi}_k\|^2 = \sum_{i=1}^n w_i |c_{ki}|^2 > 0$ and $\sum_{k=1}^m |\tilde{\psi}_k\rangle\langle \tilde{\psi}_k| = \varrho$, showing that $(P[\tilde{\psi}_k], \|\tilde{\psi}_k\|^2)_{k=1}^m$ is a decomposition of ϱ . This example goes back to [4]; it is not exhaustive since finite rank states admit also countably infinite decompositions. $\triangleleft$

For any state $\varrho \in \mathcal{S}(\mathcal{H})$, the Hilbert space $\mathcal{H}$ decomposes as (Lemma 2.1)

$$\mathcal{H} = \ker(\varrho) \oplus \ker(\varrho)^\perp = \ker(\varrho) \oplus \overline{\varrho^*(\mathcal{H})} = \ker(\varrho) \oplus \overline{\varrho(\mathcal{H})}.$$

We also recall that

$$\varrho(\mathcal{H}) \subset \varrho^{1/2}(\mathcal{H}) \subset \overline{\varrho^{1/2}(\mathcal{H})} = \overline{\varrho(\mathcal{H})}.$$

It is a restatement of a classic result for compact operators that these inclusions are equalities exactly when ϱ is of finite rank (exercise). Before stating a full character-

²The need to form statistical mixtures of (pure) states was equally evident in the early stages of quantum mechanics. The explicit formula $\varrho = \sum_n w_n P[\varphi_n]$ for a statistical mixture of pure states $P[\varphi_n]$ with the weights w_n (Gemisch von Zuständen) was already developed and investigated in detail by von Neumann [3].

isation of all the decompositions of ϱ we quote a partial result due to Hadjisavvas [5]: a vector state $P[\varphi]$ is a component of ϱ in the sense that $\varrho = wP[\varphi] + (1 - w)\varrho'$ for some $w > 0$ and $\varrho' \in \mathcal{S}(\mathcal{H})$ if and only if $\varphi \in \varrho^{1/2}(\mathcal{H})$.

To formulate the main result of this section some preparations are due. Consider again a state $\varrho \in \mathcal{S}(\mathcal{H})$, and let $\mathcal{N} = \ker(\varrho)$, $\mathcal{K} = \overline{\varrho(\mathcal{H})}$, so that $\mathcal{H} = \mathcal{N} \oplus \mathcal{K}$, and let P denote the projection onto $\mathcal{K}$. Since the restriction of $\varrho^{1/2}$ to $\mathcal{K}$ is injective, there is a unique selfadjoint operator A acting in $\mathcal{H}$, with the domain $\varrho^{1/2}(\mathcal{H})$, such that

$$\begin{aligned} A\varrho^{1/2}\varphi &= \varphi & \text{for all } \varphi \in \mathcal{K}, \\ \varrho^{1/2}A\varphi &= \varphi & \text{for all } \varphi \in \varrho^{1/2}(\mathcal{H}). \end{aligned}$$

Let $(\psi_i)_{i \in I}$ be an orthonormal family of vectors in $\mathcal{H}$. The family $(\psi_i)_{i \in I}$ is said to be *nondegenerate with respect to P* if

- (1) $P(\mathcal{H}) = \mathcal{K} \subset \overline{\text{lin}\{\psi_i \mid i \in I\}}$;
- (2) $P\psi_i \neq 0$ for all $i \in I$;
- (3) for any $i \neq j$, the vectors $P\psi_i$ and $P\psi_j$ are not collinear.

Theorem 9.2 ([6]) *Let $\varrho \in \mathcal{S}(\mathcal{H})$ and let $(\psi_i)_{i \in I}$ be a nondegenerate family with respect to the projection P onto $\mathcal{K}$, the closure of the range of ϱ . For all $i \in I$ define*

$$w_i = \langle \psi_i \mid \varrho\psi_i \rangle = \|\varrho^{1/2}\psi_i\|^2 > 0, \quad \eta_i = w_i^{-1/2}\varrho^{1/2}\psi_i.$$

Then the family $(P[\eta_i], w_i)_{i \in I}$ is a decomposition of ϱ , that is, $\varrho = \sum_i w_i P[\eta_i]$. Conversely, all decompositions of ϱ can be obtained in this way (provided that the kernel of ϱ is infinite-dimensional).

9.2 Effects

Any state $\varrho \in \mathcal{S}(\mathcal{H})$ induces an expectation functional $A \mapsto \text{tr}[\varrho A]$ on the set $\mathcal{L}(\mathcal{H})$ of bounded operators. The requirement that the numbers $\text{tr}[\varrho A]$ represent probabilities, $0 \leq \text{tr}[\varrho A] \leq 1$ for all $\varrho \in \mathcal{S}(\mathcal{H})$, implies that the operator A is positive and bounded by the unit operator: $0 \leq A \leq I$. Such operators are called effect operators or simply effects, and the number $\text{tr}[\varrho A]$ is the probability for the effect A in state ϱ . The set of all effect operators is

$$\mathbf{E}(\mathcal{H}) = \{A \in \mathcal{L}(\mathcal{H}) \mid 0 \leq A \leq I\}. \quad (9.4)$$

As a subset of $\mathcal{L}_s(\mathcal{H})$, $\mathbf{E}(\mathcal{H})$ is partially ordered, with 0 and I as its order bounds. The order on $\mathbf{E}(\mathcal{H})$ is connected with the basic probabilities of quantum mechanics. Indeed, for any $A, B \in \mathbf{E}(\mathcal{H})$, $A \leq B$ (in the sense that $B - A \geq 0$) if and only if $\text{tr}[\varrho A] \leq \text{tr}[\varrho B]$ for all $\varrho \in \mathcal{S}(\mathcal{H})$. The map $\mathbf{E}(\mathcal{H}) \ni A \mapsto A^\perp = I - A \in \mathbf{E}(\mathcal{H})$ is an involutive antiautomorphism: it reverses the order (if $A \leq B$, then $B^\perp \leq A^\perp$) and, when applied twice, yields the identity $(A^\perp)^\perp = A$. These properties guarantee

that the de Morgan laws hold in $E(\mathcal{H})$ in the sense that if, for instance, the greatest lower bound $A \wedge B$ of $A, B \in E(\mathcal{H})$ exists in $E(\mathcal{H})$, then also the least upper bound of their (pseudo) complements $A^\perp$ and $B^\perp$ exists in $E(\mathcal{H})$ and $(A \wedge B)^\perp = A^\perp \vee B^\perp$. However, $A \mapsto A^\perp$ is not an orthocomplementation since the greatest lower bound of A and $A^\perp$ need not exist at all, or, even when it does, it need not be the null effect.

An important subset of $E(\mathcal{H})$ is given by the set $P(\mathcal{H})$ of projections. For any $A \in E(\mathcal{H})$, $AA^\perp = A^\perp A$, so that $AA^\perp$ is an effect contained in both A and $A^\perp$. Therefore, the projections can be characterised as those effects A for which the set of lower bounds of A and $A^\perp$, i.e. $\{A, A^\perp\}$, contains only the null effect,

$$P(\mathcal{H}) = \{A \in E(\mathcal{H}) \mid \text{l.b. } \{A, A^\perp\} = \{0\}\}. \quad (9.5)$$

In addition to being an ordered set, $E(\mathcal{H})$ of effects is a convex subset of $\mathcal{L}(\mathcal{H})$: for any $A, B \in E(\mathcal{H})$ and $0 \leq w \leq 1$, $wA + (1-w)B \in E(\mathcal{H})$. This structure reflects the physical possibility of combining measurements into new measurements by mixing them. An effect $A \in E(\mathcal{H})$ is an extreme effect if the condition $A = wA_1 + (1-w)A_2$, with $A_1, A_2 \in E(\mathcal{H})$, $0 < w < 1$, implies that $A = A_1 = A_2$. Extreme effects arise from pure measurements, that is, measurements which cannot be obtained by mixing some other measurements.

Theorem 9.3 *The set of extreme effects coincides with the set of projections,*

$$\text{ex}(E(\mathcal{H})) = P(\mathcal{H}). \quad (9.6)$$

Proof Let $P \in P(\mathcal{H})$ be expressed as $P = wA_1 + (1-w)A_2$ for some $A_1, A_2 \in E(\mathcal{H})$ and $0 < w < 1$. Let $\varphi \in \mathcal{H}$ be such that $P\varphi = 0$. Then

$$\begin{aligned} 0 &\leq w\|A_1^{1/2}\varphi\| = w\langle \varphi | A_1 \varphi \rangle \\ &\leq w\langle \varphi | A_1 \varphi \rangle + (1-w)\langle \varphi | A_2 \varphi \rangle \\ &= \langle \varphi | P \varphi \rangle = 0, \end{aligned}$$

which implies that $A_1\varphi = 0$. Let then $\psi \in \mathcal{H}$ be such that $P\psi = \psi$, that is, $(I - P)\psi = 0$. Since $I - P = w(I - A_1) + (1-w)(I - A_2)$, one obtains that also $(I - A_1)\psi = 0$. Since $\mathcal{H} = \ker(P) \oplus P(\mathcal{H})$, this shows that $A_1 = P$. It follows immediately that $A_2 = P$, so that P is an extreme effect.

Consider now an $A \in E(\mathcal{H})$ which is not a projection. Then there is an $a \in \sigma(A)$ with $0 < a < 1$. Let f be a continuous function on $[0, 1]$ such that $0 \leq x \pm f(x) \leq 1$ for all $x \in [0, 1]$ and for which $f(a) \neq 0$. Then both $A_1 = A + f(A)$ and $A_2 = A - f(A)$ are effects, $A_1 \neq A \neq A_2$, but $A = \frac{1}{2}A_1 + \frac{1}{2}A_2$, which shows that A is not an extreme effect. $\square$

Remark 9.2 By virtue of the spectral theorem, any effect $A \in E(\mathcal{H})$ can be approximated in norm by a sum $\sum_{i=1}^n a_i P_i$, where $P_1, \dots, P_n$ are mutually orthogonal projections and $a_i \in [0, 1]$. This need not be a convex combination. However,

$\mathbf{P}(\mathcal{H}) \subset \text{conv}(\mathbf{P}(\mathcal{H})) \subset \mathbf{E}(\mathcal{H})$. Moreover, the weak closure of $\mathbf{P}(\mathcal{H})$ is the whole set of effects [7, Lemma 2.4]. $\triangleleft$

The algebraic structure of $\mathcal{L}(\mathcal{H})$ renders $\mathbf{E}(\mathcal{H})$ a partial algebra. Indeed, for any $A, B \in \mathbf{E}(\mathcal{H})$, their sum $A + B$ is an effect whenever the operator $A + B$ is bounded by the unit operator. Moreover, for each $A \in \mathbf{E}(\mathcal{H})$, there is a unique $A' \in \mathbf{E}(\mathcal{H})$ such that $A + A' = I$. Clearly, $A' = A^\perp$. This structure is closely related to the physical possibility that the effects A and B for which $A + B \leq I$ can be measured together. The partial sum leads one to define an order on $\mathbf{E}(\mathcal{H})$: for any $A, B \in \mathbf{E}(\mathcal{H})$, we write $A \leq B$ exactly when there is a $C \in \mathbf{E}(\mathcal{H})$ such that $A + C = B$. Obviously, the order so defined agrees with the order given by the notion of a positive operator. We also observe that if $P_1, P_2 \in \mathbf{P}(\mathcal{H})$, then $P_1 + P_2$ is an effect if and only if it is a projection, hence $\mathbf{P}(\mathcal{H})$ itself is endowed with a partial algebra structure by restricting on it the partially defined sum of $\mathbf{E}(\mathcal{H})$. The order defined on $\mathbf{P}(\mathcal{H})$ by this partial sum is obviously the standard one.

Remark 9.3 There is an important difference between $\mathbf{P}(\mathcal{H})$ and $\mathbf{E}(\mathcal{H})$ for what concerns the relation between their structures of partial algebras and ortho-ordered sets. In fact, given $P_1, P_2 \in \mathbf{P}(\mathcal{H})$, one has $P_1 + P_2 \in \mathbf{P}(\mathcal{H})$ if and only if $P_1 \leq P_2^\perp$ and, in this case, $P_1 + P_2 = P_1 \vee P_2$. Hence, not only does the partial algebra structure of $\mathbf{P}(\mathcal{H})$ determine its order structure, but the converse is also true. This is, however, not the case in the set of effects. In fact, there exist effects $A, B \in \mathbf{E}(\mathcal{H})$ such that $A \leq B^\perp$ and $A + B \in \mathbf{E}(\mathcal{H})$, but $A + B \neq A \vee B$, as would be required if we wanted to define the partial sum in terms of the order. This is due to the fact that $A \vee B$ need not exist at all. As an example, consider $A = \alpha P_1, B = \beta P_2$, with $0 < \alpha < \beta < 1, P_1 \leq P_2^\perp$ and $P_1, P_2 \in \mathbf{P}(\mathcal{H})$. Then $\alpha P_1 \leq (\beta P_2)^\perp, \alpha P_1 + \beta P_2 \in \mathbf{E}(\mathcal{H})$, but $\alpha P_1 \vee \beta P_2$ does not exist [8]. $\triangleleft$

The notion of the compatibility of effects is a fundamental concept in quantum mechanics, introduced to describe effects that can be measured together by measuring a single observable. Any two effects $A, B \in \mathbf{E}(\mathcal{H})$ are said to be compatible if and only if there are effects $A_1, B_1, C \in \mathbf{E}(\mathcal{H})$ such that $A = A_1 + C, B = B_1 + C$, and $A_1 + B_1 + C \leq I$. The compatibility of two or more projections $P_1, P_2, \dots$ is equivalent to their commutativity (exercise). The concept of compatibility is investigated in greater detail in Chap. 11.

States as Mappings on $\mathbf{E}(\mathcal{H})$

The basic probability measures of quantum mechanics are of the form $E_\varrho : \mathcal{A} \rightarrow [0, 1]$, with $E_\varrho(X) = \text{tr}[\varrho E(X)]$, where $\varrho \in S(\mathcal{H})$ and $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a semispectral measure, defined on a σ -algebra $\mathcal{A}$ of subsets of a nonempty set Ω , Sect. 9.3. Defining a map $m_\varrho : \mathbf{E}(\mathcal{H}) \rightarrow [0, 1]$ by $m_\varrho(A) = \text{tr}[\varrho A]$ for all $A \in \mathbf{E}(\mathcal{H})$, the probability measure E_ϱ may thus be considered as a composition of the maps $E : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{H})$ and $m_\varrho : \mathbf{E}(\mathcal{H}) \rightarrow [0, 1]$, $E_\varrho = m_\varrho \circ E$. Clearly, the map m_ϱ preserves order, complement, convexity, and partial sum:

- (i) if $A \leq B$, then $m_\varrho(A) \leq m_\varrho(B)$,
- (ii) $m_\varrho(I - A) = 1 - m_\varrho(A)$,
- (iii) $m_\varrho(wA + (1 - w)B) = wm_\varrho(A) + (1 - w)m_\varrho(B)$,
- (iv) if $A, B \in \mathbf{E}(\mathcal{H})$ and $A + B \in \mathbf{E}(\mathcal{H})$, then $m_\varrho(A + B) = m_\varrho(A) + m_\varrho(B)$.

Moreover, since in $\mathbf{E}(\mathcal{H})$ the weak and the ultraweak operator topologies coincide,

- (v) the map $A \mapsto m_\varrho(A)$ on $\mathbf{E}(\mathcal{H})$ is continuous with respect to the weak operator topology.

Furthermore, due to the normality of m_ϱ ,

- (vi) if $(A_i)_{i \in I} \subset \mathbf{E}(\mathcal{H})$ is an increasing net, $m_\varrho(\sup_{i \in I} A_i) = \sup_{i \in I} m_\varrho(A_i)$.

In view of the probability structures at hand, one might ask if a state ϱ could be characterised as a kind of generalised probability measure on $\mathbf{E}(\mathcal{H})$. The pseudo ortho-ordered effect structure $(\mathbf{E}(\mathcal{H}), \leq, \perp, 0, I)$ allows one to define a map $m : \mathbf{E}(\mathcal{H}) \rightarrow [0, 1]$ which preserves the order (i) and pseudo-complement (ii). However, since for $A, B \in \mathbf{E}(\mathcal{H})$, the condition $A \leq B^\perp$, that is $A + B \leq I$, does not guarantee the existence of $A \vee B$ in $\mathbf{E}(\mathcal{H})$, one cannot require that $m(A \vee B) = m(A) + m(B)$ for “mutually orthogonal elements” $A, B \in \mathbf{E}(\mathcal{H})$, $A \leq B^\perp$. In other words, the above additivity property (iv) does not relate to the order structure of $\mathbf{E}(\mathcal{H})$: for $A, B \in \mathbf{E}(\mathcal{H})$, if $A \leq I - B$, then $A + B$ is an upper bound of A and B but it need not coincide with $A \vee B$ (if the latter exists at all). Hence there seems to be no natural way to define a generalised probability measure on $\mathbf{E}(\mathcal{H})$ (in the usual full sense of the concept).

The situation changes completely if one considers the restriction of such a map to $\mathbf{P}(\mathcal{H})$. Indeed, for any $P, R \in \mathbf{P}(\mathcal{H})$, P and R are mutually orthogonal in the sense that $P \leq R^\perp$ if and only if $P + R$ is a projection, in which case $P + R = P \vee R$. With respect to the ortho-order structure $(\mathbf{P}(\mathcal{H}), \leq, \perp, 0, I)$ of projection operators one may thus define a generalised probability measure $m : \mathbf{P}(\mathcal{H}) \rightarrow [0, 1]$ for which $m(I) = 1$ and σ -additivity holds, in the sense that for a sequence $(P_i)_{i \in I}$ of pairwise orthogonal ($P_i \leq P_j^\perp$, $i \neq j$) projections, $m(\vee P_i) = m(\sum P_i) = \sum m(P_i)$. Clearly, the restriction of any m_ϱ to $\mathbf{P}(\mathcal{H})$ is a generalised probability measure. The celebrated Gleason theorem gives the converse result:

Theorem 9.4 (Gleason [9]) *Let $\mathcal{H}$ be a complex separable Hilbert space of $\dim(\mathcal{H}) \geq 3$. If $m : \mathbf{P}(\mathcal{H}) \rightarrow [0, 1]$ is a generalised probability measure, then there is a unique $\varrho \in \mathcal{S}(\mathcal{H})$ such that $m = m_\varrho$.*

Disregarding the fact that there is no obvious way to define a generalised probability measure (in the full sense described) on $\mathbf{E}(\mathcal{H})$ it is of interest to note that the additivity and continuity properties (iv) and (vi) suffice to determine the structure of a map $m : \mathbf{E}(\mathcal{H}) \rightarrow [0, 1]$.

Proposition 9.1 *Let $m : \mathbf{E}(\mathcal{H}) \rightarrow [0, 1]$ be a function with the following properties:*

- (i) *if $A + B \leq I$, then $m(A + B) = m(A) + m(B)$,*
- (ii) *if $(A_i)_{i \in I}$ is an increasing net in $\mathbf{E}(\mathcal{H})$, then $m(\sup_{i \in I} A_i) = \sup_{i \in I} m(A_i)$.*

Then there is a unique $\varrho \in \mathcal{S}(\mathcal{H})$ such that $m = m_\varrho$.

Proof We note first that $m(A) = m(A + 0) = m(A) + m(0)$, so that $m(0) = 0$. We prove next that for all $A \in \mathbf{E}(\mathcal{H})$ and $0 < \lambda < 1$,

$$m(\lambda A) = \lambda m(A).$$

If λ is rational, this follows from the additivity of m . Let $0 < \lambda < 1$ and let (r_n) be an increasing sequence of positive rationals converging to λ . Then $\sup_n (r_n A) = \lambda A$, and this implies that

$$m(\lambda A) = m\left(\sup_n \{r_n A\}\right) = \sup_n m(r_n A) = \sup_n (r_n m(A)) = \lambda m(A).$$

The (unique) extension of m to a positive linear map $\hat{m} : \mathcal{L}_s(\mathcal{H}) \rightarrow \mathbb{R}$ is straightforward, and is left as an exercise. By Theorem 6.1, condition (ii) implies that the map $\hat{m}$ is normal and thus by Proposition 6.1 it arises from a unique trace class operator ϱ , which is positive (Theorem 3.14). Since $\hat{m}$ is uniquely defined by its restriction m on $\mathbf{E}(\mathcal{H})$, the proof is complete. $\square$

For any state $\varrho \in \mathcal{S}(\mathcal{H})$ there is the smallest projection $s(\varrho) \in \mathbf{P}(\mathcal{H})$ such that $\text{tr}[\varrho s(\varrho)] = 1$. This is the *support projection* of the state ϱ . The support function $\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto s(\varrho) \in \mathbf{P}(\mathcal{H})$ is surjective. When restricted to the set $\mathbf{P}_1(\mathcal{H})$ of pure states this function is a bijection between the pure states in $\mathcal{S}(\mathcal{H})$ and the atoms, the one-dimensional projections on $\mathbf{P}(\mathcal{H})$. This leads to the important results that a pure state can be identified with the collection of properties (extreme effects) that the system has in that state in the sense of probabilistic certainty.

The theorem of Gleason assumes that the Hilbert space is at least three dimensional. For $\mathcal{H} = \mathbb{C}^2$ it is still true that any state ϱ defines a probability measure on $\mathbf{P}(\mathbb{C}^2)$. In this case there is also a lavish supply of other probability measures on $\mathbf{P}(\mathbb{C}^2)$ (exercise). None of them has a support in $\mathbf{P}(\mathbb{C}^2)$. Neither is it known that any of them would have some physical relevance.

9.3 Observables

Observables as Operator Measures

If one adopts the representation of the states of a quantum system $\mathcal{S}$ as density operators $\varrho \in \mathcal{S}(\mathcal{H})$ in the associated Hilbert space, the description of observables as normalised positive operator measures need not be postulated separately but follows by implementation of the idea of measurement statistics. Indeed, assume that a physical quantity O with value space $(\Omega, \mathcal{A})$ is measured on the system prepared in a state ϱ , and a result in the set $X \in \mathcal{A}$ is registered. Then for each state it should be possible to assign a number $p_\varrho^O(X)$ that gives the corresponding probability for this outcome set, thus constituting a mapping

$$\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto p_\varrho^O(X) \in [0, 1]. \quad (9.7)$$

If the state ϱ is a mixture $\varrho = \sum_i w_i \varrho_i$, it is natural to require³ that

$$p_\varrho^O(X) = \sum_i w_i p_{\varrho_i}^O(X). \quad (9.8)$$

Assuming that (9.8) holds for any state ϱ and its possible decompositions, then it is an easy exercise to show that the map (9.7) extends to a positive linear functional $\Phi_X^O : \mathcal{T}_s(\mathcal{H}) \rightarrow \mathbb{R}$ such that $\Phi_X^O(\varrho) = p_\varrho^O(X)$ for any state ϱ . By the duality $\mathcal{T}_s(\mathcal{H})^* \simeq \mathcal{L}_s(\mathcal{H})$, there is a unique bounded selfadjoint operator $E(X)$ such that $\Phi_X^O(\varrho) = \text{tr}[\varrho E(X)]$ for all $\varrho \in \mathcal{T}_s(\mathcal{H})$. Since these numbers are in the interval $[0, 1]$ for all $\varrho \in \mathcal{S}(\mathcal{H})$, it follows that the operator $E(X)$ is positive and bounded by I , i.e., it is an effect operator. Varying the set X , one thus obtains an effect valued map

$$\mathcal{A} \ni X \mapsto E(X) \in E(\mathcal{H}). \quad (9.9)$$

Since for any state ϱ the composite map

$$\mathcal{A} \ni X \mapsto E(X) \mapsto E_\varrho(X) \in [0, 1]$$

is to give the probability measure for the measurement outcomes in that state ϱ , it now follows from Lemma 9.1 that

$$E_\varrho(X) = \text{tr}[\varrho E(X)] \quad (9.10)$$

and that the effect valued map (9.9) is, in fact, a normalised positive operator measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ (exercise).

The above arguments lead one to accept the point of view that an *observable* of a quantum system $\mathcal{S}$, described by a Hilbert space $\mathcal{H}$, is represented as and identified with a semispectral measure $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$.⁴ The space $(\Omega, \mathcal{A})$ is the value space of the observable E . Given an observable E and a state ϱ , the probabilities $E_\varrho(X)$ are given by the trace formula (9.10) and they are interpreted as the measurement outcome probabilities, with the intended meaning that if the same E -measurement is performed on system $\mathcal{S}$ in a state ϱ N times (or on the identically prepared independent copies of $\mathcal{S}$) and if the result is registered $N(X)$ times in the set X , then, for large N , $E_\varrho(X) \approx N(X)/N$.

³As discussed in Sect. 1.2, this assumption is built into the very idea of a mixed state.

⁴This argument is essentially the same as that used by von Neumann in [1] to show that the observables of a quantum system are to be represented as hypermaximal symmetric operators (which are now called selfadjoint operators). The sole difference between von Neumann's and the present approach is that here we do not adopt his assumption that the measurement outcome statistics of an observable are to be represented in terms of a single (selfadjoint) operator.

Let $M(\Omega)_1^+$ denote the set of probability measures on $\mathcal{A}$. From the results of Sect. 4.6 one may easily check that an observable $\mathbf{E} : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{H})$ is uniquely defined by any of the three mappings

$$\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto \mathbf{E}_\varrho \in M_1^+(\Omega) \quad (9.11)$$

$$\mathbf{P}_1(\mathcal{H}) \ni P \mapsto \mathbf{E}_P \in M_1^+(\Omega) \quad (9.12)$$

$$\mathcal{H}_1 \ni \varphi \mapsto \mathbf{E}_{P[\varphi]} = \mathbf{E}_\varphi \in M_1^+(\Omega). \quad (9.13)$$

They express the fact that an observable is uniquely determined by the totality of its measurement outcome probabilities, and that the pure states suffice to identify an observable.

As indicated in (9.13), we may also write $\mathbf{E}_\varphi$ instead of $\mathbf{E}_{P[\varphi]}$ for a vector state $P = P[\varphi]$. In that case,

$$\mathbf{E}_\varphi(X) = \text{tr}[P[\varphi]\mathbf{E}(X)] = \langle \varphi | \mathbf{E}(X)\varphi \rangle.$$

This formula underlines the fact that any two unit vectors φ and ψ which differ only by a phase factor, that is, $\varphi = e^{i\lambda}\psi$ for some $\lambda \in \mathbb{R}$, give the same probabilities $\langle \varphi | \mathbf{E}(X)\varphi \rangle = \langle \psi | \mathbf{E}(X)\psi \rangle$ for all observables $\mathbf{E}$ and thus they represent the same state.

In all our physical applications the value space $(\Omega, \mathcal{A})$ of an observable $\mathbf{E}$ is the real Borel space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, a Cartesian product space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $n = 1, 2, \dots$, or a closed or open subspace of it. We let $\mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ denote the set of all observables $\mathbf{E}$ with the value space $(\Omega, \mathcal{A})$. If the value space Ω is a topological space and $\mathcal{A}$ is its Borel σ -algebra, we suppress $\mathcal{B}(\Omega)$ from the notation $\mathcal{O}(\Omega, \mathcal{B}(\Omega), \mathcal{H})$.

Among the observables there are those given by normalised projection measures, spectral measures. For reasons to be explained later we call such observables *sharp observables*. All other observables will be called *unsharp*. We recall from Proposition 4.6 that an observable $\mathbf{E} \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ is sharp, that is, $\mathbf{E}(X)^2 = \mathbf{E}(X)$ for all $X \in \mathcal{A}$, if and only if $\mathbf{E}$ is multiplicative, that is, for any $X, Y \in \mathcal{A}$, $\mathbf{E}(X \cap Y) = \mathbf{E}(X)\mathbf{E}(Y)$. In addition to these technical conditions, we shall meet some physical characterisations for an observable to be sharp in the next two sub-sections.

For later use we introduce the following distinction. An observable $\mathbf{E} \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ is *discrete* if there is a countable set $\Omega_0 \subset \Omega$ such that $\{\omega\} \in \mathcal{A}$ for all $\omega \in \Omega_0$ and $\mathbf{E}(\Omega_0) = I$. Then $\mathbf{E}(X) = \mathbf{E}(\Omega_0 \cap X) = \sum_{\omega_i \in \Omega_0 \cap X} \mathbf{E}(\{\omega_i\})$ for any $X \in \mathcal{A}$ and $\varrho \in \mathcal{S}(\mathcal{H})$. We will often denote a discrete observable simply by means of the set of its generating effects $E_i = \mathbf{E}(\{\omega_i\})$, that is, $\mathbf{E} = \{E_1, E_2, \dots\}$. One may also distinguish observables with respect to their ability to identify the state which is being measured. An observable $\mathbf{E}$ is *informationally complete* if its probability distributions identify the state unambiguously, that is, the map $\varrho \mapsto \mathbf{E}_\varrho$ is one-to-one. This class of observables will be studied in Chap. 18. It will be seen there that an informationally complete observable is necessarily unsharp. At the other end of the spectrum, one finds observables that are completely uninformative. An observable

$\mathbf{E}$ is *trivial* if its probability distributions do not depend on the states, that is, there is a probability measure λ on the outcome space of $\mathbf{E}$ such that $\mathbf{E} = \lambda I$.

Observables on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and Selfadjoint Operators

According to the spectral theorem for selfadjoint operators, each sharp observable $\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$ defines a unique selfadjoint operator $L(x, \mathbf{E}) = \mathbf{E}[1]$, the first moment operator of $\mathbf{E}$, with its respective domain $\mathcal{D}(x, \mathbf{E})$ (see Lemma 5.5), and any selfadjoint operator A , with the domain $\mathcal{D}(A)$, is the first moment operator of a unique sharp observable $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{E}(\mathcal{H})$, that is, $A = \mathbf{E}[1]$, with $\mathcal{D}(A) = \mathcal{D}(x, \mathbf{E})$. To emphasise this connection between sharp observables $\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$ and selfadjoint operators A acting in $\mathcal{H}$ we may write $\mathbf{E}^A$ for the sharp observable having A as its first moment operator, i.e. $\mathbf{E}^A[1] = A$. For selfadjoint operators $A, B, \dots$ we also use the corresponding *Sans Serif* letters $A, B, \dots$ to denote their spectral measures.

The importance of the selfadjoint operator representation of a sharp observable is twofold. First, the measurement outcome statistics of such an observable $\mathbf{E}^A$ in a state ϱ is often uniquely determined by the n th moments, that is, the expectations of the powers of A , $\text{tr}[\varrho A^n]$, $\varrho \in \mathcal{S}(\mathcal{H})$, $n = 1, 2, \dots$; moreover, if the measured observable is sharp, $\mathbf{E} = \mathbf{E}^A$, then the numbers $\text{tr}[\varrho A]$, $\text{tr}[\varrho A^2]$, $\varrho \in \mathcal{S}(\mathcal{H})$ (or even only $\varrho \in \mathcal{P}_1(\mathcal{H})$), suffice to determine the observable (exercise). (Note that the notation $\text{tr}[\varrho B]$ is well defined for bounded operators B , but if B is an unbounded operator on $\mathcal{H}$, then it must be understood as a formal notation or a basis dependent trace. For example, if $\varrho = \sum_i w_i P[\varphi_i]$ where $\{\varphi_i\}$ is an orthonormal basis of $\mathcal{H}$ then one may define $\text{tr}[\varrho B] = \sum_i w_i \langle \varphi_i | B \varphi_i \rangle$ when the vectors φ_i belong to the domain of B and the series converges.)

Second, it is common practice in physics to define concrete observables in terms of their first moment operators, often with reference to classical correspondence; this is clearly the case with the position and momentum operators of a particle and, in particular, the energy of mutually interacting systems. We elaborate somewhat further on the observable-operator correspondence.

As noted above, an observable $\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$ is associated with a sequence of symmetric operators, its moment operators $\mathbf{E}[k] = L(x^k, \mathbf{E})$, $k \in \mathbb{N}_0$ (cf. Sect. 8.5). In particular, its first and second moment operators are the operators

$$\mathbf{E}[1] = \int_{\mathbb{R}} x d\mathbf{E}(x), \quad \mathbf{E}[2] = \int_{\mathbb{R}} x^2 d\mathbf{E}(x),$$

with the domains $\mathcal{D}(x, \mathbf{E})$ and $\mathcal{D}(x^2, \mathbf{E})$, which contain the respective square integrability domains.

These operators are very important since they allow one to express the expectation and the variance (or its square root, the standard deviation) of the measurement outcome statistics of the observable. Indeed, for any state $\varrho \in \mathcal{S}(\mathcal{H})$, the expectation and the variance of the probability measure $\mathbf{E}_\varrho$ are given by the integrals

$$\begin{aligned}\text{Exp}(\mathbf{E}_\varrho) &= \int_{\mathbb{R}} x d\mathbf{E}_\varrho(x) = \text{tr}[\varrho \mathbf{E}[1]], \\ \text{Var}(\mathbf{E}_\varrho) &= \int_{\mathbb{R}} x^2 d\mathbf{E}_\varrho(x) - \left(\int_{\mathbb{R}} x d\mathbf{E}_\varrho(x) \right)^2 = \text{tr}[\varrho \mathbf{E}[2]] - \text{tr}[\varrho \mathbf{E}[1]]^2,\end{aligned}$$

where, in both cases, the second equalities require that the state ϱ can be decomposed in the square integrability domain of the relevant operator. Defining the operator

$$N(\mathbf{E}) = \mathbf{E}[2] - \mathbf{E}[1]^2, \quad (9.14)$$

the variance of the statistics $\mathbf{E}_\varrho$ takes a suggestive form, the sum of two nonnegative terms,

$$\text{Var}(\mathbf{E}_\varrho) = (\text{tr}[\varrho \mathbf{E}[1]^2] - \text{tr}[\varrho \mathbf{E}[1]]^2) + \text{tr}[\varrho N(\mathbf{E})] \quad (9.15)$$

(which again is valid for states supported by the relevant operator domains). It must be emphasised, however, that neither of the two terms on the right hand side of (9.15) can be determined from the statistics $\mathbf{E}_\varrho$. Therefore, special care is required with the applicability and interpretation of this splitting.

Assume now that the first moment operator $\mathbf{E}[1]$ of the observable $\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$ is actually selfadjoint, and let $\mathbf{E}^A, A = \mathbf{E}[1]$, be the corresponding sharp observable. In this case the variance of $\mathbf{E}$ in a state ϱ can be written in the form

$$\text{Var}(\mathbf{E}_\varrho) = \text{Var}(\mathbf{E}_\varrho^A) + \text{tr}[\varrho N(\mathbf{E})] \geq \text{Var}(\mathbf{E}_\varrho^A). \quad (9.16)$$

In general, the term $\text{tr}[\varrho N(\mathbf{E})]$ is nonzero, and thus the above inequality is strict. The operator $N(\mathbf{E})$ is occasionally called the *noise* or *intrinsic noise* operator of the observable $\mathbf{E}$. We shall meet examples where this term can directly be associated with the accuracy of the measurement (Chap. 15) and examples where such an interpretation is questionable (Chap. 16). According to Theorem 8.5, an observable $\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$ with selfadjoint first moment operator is sharp exactly when its noise operator is zero.

Let A be a selfadjoint operator, with domain $\mathcal{D}(A)$, acting in $\mathcal{H}$, and define

$$\mathcal{O}(A) = \{\mathbf{E} \in \mathcal{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H}) \mid \mathbf{E}[1] = A\}.$$

The set $\mathcal{O}(A)$ consists of all the observables, with the value space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which are indistinguishable from the sharp observable $\mathbf{E}^A$ in terms of their statistical expectations. The above results lead to the following corollary:

Corollary 9.1 (Ingarden [10]) *The sharp observable $\mathbf{E}^A$ is characterised as the noiseless observable in $\mathcal{O}(A)$. In particular, for each $\mathbf{E} \in \mathcal{O}(A)$,*

$$\text{Var}(\mathbf{E}_\varrho) \geq \text{Var}(\mathbf{E}_\varrho^A)$$

for any $\varrho \in \mathcal{S}(\mathcal{H})$.

Extremal Observables

The set $E(\mathcal{H})$ of effects is convex with its extreme elements given by the set $P(\mathcal{H})$ of sharp effects (projection operators). The convex structure of $E(\mathcal{H})$ reflects the possibility of the statistical mixing of (simple) measurements. It induces a convex structure on the set $O(\Omega, \mathcal{A}, \mathcal{H})$ of observables $E : \mathcal{A} \rightarrow E(\mathcal{H})$ with a common value space $(\Omega, \mathcal{A})$. In fact, for $E_1, E_2 \in O(\Omega, \mathcal{A}, \mathcal{H})$ and any $\lambda \in [0, 1]$, then $E = \lambda E_1 + (1 - \lambda)E_2$ is an observable, defined by

$$\mathcal{A} \ni X \mapsto E(X) = \lambda E_1(X) + (1 - \lambda)E_2(X) \in E(\mathcal{H}). \quad (9.17)$$

Physically, the convexity of $O(\Omega, \mathcal{A}, \mathcal{H})$ represents the randomisation of measurements. However, it is important to note that the decomposition (9.17) is never unique. Hence, as in the case of states, the “random mixing” interpretation may in some contexts be an oversimplification.

Let $\text{ex}(O(\Omega, \mathcal{A}, \mathcal{H}))$ denote the set of extremal elements in $O(\Omega, \mathcal{A}, \mathcal{H})$, and let $O_P(\Omega, \mathcal{A}, \mathcal{H}) \subset O(\Omega, \mathcal{A}, \mathcal{H})$ denote the set of the sharp observables. Using the fact that $P(\mathcal{H}) = \text{ex}(S(\mathcal{H}))$ (cf. Theorem 9.3), it is an easy exercise to show that

$$O_P(\Omega, \mathcal{A}, \mathcal{H}) \subset \text{ex}(O(\Omega, \mathcal{A}, \mathcal{H})). \quad (9.18)$$

Thus sharp observables are free from any classical randomisation in the sense described above. However, only for two-valued observables is the inclusion in (9.18) an equality. Indeed, let $\Omega = \{a, b\}$, $\mathcal{A} = 2^\Omega$, and $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ an observable. If, say, $E(\{a\})$ is not a projection, then one has a nontrivial convex combination $E(\{a\}) = \frac{1}{2}A_1 + \frac{1}{2}A_2$, $A_1 \neq A_2 \in E(\mathcal{H})$ (see Theorem 9.3), and we can write $E = \frac{1}{2}E_1 + \frac{1}{2}E_2$ where E_i , $i = 1, 2$, is an observable defined by $E_i(\{a\}) = A_i$. For all other cases, the inclusion is proper, as shown, for instance in [11, p. 44].

In many physical applications the set of relevant observables in $O(\Omega, \mathcal{A}, \mathcal{H})$ is further specified, for example, by a symmetry requirement. The restricted set of observables $O_R(\Omega, \mathcal{A}, \mathcal{H})$ may still be convex and one may search for its extremal elements. Clearly,

$$\text{ex}(O(\Omega, \mathcal{A}, \mathcal{H})) \cap O_R(\Omega, \mathcal{A}, \mathcal{H}) \subset \text{ex}(O_R(\Omega, \mathcal{A}, \mathcal{H})). \quad (9.19)$$

Example 9.2 In the set of real observables $O(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})$, the set $O_P(\mathbb{R}, \mathcal{H})$ of sharp observables is a proper subset of $\text{ex}(O(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H}))$. Consider the case $\mathcal{H} = L^2(\mathbb{R})$ and let M denote the canonical spectral measure (position in the position representation, momentum in the momentum representation). Let $M(\mathbb{R})_1^+$ denote the set of all probability measures $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$, and form the set $O_u(\mathbb{R}, \mathcal{H}) = \{\mu * M \mid \mu \in M(\mathbb{R})_1^+\}$ of unsharp canonical observables (position or momentum). This is a convex set whose extremal elements are exactly those $\mu * M$ with $\mu \in \text{ex}(M(\mathbb{R})_1^+)$. But $\mu \in \text{ex}(M(\mathbb{R})_1^+)$ exactly when μ is the Dirac measure δ_{x_o} concentrated at a point $x_o \in \mathbb{R}$. Since $\delta_{x_o} * M(X) = M(X - x_o)$, we conclude that in this case

$$\text{ex}(\mathbf{O}_u(\mathbb{R}, \mathcal{H})) \subset \mathbf{O}_P(\mathbb{R}, \mathcal{H}) \subset \text{ex}(\mathbf{O}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{H})). \quad \triangleleft$$

Example 9.3 Consider the set $\mathbf{O}(\mathbb{R}^2, \mathcal{H})$ with $\mathcal{H} = L^2(\mathbb{R})$. The covariant phase space observables defined in Sect. 8.4 form an important subset of physical interest, as will become evident in later chapters. Denote this set by $\mathbf{O}_W(\mathbb{R}^2, \mathcal{H})$. According to Theorem 8.6, an observable $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathbf{E}(\mathcal{H})$ is a covariant phase space observable if and only if it is of the form $\mathbf{E} = \mathbf{G}^S$ for some generating operator $S \in \mathcal{T}(\mathcal{H})_1^+$. None of the observables in $\mathbf{O}_W(\mathbb{R}^2, \mathcal{H})$ is sharp. Due to the bijective convexity preserving correspondence between the sets $\mathbf{O}_W(\mathbb{R}^2, \mathcal{H})$ and $\mathcal{S}(\mathcal{H})$, and due to the fact that $\text{ex}(\mathcal{S}(\mathcal{H})) = \mathbf{P}_1(\mathcal{H})$, one gets

$$\text{ex}(\mathbf{O}_W(\mathbb{R}^2, \mathcal{H})) = \{\mathbf{G}^S \in \mathbf{O}_W(\mathbb{R}^2, \mathcal{H}) \mid S \in \mathbf{P}_1(\mathcal{H})\}. \quad \triangleleft$$

The minimal diagonal Naimark dilation of a semispectral measure can be used to characterise the set $\text{ex}(\mathbf{O}(\Omega, \mathcal{A}, \mathcal{H}))$ of extremal observables. Let $\mathbf{E} \in \mathbf{O}(\Omega, \mathcal{A}, \mathcal{H})$, and let $(\mathcal{H}_\oplus, \mathbf{M}, V_\oplus)$ and

$$\mathbf{E}(X) = V_\oplus^* \mathbf{M}(X) V_\oplus = \sum_{n,m=1}^{\dim \mathcal{H}} \int_X \langle \eta_n(x) | \eta_m(x) \rangle d\mu(x) |\varphi_n\rangle \langle \varphi_m|$$

be its minimal diagonal Naimark dilation and Kolmogorov decomposition, respectively (see Sect. 8.6). Here the vectors $\eta_n = V_\oplus \varphi_n$ constitute an orthonormal set in the direct integral Hilbert space $\mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}_{m(x)} d\mu(x)$ where μ is a (probability) measure such that $\mathbf{E}$ is absolutely continuous with respect to it; moreover, the linear combinations of the vectors $\chi_X \eta_n$ are dense in $\mathcal{H}_\oplus$.

Theorem 9.5 *An observable $\mathbf{E} \in \mathbf{O}(\Omega, \mathcal{A}, \mathcal{H})$ is an extremal observable if and only if, for any bounded decomposable operator*

$$D = \int_\Omega^\oplus D(x) d\mu(x)$$

on $\int_\Omega^\oplus \mathcal{H}_{m(x)} d\mu(x)$ the condition $V_\oplus^* D V_\oplus = 0$, i.e. the condition

$$\int_\Omega \langle \eta_n(x) | D(x) \eta_m(x) \rangle d\mu(x) = 0$$

for all n, m , implies that $D = 0$. In this case, the linear map $L^\infty(\mu) \ni f \mapsto \int f d\mathbf{E} \in \mathcal{L}(\mathcal{H})$ is injective.

Proof This theorem is an immediate consequence of Theorem 7.13 (see Example 7.2), Remark 8.1, and Proposition 8.11. The last claim follows from the fact that any $f \in L^\infty(\mu)$ corresponds to a multiplicative (or diagonalisable) bounded

operator $\hat{f} \in \mathcal{L}(\mathcal{H}_\oplus)$, $(\hat{f}\eta)(x) = f(x)\eta(x)$, $\eta \in \mathcal{H}_\oplus$, which is decomposable, $\hat{f} = \int_\Omega^\oplus f(x)I_{\mathcal{H}_{m(x)}}d\mu(x)$. In addition, $V_\oplus^*\hat{f}V_\oplus = \int f d\mathbf{E}$. $\square$

If $\mathbf{E} \in \mathbf{O}_P(\Omega, \mathcal{A}, \mathcal{H})$, then $\{\eta_n\}$ is an orthonormal basis of $\int_\Omega^\oplus \mathcal{H}_{m(x)}d\mu(x)$ and one sees immediately that $\mathbf{E} \in \text{ex}(\mathbf{O}(\Omega, \mathcal{A}, \mathcal{H}))$.

If $\mathbf{E}$ is of rank 1 (i.e. $m(x) = 1$ for μ -almost all $x \in \Omega$) then any decomposable D corresponds a function $f \in L^\infty(\mu)$ (i.e. it is multiplicative) and $\mathbf{E}$ is extremal if and only if the linear map $L^\infty(\mu) \ni f \mapsto \int f d\mathbf{E} \in \mathcal{L}(\mathcal{H})$ is injective.

Example 9.4 The set $\text{ex}(\mathbf{O}_W(\mathbb{R}^2, \mathcal{H}))$ of extremal covariant phase space observables of Example 9.3 contains no projection valued observables, that is, the intersection $\mathbf{O}_P(\mathbb{R}^2, \mathcal{H}) \cap \text{ex}(\mathbf{O}_W(\mathbb{R}^2, \mathcal{H}))$ is empty. We go on to show that

$$\text{ex}(\mathbf{O}(\mathbb{R}^2, \mathcal{H})) \cap \text{ex}(\mathbf{O}_W(\mathbb{R}^2, \mathcal{H})) \neq \emptyset.$$

Let h_0 be the 0th Hermite function, $S = |0\rangle\langle 0|$ the projection onto $[h_0]$, and consider the (rank-1) phase space observable $\mathbf{G}^{(0)}$, which in the complex representation assumes the form

$$\mathbf{G}^{(0)}(Z) = \frac{1}{\pi} \int_Z |z\rangle\langle z| d^2z = \sum_{n,m=0}^{\infty} \int_Z \frac{r^{n+m} e^{i(n-m)\theta}}{\sqrt{n!m!}} e^{-r^2} dr^2 \frac{d\theta}{2\pi} |n\rangle\langle m|,$$

$Z \in \mathcal{B}(\mathbb{C})$, with $z = re^{i\theta} \in \mathbb{C} \cong \mathbb{R}^2$, $r = |z|$, and $|z\rangle = e^{-|z|^2/2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle$ is the coherent state. For all $s = 0, 1, \dots$, define a measure $\mu_s(X) = \int_X t^{s/2} e^{-t} dt$, $X \in \mathcal{B}([0, \infty))$, and note that the polynomials are dense in $L^2([0, \infty), \mu_s)$ since the (normalised) associated Laguerre polynomials $L_n^{s/2}(t)$, $n = 0, 1, \dots$, constitute an orthonormal basis of $L^2([0, \infty), \mu_s)$.

Denoting $t = r^2$ we get

$$r^{n+m} e^{i(n-m)\theta} = t^m e^{ik\theta} t^{l/2} = t^n e^{-il\theta} t^{l/2}$$

where $k = n - m$ and $l = m - n$. From Theorem 9.5 one sees that $\mathbf{G}^{(0)}$ belongs to $\text{ex}(\mathbf{O}(\mathbb{R}^2, \mathcal{H}))$ if

$$\begin{aligned} \int_0^\infty \left[\int_0^{2\pi} \lambda(\theta, t) e^{ik\theta} d\theta \right] t^m d\mu_k(t) &= 0, & k, m \in \mathbb{N}_0, \\ \int_0^\infty \left[\int_0^{2\pi} \lambda(\theta, t) e^{-il\theta} d\theta \right] t^n d\mu_l(t) &= 0, & l, n \in \mathbb{N}_0, \end{aligned}$$

implies $\lambda = 0$ (where λ is an essentially bounded (measurable) complex function). Since the polynomials are dense in $L^2([0, \infty), \mu_s)$ and both m and n run from 0 to ∞ , it follows that

$$\int_0^{2\pi} \lambda(\theta, t) e^{ik\theta} d\theta = 0$$

for all $k \in \mathbb{Z}$, i.e. $\lambda = 0$. Hence $\mathbf{G}^{(0)} \in \text{ex}(\mathbf{O}(\mathbb{R}^2, \mathcal{H}))$.

It can be shown that any $\mathbf{G}^\psi$ (with $\psi \in \mathcal{H}$, $\|\psi\| = 1$) belongs to $\text{ex}(\mathbf{O}(\mathbb{R}^2, \mathcal{H}))$ if and only if the Fourier transform of $z \mapsto |\langle z|\psi\rangle|^2$ is everywhere nonzero [12]. This implies that $\mathbf{G}^{(s)} \notin \text{ex}(\mathbf{O}(\mathbb{R}^2, \mathcal{H}))$ if $s = 1, 2, \dots$ $\triangleleft$

9.4 State Changes

The description of a quantum system, with associated Hilbert space $\mathcal{H}$, is based on the measurement outcome probabilities $\mathbf{E}_\varrho$ defined by states $\varrho \in \mathcal{S}(\mathcal{H})$ and observables $\mathbf{E} : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{H})$. Therefore, any change experienced by the system, be it due to autonomous dynamics or external influences, spontaneous or induced, if it does not destroy the system, is reflected in a transformation of these probabilities

$$\mathbf{E}_\varrho \mapsto m(\mathbf{E}_\varrho). \quad (9.20)$$

In the most general situation nothing specific can be said about this transformation. Assuming that the system preserves its identity during the course of its evolution and that the value spaces of observables are not influenced by that, then the probabilities $m(\mathbf{E}_\varrho)(X)$, $X \in \mathcal{A}$, are still of the form $p_{\tilde{\varrho}}^{\tilde{\mathbf{E}}}(X) = \text{tr}[\tilde{\varrho}\tilde{\mathbf{E}}(X)]$ for some observable $\tilde{\mathbf{E}} : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{H})$ and state $\tilde{\varrho} \in \mathcal{S}(\mathcal{H})$. Since any mapping $\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto \tilde{\varrho} \in \mathcal{S}(\mathcal{H})$ induces a mapping $\mathbf{E}(\mathcal{H}) \ni \hat{A} \mapsto \tilde{\hat{A}} \in \mathbf{E}(\mathcal{H})$ and conversely, the change (9.20) can be given either as a transformation of states or of observables, or effects, that is, in one of the (dual) forms $\tilde{m} : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$ in the Schrödinger picture or $\tilde{m} : \mathbf{E}(\mathcal{H}) \rightarrow \mathbf{E}(\mathcal{H})$ in the Heisenberg picture. Here we shall primarily follow the Schrödinger representation, switching to the Heisenberg picture whenever convenient.

Remark 9.4 The assumption that the system preserves its identity during the course of the change can easily be relaxed to the case where the new system is described by another Hilbert space $\mathcal{K}$. Then the state and observable changes are simply given by the maps $\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto \tilde{\varrho} \in \mathcal{S}(\mathcal{K})$ and $(\mathbf{E} : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{H})) \mapsto (\tilde{\mathbf{E}} : \mathcal{A} \rightarrow \mathbf{E}(\mathcal{K}))$ with the relevant properties describing the change in question. $\triangleleft$

The actual change dictates the structure of the mapping $\mathcal{S}(\mathcal{H}) \ni \varrho \mapsto \tilde{m}(\varrho) \in \mathcal{S}(\mathcal{H})$. Though here $\tilde{m}(\varrho)$ is assumed to be a state, that is, a positive trace one operator, it is useful to allow for the possibility that the trace of the state may decrease in the process, that is, we assume that for any $\varrho \in \mathcal{S}(\mathcal{H})$

$$\tilde{m}(\varrho) \geq 0, \quad \text{tr}[\tilde{m}(\varrho)] \leq 1. \quad (9.21)$$

Considering a mixture $w\rho_1 + (1 - w)\rho_2$ of any two states $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$, with weights $w, 1 - w$, the change $\tilde{m}$ gives us the two possibly nonnormalised states $\tilde{m}(w\rho_1 + (1 - w)\rho_2)$ and $w\tilde{m}(\rho_1) + (1 - w)\tilde{m}(\rho_2)$. In some processes these two states may be the same, in some others not. We now consider the case where

$$\tilde{m}(w\rho_1 + (1 - w)\rho_2) = w\tilde{m}(\rho_1) + (1 - w)\tilde{m}(\rho_2) \quad (9.22)$$

for any states $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$ and $0 \leq w \leq 1$

It is a simple exercise to show that a mapping $\tilde{m}$ with the properties (9.21) and (9.22) defines a unique operation in the sense of Definition 7.2.

Important classes of operations arise from projections $P \in \mathcal{P}(\mathcal{H})$, effect operators $E \in \mathcal{E}(\mathcal{H})$, and unitary operators $U \in \mathcal{U}(\mathcal{H})$:

$$\begin{aligned} \varrho &\mapsto P\varrho P = \Phi_P(\varrho) \\ \varrho &\mapsto \sqrt{E}\varrho\sqrt{E} = \Phi_E(\varrho) \\ \varrho &\mapsto U\varrho U^* = \Phi_U(\varrho). \end{aligned}$$

Operations of the form Φ_P and Φ_E typically arise from the measurement context, to be studied in greater detail in Chap. 10, whereas operations of the form Φ_U arise typically from symmetries, including autonomous time evolutions, as will be described below.

Let $\mathcal{O}(\mathcal{H})$ be the set of operations of a quantum system with the Hilbert space $\mathcal{H}$. Any operation $\Phi \in \mathcal{O}(\mathcal{H})$ defines a unique effect $\Phi^*(I) \in \mathcal{E}(\mathcal{H})$, and any effect $E \in \mathcal{E}(\mathcal{H})$ can be realised as $E = \Phi^*(I)$ for some operation $\Phi \in \mathcal{O}(\mathcal{H})$. The correspondence

$$\mathcal{O}(\mathcal{H}) \ni \Phi \mapsto \Phi^*(I) = E \in \mathcal{E}(\mathcal{H})$$

is many-to-one, accounting for the fact that any effect can be realised in many different processes.

The set $\mathcal{O}(\mathcal{H})$ of operations is closed under compositions and convex combinations, that is, for any $\Phi_1, \Phi_2 \in \mathcal{O}(\mathcal{H})$, $0 \leq w \leq 1$, $\Phi_1 \circ \Phi_2 \in \mathcal{O}(\mathcal{H})$ and $w\Phi_1 + (1 - w)\Phi_2 \in \mathcal{O}(\mathcal{H})$, and it is also equipped with a partial order: for any two operations $\Phi_1, \Phi_2 \in \mathcal{O}(\mathcal{H})$, $\Phi_1 \leq \Phi_2$ if and only if $\Phi_1^*(I) \leq \Phi_2^*(I)$. All these structures have obvious physical interpretations.

Wigner's Theorem

We recall next a theorem due to Wigner that characterises the class of operations that is commonly used to represent symmetry actions. This class may be described in various different ways, namely, as automorphisms of the sets of states $\mathcal{S}(\mathcal{H})$, pure states $\mathcal{P}_1(\mathcal{H})$, effects $\mathcal{E}(\mathcal{H})$, or sharp effects $\mathcal{P}(\mathcal{H})$ (see e.g. [13, 14]). For instance, we say that a map $s : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H})$ is a *state automorphism* if it is bijective and $s(w\rho_1 + (1 - w)\rho_2) = ws(\rho_1) + (1 - w)s(\rho_2)$ for all $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$, $w \in [0, 1]$, and a map $p : \mathcal{P}_1(\mathcal{H}) \rightarrow \mathcal{P}_1(\mathcal{H})$ is a *pure state automorphism* if p is a bijection and $\text{tr}[p(P_1)p(P_2)] = \text{tr}[P_1P_2]$ for all $P_1, P_2 \in \mathcal{P}_1(\mathcal{H})$.

The theorem relates symmetry actions to the family of operations of the form Φ_U , where U may be either unitary or antiunitary. A map $U : \mathcal{H} \rightarrow \mathcal{H}$ is *antiunitary* if it is antilinear (i.e. $U(\psi + c\varphi) = U\psi + \bar{c}U\varphi$, $\psi, \varphi \in \mathcal{H}, c \in \mathbb{C}$) and $\langle U\psi | U\varphi \rangle = \overline{\langle \psi | \varphi \rangle}$ for all $\psi, \varphi \in \mathcal{H}$. If U is antiunitary then U^2 is unitary.

Since any state automorphism is an operation, and any pure state automorphism is a restriction of a unique operation, we may formulate the Wigner theorem as follows (for a proof, see. e.g. [13]).

Theorem 9.6 *For any operation $\Phi \in \mathcal{O}(\mathcal{H})$, the following conditions are equivalent:*

- (i) Φ is a state automorphism;
- (ii) Φ is a pure state automorphism;
- (iii) Φ is of the form Φ_U for some unitary or antiunitary operator $U : \mathcal{H} \rightarrow \mathcal{H}$ (which is unique up to a phase factor $e^{i\theta}$, $\theta \in \mathbb{R}$).

As an application of Theorem 9.6, we formulate the time evolution of a closed quantum system. We let $\text{Aut}(\mathcal{S}(\mathcal{H}))$ denote the group of state automorphisms of the system and endow $\text{Aut}(\mathcal{S}(\mathcal{H}))$ with the weakest topology in which the functions $\text{Aut}(\mathcal{S}(\mathcal{H})) \ni s \mapsto \text{tr}[s(\varrho)E] \in [0, 1]$, $\varrho \in \mathcal{S}(\mathcal{H})$, $E \in \mathcal{E}(\mathcal{H})$, are continuous. As is well known, the evolution of a closed system can most conveniently be expressed as a continuous group homomorphism (also called a *symmetry action*),

$$\mathbb{R} \ni t \mapsto s_t \in \text{Aut}(\mathcal{S}(\mathcal{H})),$$

with the interpretation that $s_t(\varrho)$ is the state of the system at time t if ϱ is its state at time $t = 0$.

By Theorem 9.6, each s_t is of the form $s_t = \Phi_{U_t}$ for a unitary or antiunitary operator U_t which is unique up to a phase factor. On the other hand, any t can be written as $t = \frac{t}{2} + \frac{t}{2}$, so that $s_t = s_{t/2} \circ s_{t/2}$. Since the square of an antiunitary operator is unitary, it follows that each s_t is actually implemented by a unitary operator U_t , that is,

$$s_t(\varrho) = U_t \varrho U_t^*, \quad U_t \in \mathcal{U}(\mathcal{H}),$$

and U_t is determined up to a phase factor.

Let $\Sigma_0 = \mathcal{U}(\mathcal{H})/\simeq$ denote the set of the equivalence classes $[U]$ of unitary operators $U \in \mathcal{U}$ differing from each other by a phase factor, that is, for any $U, U' \in \mathcal{U}(\mathcal{H})$, $U' \in [U]$ if and only if $U' = e^{i\lambda}U$ for some $\lambda \in \mathbb{R}$. For any s_t there is a unique $\sigma_t \in \Sigma_0$. Let $\pi : \mathcal{U}(\mathcal{H}) \rightarrow \Sigma_0$ be the canonical projection. To pick up a single operator U_t corresponding to s_t one has to choose a section c for the projection π , that is, a (measurable) function $c : \Sigma_0 \rightarrow \mathcal{U}(\mathcal{H})$ for which $\pi(c(\sigma)) = \sigma$ for any $\sigma \in \Sigma_0$. The existence of such a function is guaranteed by a theorem due to Mackey (for instance, [15, Theorem 7.4]). If c and c' are two different sections for π , the resulting unitary operators $c(\sigma_t) = U_t$ and $c'(\sigma_t) = U'_t$ representing s_t differ only by a phase factor and therefore give the same state change $\varrho \mapsto U_t \varrho U_t^* = U'_t \varrho U'^*_t$.

By virtue of Wigner's theorem, a group homomorphism $\mathbb{R} \ni t \mapsto s_t \in \text{Aut}(\mathcal{S}(\mathcal{H}))$ together with a section $c : \Sigma_0 \rightarrow \mathcal{U}(\mathcal{H})$ define a (measurable) function $\mathbb{R} \ni t \mapsto$

$U_t \in \mathcal{U}(\mathcal{H})$ such that for any $t \in \mathbb{R}$ and $\varrho \in \mathcal{S}(\mathcal{H})$, $s_t(\varrho) = U_t \varrho U_t^*$. For $t_1, t_2 \in \mathbb{R}$, let U_{t_1} , U_{t_2} and $U_{t_1+t_2}$ be the unitary operators representing s_{t_1} , s_{t_2} and $s_{t_1+t_2}$, respectively. Since s is a group homomorphism one can conclude that

$$U_{t_1+t_2} = m(t_1, t_2) U_{t_1} U_{t_2}$$

for some $m(t_1, t_2) \in \mathbb{T}$, showing that the map $\mathbb{R} \ni t \mapsto U_t \in \mathcal{U}(\mathcal{H})$ is a projective representation of $\mathbb{R}$. Since all the multipliers (2-cocycles) m of $\mathbb{R}$ are exact, one may choose the section c such that the map $\mathbb{R} \ni t \mapsto U_t \in \mathcal{U}(\mathcal{H})$ is a (measurable) unitary representation (i.e. $m \equiv 1$). By a theorem of von Neumann [15, Lemma 5.28], the measurability of the map $t \mapsto U_t$ is equivalent to its continuity and so one may apply Stone's theorem (Theorem 5.6) to get

$$U_t = e^{itH/\hbar}$$

for a unique selfadjoint operator H acting in $\mathcal{H}$.

Using the differential form of Stone's theorem one obtains the familiar result that a continuous evolution of a closed system can be given by the equation

$$-i\hbar \frac{\partial}{\partial t} \varphi_t = H \varphi_t,$$

where $\varphi_t = e^{itH/\hbar} \varphi$ is the vector state of the system at time t if the state of the system at $t = 0$ is $\varphi \in \mathcal{D}(H)$. This is the abstract form of the Schrödinger equation. Similarly, if $\varrho_t = U_t \varrho U_t^*$, then by differentiating one gets the von Neumann-Liouville equation

$$-i\hbar \frac{\partial}{\partial t} \varrho_t = H \varrho_t - \varrho_t H,$$

valid under appropriate domain conditions. Assume that the physical system under consideration carries the Galilei symmetry, that is, there is a symmetry action of the Galilei group on the group of state automorphisms of the system, so that the evolution $t \mapsto s_t$ represents the time shifts $t \mapsto t + t'$. Then one can show that the evolution generating operator H represents the energy of the system (see, for instance, [16, 17], or [13]).

Conditioning and Lüders Operations

Operations of the form $\Phi_E : \varrho \mapsto \sqrt{E} \varrho \sqrt{E}$, $E \in \mathcal{E}(\mathcal{H})$, are special types of operations arising in a measurement context; they are called *Lüders operations*. Their full analysis is given in Chap. 10. Here we study the connection of the subclass of operations Φ_P , $P \in \mathcal{P}(\mathcal{H})$, with the notion of conditional probability, following the presentation in [18, 19].

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, and let $B \in \mathcal{A}$ be a fixed event for which $\mu(B) \neq 0$. The classical conditional probability of an event $A \in \mathcal{A}$, given the event B , is defined as

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)}, \quad (9.23)$$

and one easily verifies that $\mu(\cdot|B) : \mathcal{A} \rightarrow [0, 1]$ is a probability measure with the property

$$\mu(A|B) = \frac{\mu(A)}{\mu(B)} \quad (9.24)$$

for any $A \in \mathcal{A}$ such that $A \subset B$. In fact, if ν is any probability measure on $\mathcal{A}$ with the property (9.24), that is, $\nu(A) = \mu(A)/\mu(B)$ for any $A \in \mathcal{A}$, $A \subset B$, then $\nu = \mu(\cdot|B)$. This is to say that the conditional probability $\mu(\cdot|B)$ is the unique probability measure with the property (9.24).

Now consider the quantum case. Any state $\varrho \in \mathcal{S}(\mathcal{H})$ defines a generalised probability measure $m_\varrho : \mathcal{P}(\mathcal{H}) \rightarrow [0, 1]$ with $m_\varrho(P) = \text{tr}[\varrho P]$. For a fixed $\varrho \in \mathcal{S}(\mathcal{H})$, choose a sharp effect $R \in \mathcal{P}(\mathcal{H})$ for which $m_\varrho(R) \neq 0$. The function

$$\mathcal{P}(\mathcal{H}) \ni P \mapsto \frac{m_\varrho(P \wedge R)}{m_\varrho(R)} \in [0, 1]$$

is well defined but it is not a generalised probability measure (exercise).

On the other hand, if ν is any generalised probability measure on $\mathcal{P}(\mathcal{H})$ with the property

$$\nu(P) = \frac{m_\varrho(P)}{m_\varrho(R)}$$

for any $P \in \mathcal{P}(\mathcal{H})$, $P \leq R$, then by Gleason's theorem ν is of the form $\nu = m_{\varrho_R}$ for a unique $\varrho_R \in \mathcal{S}(\mathcal{H})$, and with a direct computation one may confirm that

$$\varrho_R = \frac{R\varrho R}{m_\varrho(R)} = \frac{1}{m_\varrho(R)} \Phi_R(\varrho). \quad (9.25)$$

For any state ϱ and for any $R \in \mathcal{P}(\mathcal{H})$, for which $m_\varrho(R) \neq 0$, one may define the generalised conditional probability of $P \in \mathcal{P}(\mathcal{H})$, given R , in the state ϱ as the number

$$m_\varrho(P|R) = m_{\varrho_R}(P) \quad (9.26)$$

consistently with the defining property (9.24) of the classical conditional probability.

The classical conditional probability $\mu(\cdot|B)$ is additive with respect to the disjoint partitions of the conditioning event. Indeed, if $B = \cup B_i$ is a partition of B , with $B_i \cap B_j = \emptyset$ for all $i \neq j$, and if $\mu(B_i) \neq 0$ for each i , then for any $A \in \mathcal{A}$,

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)} = \sum_i \frac{\mu(A \cap B_i)}{\mu(B)} = \sum_i \frac{\mu(B_i)}{\mu(B)} \mu(A|B_i),$$

that is, the measure $\mu(\cdot|B)$ is a convex combination of the measures $\mu(\cdot|B_i)$ with the weights $\frac{\mu(B_i)}{\mu(B)}$. Again, if $\nu : \mathcal{A} \rightarrow [0, 1]$ is any probability measure which has the property

$$\nu(A) = \frac{\mu(A)}{\mu(B)} \quad (9.27)$$

for all A in any interval $\emptyset \subset A \subset B_i$, then $\nu = \mu(\cdot|B)$.

Consider again the quantum case, with a given $\varrho \in \mathcal{S}(\mathcal{H})$ and $R \in \mathcal{P}(\mathcal{H})$ such that $m_\varrho(R) \neq 0$. Let (R_i) be a partition of R into mutually orthogonal elements, that is, $R_i R_j = 0$, and assume that $m_\varrho(R_i) \neq 0$ for each i . Then $m_\varrho(R) = \sum_i m_\varrho(R_i)$ but $\Phi_{\Sigma_i R_i} \neq \sum_i \Phi_{R_i}$. However, if $\nu : \mathcal{P}(\mathcal{H}) \rightarrow [0, 1]$ is any generalised probability measure which takes the values

$$\nu(P) = \frac{m_\varrho(P)}{m_\varrho(R)} \quad (9.28)$$

for any P in each interval $[0, R_i]$, then, by Gleason's theorem, ν is the unique generalised probability measure defined by the state

$$\varrho_{(R_i)} = \frac{1}{m_\varrho(R)} \sum_i \Phi_{R_i}(\varrho) = \sum_i \frac{m_\varrho(R_i)}{m_\varrho(R)} \frac{\Phi_{R_i}(\varrho)}{m_\varrho(R_i)}, \quad (9.29)$$

that is, the state $\varrho_{(R_i)}$ is the mixture of the states $\Phi_{R_i}(\varrho)/m_\varrho(R_i)$ with the weights $m_\varrho(R_i)/m_\varrho(R)$.

For any state ϱ and for any disjoint partition (R_i) of an $R \in \mathcal{P}(\mathcal{H})$, such that $m_\varrho(R_i) \neq 0$ for all i , one may define the generalised conditional probability of a $P \in \mathcal{P}(\mathcal{H})$, given the partition (R_i) , in the state ϱ as the number

$$m_\varrho(P|(R_i)) = m_{\varrho_{(R_i)}}(P) \quad (9.30)$$

consistently with the defining property (9.27) of the classical conditional probability.

The fact that the conditional probabilities (9.26) and (9.30) are different is an important feature of quantum mechanics; they differ by the so-called *interference term*.

9.5 Compound Systems

We have so far presented the quantum mechanical description of a closed physical system $\mathcal{S}$ which can be prepared in specific states and subjected to measurements but are otherwise isolated from the rest of the world. In particular, the observers as the systems carrying out the preparations and measurements are not part of the description.

In many situations, however, the isolation of the system $\mathcal{S}$ can be achieved only partially, the system cannot completely be separated from its environment $\mathcal{E}$. It is also frequently natural to consider the system as composed of some simpler parts $\mathcal{S}_1, \dots, \mathcal{S}_n$. In both cases it is assumed, however, that the whole, be it the system with its environment $(\mathcal{S}, \mathcal{E})$, or the system composed of subsystems $\mathcal{S} = (\mathcal{S}_1, \dots, \mathcal{S}_n)$, can, in principle, be described in terms of quantum mechanics with its basic rules.

The quantum theory of compound systems builds on such intuitive ideas and it describes the relations between the Hilbert space descriptions of the whole and its constituents. There are two new fundamental nonclassical features to be learned from the quantum theory of compound systems. One is entanglement: subsystems may be strongly correlated even when they are noninteracting. The other feature arises when the subsystems are identical: there is no way to keep track of identical systems—they cannot be identified continuously over time. Both of these features—entanglement and the indistinguishability of identical systems—have important consequences in quantum mechanics; among many other things, the former gives rise to quantum nonlocality and has been identified as a resource for quantum information processing, and the latter gives rise to the Pauli exclusion principle and is responsible for nothing less than the stability of matter.

The Tensor Product Structure

We consider a quantum system $\mathcal{S}$ with associated Hilbert space $\mathcal{H}$ and formulate now a set of *composition rules* which should hold if $\mathcal{S}$ is composed of two other systems $\mathcal{S}_1$ and $\mathcal{S}_2$, with their respective Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. The identification of $\mathcal{S}_1, \mathcal{S}_2$ as parts of $\mathcal{S}$ can be formalised in terms of an embedding of the set of properties (sharp or unsharp) of $\mathcal{S}_1$ into the set of properties of $\mathcal{S}$, that is, via a map $\mathbf{E}(\mathcal{H}_1) \ni E \mapsto h_1(E) \in \mathbf{E}(\mathcal{H})$ which preserves all the physically relevant structures of $\mathbf{E}(\mathcal{H}_1)$, and similarly for $\mathcal{S}_2$. In addition, any two properties (sharp or unsharp) of $\mathcal{S}_1$ and $\mathcal{S}_2$ must be jointly measurable when viewed as properties of $\mathcal{S}$, that is, for all properties $E \in \mathbf{E}(\mathcal{H}_1)$ and $F \in \mathbf{E}(\mathcal{H}_2)$, the properties $h_1(E), h_2(F) \in \mathbf{E}(\mathcal{H})$ must be jointly measurable, the exact sense of which will be explained in Chap. 11.

To describe the idea that complete subsystem descriptions should also exhaust the whole system we refer to the fact that a pure state is a maximal information state, in the sense that it can be identified with the collection of sharp properties that the system possesses in that state (that is, the properties that have probability equal to one). Therefore, we require that maximal information about the subsystems determines maximal information about the whole system; in other words, if both subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$ are in pure states $P[\varphi_1], P[\varphi_2]$ then also the whole system $\mathcal{S}$ is in a pure state $P[\varphi]$.

Example 9.5 Consider the case where $\mathcal{H}_1 = \mathbb{C}^n$ and $\mathcal{H}_2 = \mathbb{C}^m$. Let $\mathcal{H} = \mathbb{C}^{nm}$ and define the mappings $h_1(A) = A \otimes I_m$ and $h_2(B) = I_n \otimes B$, where $\otimes$ is the Kronecker product of matrices and I_m, I_n are the identity matrices. Clearly, the above mentioned composition rules hold in this model. In particular, if $A \in M_n(\mathbb{C})$ and $B \in M_m(\mathbb{C})$ are pure states, that is, one-dimensional projections, then their Kronecker product $A \otimes B \in M_{nm}(\mathbb{C})$ is a pure state. Therefore, in this case, the systems $\mathcal{S}_1$ and $\mathcal{S}_2$ with

the Hilbert spaces $\mathbb{C}^n$ and $\mathbb{C}^m$ can be interpreted as subsystems of a system $\mathcal{S}$ with the Hilbert space $\mathbb{C}^{nm}$. $\triangleleft$

For any Hilbert space $\mathcal{H}$, its dual $\mathcal{H}^*$ will be considered as a Hilbert space equipped with the inner product $\langle f | g \rangle = \langle f | g \rangle_{\mathcal{H}^*} = \langle \varphi_g | \varphi_f \rangle$, where by definition e.g. $\varphi_f \in \mathcal{H}$, for $f \in \mathcal{H}^*$, is the vector satisfying $\langle \varphi_f | \psi \rangle = f(\psi)$ for all $\psi \in \mathcal{H}$. We formulate the basic result of the theory of compound systems in the following theorem. Let $E^* : \mathcal{H}_1^* \rightarrow \mathcal{H}_1^*$ denote the dual of an effect $E : \mathcal{H}_1 \rightarrow \mathcal{H}_1$.

Theorem 9.7 ([20, 21]) *Assume that a quantum system $\mathcal{S}$ with the Hilbert space $\mathcal{H}$ consists of two subsystems with the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, where $\dim(\mathcal{H}_i) \geq 2$ for $i = 1, 2$. Then either $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and $h_1(E) = E \otimes I$ and $h_2(F) = I \otimes F$ for all $E \in \mathbf{E}(\mathcal{H}_1)$, $F \in \mathbf{E}(\mathcal{H}_2)$, or $\mathcal{H} = \mathcal{H}_1^* \otimes \mathcal{H}_2$ and $h_1(E) = E^* \otimes I$ and $h_2(F) = I^* \otimes F$ for all $E \in \mathbf{E}(\mathcal{H}_1)$, $F \in \mathbf{E}(\mathcal{H}_2)$.*

Remark 9.5 There are two more solutions for the composition rules, $\mathcal{H} = \mathcal{H}_1^* \otimes \mathcal{H}_2^*$ and $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2^*$, with the corresponding embeddings, but these are canonically equivalent to $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\mathcal{H}_1^* \otimes \mathcal{H}_2$, respectively. $\triangleleft$

The two solutions of Theorem 9.7 are inequivalent in the sense that the projection lattices $\mathbf{P}(\mathcal{H})$ as well as the sets of effects $\mathbf{E}(\mathcal{H})$ for the two cases $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and $\mathcal{H} = \mathcal{H}_1^* \otimes \mathcal{H}_2$ cannot be isomorphic since neither of the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is one-dimensional.

The known composition rules [20, 21] are not quite enough to fix the Hilbert space of the compound systems to be the tensor product of the Hilbert spaces of the constituent systems. Here we follow the traditional approach and assume that for a quantum system $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ with the Hilbert spaces $\mathcal{H}$ and $\mathcal{H}_1, \mathcal{H}_2$, respectively, one has $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. A generalisation for a quantum system consisting of finitely many, or even a countably infinite number of subsystems is straightforward. For notational simplicity we consider here only the generic case of a two-part system, which is also the basic setup for measuring processes.

With $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ the description of $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ follows the usual rules of the theory: the states of $\mathcal{S}$ are given as the positive trace one operators ϱ acting on $\mathcal{H}$, its observables are given as the semispectral measures $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$, and the measurement outcome probabilities are $X \mapsto \text{tr}[\varrho \mathbf{E}(X)]$.

The set $\mathbf{S}(\mathcal{H})$ of states of $\mathcal{S}$ contains states which are of the product form $\varrho_1 \otimes \varrho_2$, with $\varrho_i \in \mathbf{S}(\mathcal{H}_i)$, $i = 1, 2$, or finite or countably infinite convex combinations of such states. However, they do not exhaust the whole set of states: there are states $\varrho \in \mathbf{S}(\mathcal{H})$ which are not of that form. In particular, there are pure states $P[\xi] \in \mathbf{P}_1(\mathcal{H})$ which are not decomposable, that is, of the product form $P[\xi] = P[\xi_1] \otimes P[\xi_2] = P[\xi_1 \otimes \xi_2]$.

Consider any two observables $\mathbf{E}_i : \mathcal{A}_i \rightarrow \mathcal{L}(\mathcal{H}_i)$ of the subsystem $\mathcal{S}_i$, $i = 1, 2$. The map

$$\mathcal{A}_1 \times \mathcal{A}_2 \ni (X_1, X_2) \mapsto \mathbf{E}_1(X_1) \otimes \mathbf{E}_2(X_2) \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$$

is a well-defined operator bimeasure. The extension of it to a unique observable $\mathbf{E} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ is now obtained without the continuity assumption (D) of

Definition 4.7. This is due to the tensor product structure. Indeed, for any decomposable vectors $\Psi = \varphi \otimes \psi$, $\Phi = \xi \otimes \eta$ one has

$$\langle \Psi | E_1(X_1) \otimes E_2(X_2) \Phi \rangle_{\mathcal{H}} = \langle \varphi | E_1(X_1) \xi \rangle_{\mathcal{H}_1} \langle \psi | E_2(X_2) \eta \rangle_{\mathcal{H}_2},$$

and the map $(X_1, X_2) \mapsto \langle \varphi | E_1(X_1) \xi \rangle_{\mathcal{H}_1} \langle \psi | E_2(X_2) \eta \rangle_{\mathcal{H}_2}$ extends by basic measure theory to the product measure $\mu_{\Psi, \Phi} = E_{\varphi, \xi}^1 \times E_{\psi, \eta}^2$ on $\mathcal{A}_1 \otimes \mathcal{A}_2$. Now $\mu_{\Psi, \Phi}(\Omega_1 \times \Omega_2) = \langle \varphi | \xi \rangle_{\mathcal{H}_1} \langle \psi | \eta \rangle_{\mathcal{H}_2}$. (From now on, we do not write the subscript for the inner product to indicate the Hilbert space in question.) Extending the map $(\Psi, \Phi) \mapsto \mu_{\Psi, \Phi}$ by sesquilinearity and using the fact that the decomposable vectors generate the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, Proposition 2.1 allows one to conclude that there is an observable E such that $E(X_1 \times X_2) = E_1(X_1) \otimes E_2(X_2)$ for all $X_1 \in \mathcal{A}_1$, $X_2 \in \mathcal{A}_2$ so that E is the joint observable of the observables E_1 and E_2 . The measurement outcome probabilities for a product observable E in a product state $\varrho = \varrho_1 \otimes \varrho_2$, are simply of the product form $E_{\varrho} = E_{\varrho_1}^1 \times E_{\varrho_2}^2$.

Again, an observable E of $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ is not necessarily of the decomposable form. Such observables couple the subsystems and thus often appear as parts of an interaction between the constituents, a typical example being a Coulomb interaction between two particles, see Example 9.6.

An important special case of decomposable observables are the spectral measures of decomposable selfadjoint operators $A_1 \otimes A_2$. To recall the mathematical structure of such an observable, let A_1 and A_2 be the spectral measures of the operators A_1 and A_2 , with the domains $\mathcal{D}(A_1) \subset \mathcal{H}_1$ and $\mathcal{D}(A_2) \subset \mathcal{H}_2$. Let E be the joint spectral measure, that is, $E(X \times Y) = A_1(X) \otimes A_2(Y)$ for all $X, Y \in \mathcal{B}(\mathbb{R})$. Consider the product function $\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) = xy \in \mathbb{R}$ and the induced spectral measure $E^f : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Let A be the unique selfadjoint operator, with the domain $\mathcal{D}(A) \subset \mathcal{H}$, defined by E^f . Since the linear span of the vectors $\varphi \otimes \psi$ is dense in $\mathcal{H}$, and the domains of A_1 and A_2 are dense in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, the operator A is uniquely defined by

$$\begin{aligned} \langle \varphi \otimes \psi | A \xi \otimes \eta \rangle &= \int_{\mathbb{R}} \lambda \langle \varphi \otimes \psi | E^f(d\lambda) \xi \otimes \eta \rangle \\ &= \int_{\mathbb{R}^2} xy \langle \varphi \otimes \psi | A_1(dx) \otimes A_2(dy) \xi \otimes \eta \rangle \\ &= \int_{\mathbb{R}^2} xy \langle \varphi | A_1(dx) \xi \rangle \langle \psi | A_2(dy) \eta \rangle \\ &= \int_{\mathbb{R}} x \langle \varphi | A_1(dx) \xi \rangle \int_{\mathbb{R}} y \langle \psi | A_2(dy) \eta \rangle \\ &= \langle \varphi | A_1 \xi \rangle \langle \psi | A_2 \eta \rangle, \end{aligned}$$

with $\varphi \in \mathcal{H}_1$, $\psi \in \mathcal{H}_2$, $\xi \in \mathcal{D}(A_1)$, $\eta \in \mathcal{D}(A_2)$. The domain of A is given by the spectral theorem in the usual way. In particular, A is the unique selfadjoint operator which has the property $A \xi \otimes \eta = A_1 \xi \otimes A_2 \eta$ for all $\xi \in \mathcal{D}(A_1)$, $\eta \in \mathcal{D}(A_2)$.

Therefore, we denote A as $A_1 \otimes A_2$. Clearly, if A_1 and A_2 are bounded, then this definition coincides with the direct algebraic definition of $A_1 \otimes A_2$ of Sect. 3.7.

Example 9.6 Let a system $\mathcal{S}$ consist of two parts $\mathcal{S}_1$ and $\mathcal{S}_2$, with the Hilbert spaces $\mathcal{H}_1 = L^2(\mathbb{R}^3)$ and $\mathcal{H}_2 = L^2(\mathbb{R}^3)$, so that $\mathcal{H} = L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3) \simeq L^2(\mathbb{R}^3 \times \mathbb{R}^3)$. Consider the operator M_f acting in $\mathcal{H}$ and defined by the function $(\mathbf{x}, \mathbf{y}) \mapsto |\mathbf{x} - \mathbf{y}|^{-1}$. Let $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be the triples of position operators defined by the functions $\mathbf{x} \mapsto \mathbf{x}$ and $\mathbf{y} \mapsto \mathbf{y}$ and acting in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. The selfadjoint operator M_f can now be written in the form $M_f = |\mathbf{Q}_1 \otimes I - I \otimes \mathbf{Q}_2|^{-1}$, or just $M_f = |\mathbf{Q}_1 - \mathbf{Q}_2|^{-1}$ dropping the trivial parts of the operator. Clearly, this not of product form. $\triangleleft$

Subsystem States

Consider a system $\mathcal{S}$ with two parts $\mathcal{S}_1$ and $\mathcal{S}_2$, and assume that $\mathcal{S}$ is in a state $\varrho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. The requirement that $\mathcal{S}_1$ can be identified as a subsystem of $\mathcal{S}$ leads to the following requirement concerning the state ϱ_1 of $\mathcal{S}_1$: measuring an observable E_1 of $\mathcal{S}_1$ must be the same as measuring the observable $X \mapsto E_1(X) \otimes I_2$ of $\mathcal{S}$. In other words, with the whole system in state ϱ the subsystem $\mathcal{S}_1$ must be in a state ϱ_1 such that for each E_1 and for any $X \in \mathcal{A}_1$ the measurement outcome probabilities $\text{tr}[\varrho_1 E_1(X)]$ and $\text{tr}[\varrho(E_1(X) \otimes I_2)]$ are the same. It is a simple exercise to check that this requirement is equivalent to the condition

$$\text{tr}[\varrho_1 A_1] = \text{tr}[\varrho(A_1 \otimes I_2)] \quad (9.31)$$

for any $A_1 \in \mathcal{L}(\mathcal{H}_1)$. Similarly, the subsystem $\mathcal{S}_2$ must be in such a state ϱ_2 for which

$$\text{tr}[\varrho_2 A_2] = \text{tr}[\varrho(I_1 \otimes A_2)] \quad (9.32)$$

for all $A_2 \in \mathcal{L}(\mathcal{H}_2)$. These two equations define the subsystem states ϱ_1 and ϱ_2 as the partial traces (Sect. 3.7) of the full state ϱ , that is, by the mappings

$$\begin{aligned} \text{tr}_{\mathcal{H}_2} : \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2) &\rightarrow \mathcal{S}(\mathcal{H}_1), & \text{tr}_{\mathcal{H}_2}[\varrho] &= \varrho_1 \\ \text{tr}_{\mathcal{H}_1} : \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2) &\rightarrow \mathcal{S}(\mathcal{H}_2), & \text{tr}_{\mathcal{H}_1}[\varrho] &= \varrho_2. \end{aligned}$$

The dual maps of the partial traces are the embeddings $A_1 \mapsto A_1 \otimes I_2$ and $A_2 \mapsto I_1 \otimes A_2$.

The state ϱ of the compound system uniquely defines the states ϱ_1 and ϱ_2 of the subsystems. Clearly, their product $\varrho_1 \otimes \varrho_2$ fulfills the identification Eqs. (9.31) and (9.32). However, it is an important fact that the map $\varrho \mapsto (\varrho_1, \varrho_2)$ is not invertible, in general. There is one important special case where the subsystem states define the whole state and it is a reflection of the composition rule according to which maximal information on the subsystems define maximal information of the whole system. This result is an immediate consequence of Proposition 3.4.

Proposition 9.2 *A pure state ϱ of the compound system $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ is uniquely determined by the subsystem states $\varrho_1 = \text{tr}_{\mathcal{H}_2}[\varrho]$ and $\varrho_2 = \text{tr}_{\mathcal{H}_1}[\varrho]$ if and only if one of*

the states ϱ_1 and ϱ_2 is a pure state. In that case, they both are pure, that is, of the form $\varrho_1 = P[\varphi]$ and $\varrho_2 = P[\psi]$, and then also $\varrho = \varrho_1 \otimes \varrho_2 = P[\varphi] \otimes P[\psi] = P[\varphi \otimes \psi]$.

Consider now a pure state $P[\Psi]$, with $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ being a unit vector. Using the Schmidt decomposition of Ψ one immediately computes the subsystem states ϱ_1 and ϱ_2 to be

$$\varrho_1 = \sum_i \lambda_i^2 P[\xi_i], \quad \varrho_2 = \sum_i \lambda_i^2 P[\eta_i].$$

These are the spectral decompositions of the states ϱ_1 and ϱ_2 (cf. Remark 3.1). In particular, they have the same eigenvalues. Therefore, if one of them is pure, then both are pure and the polar decomposition of Ψ is the corresponding product state $\Psi = \xi \otimes \eta$. Moreover, the von Neumann entropies of ϱ_1 and ϱ_2 are the same: $S(\varrho_1) = S(\varrho_2) = -\text{tr}[\varrho_2 \ln \varrho_2] = -\sum_i \lambda_i^2 \ln(\lambda_i^2)$ (with the convention $0 \ln 0 = 0$).

Remark 9.6 In the quantum theory of measurement, which is the main application of the theory of compound systems in this monograph, the total system consists typically of two parts, the object system $\mathcal{S}$ to be measured, and the probe or the measuring apparatus $\mathcal{P}$. These systems play quite different roles and they are easily distinguishable from each other. Hence the additional constraints arising for identical systems are not relevant here. $\triangleleft$

Subsystem Dynamics

The states of the subsystems are obtained from the state of the whole system by the partial trace operation which is a completely positive, linear trace preserving mapping on $\mathcal{T}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ taking values in $\mathcal{T}(\mathcal{H}_1)$ and $\mathcal{T}(\mathcal{H}_2)$, respectively. Combined with any operation $\Phi \in \mathcal{O}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, one thus gets positive linear contracting mappings from $\mathcal{T}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ to $\mathcal{T}(\mathcal{H}_1)$ and $\mathcal{T}(\mathcal{H}_2)$ that are trace preserving and completely positive whenever Φ is such. The induced state change of the subsystem $\mathcal{S}_1$, say, is then $\text{tr}_{\mathcal{H}_2}[\varrho] \mapsto \text{tr}_{\mathcal{H}_2}[\Phi(\varrho)]$. This is particularly useful if the two subsystems are independent of each other before the change, that is, if they are initially in a product state $\varrho = \varrho_1 \otimes \varrho_2$. In that case the composite mapping

$$\varrho_1 \mapsto \varrho_1 \otimes \varrho_2 \mapsto \Phi(\varrho_1 \otimes \varrho_2) \mapsto \text{tr}_{\mathcal{H}_2}[\Phi(\varrho_1 \otimes \varrho_2)]$$

defines an operation in $\mathcal{O}(\mathcal{H}_1)$ which is trace preserving and completely positive whenever Φ is such. Similarly, one obtains an operation for $\mathcal{S}_2$. The analysis of measurement processes within quantum mechanics gives rise to many similar examples. Here we briefly describe the application of the partial trace operation to the dynamics of a subsystem.

Assume that the compound system $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$, such as an object with its environment or a probe, is isolated from the rest of the universe. The evolution of the whole system can then be given as a continuous group homomorphism

$$\mathbb{R} \ni t \mapsto s_t \in \text{Aut}(S(\mathcal{H}_1 \otimes \mathcal{H}_2)),$$

which allows a unitary implementation $\mathbb{R} \ni t \mapsto U_t \in \mathcal{U}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ as described in Sect. 9.4. The induced subsystem evolutions are given by the state transformation valued map

$$t \mapsto \text{tr}_{\mathcal{H}_2} \circ s_t,$$

and similarly for $\mathcal{S}_2$. Though this map is continuous and each $\text{tr}_{\mathcal{H}_2} \circ s_t$ is a completely positive trace preserving operation, the operations $\text{tr}_{\mathcal{H}_2} \circ s_t$ are not, in general, invertible. Neither do the maps $\text{tr}_{\mathcal{H}_2} \circ s_t, t \geq 0$, constitute, in general, a semigroup. Therefore the subsystem evolutions are typically not unitary.

As an illustration, assume that at $t = 0$ the subsystems are in a pure product state, given by a vector $\varphi \otimes \psi$, so that the state of the whole system at time t is given by the vector $U_t(\varphi \otimes \psi)$, which usually is not of the product form, and therefore the subsystems are, in general, in mixed states and their degree of mixedness usually varies in the course of time. In fact, if $U_t(\varphi \otimes \psi)$ were a product state for all t , then necessarily the evolution would also be decomposable, that is, of the form $U_t = U_t^1 \otimes U_t^2$, for all t , which would mean that the systems are dynamically independent.

Entanglement, Correlations, Separability

The fact that states and, in particular, pure states of a compound system $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ are not necessarily of the product form leads to a peculiar mutual statistical dependence of the subsystems referred to as *entanglement*. Mathematically this phenomenon is due to the fact that the Hilbert tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the closed linear span of the product vectors $\varphi \otimes \psi$. Therefore, entanglement is closely related to the notion of superposition of states (see Sect. 9.1). More precisely, a pure state is called entangled if it is not of product form. A general state is called *separable* if it can be decomposed as a mixture of states of product form, that is,

$$\varrho = \sum_i w_i \varrho_1^i \otimes \varrho_2^i. \quad (9.33)$$

Conversely, ϱ is called *entangled* if it is not separable.

If ϱ is separable, the subsystem states are $\sum_i w_i \varrho_1^i$ and $\sum_i w_i \varrho_2^i$. Separable states are known to admit a hidden variable model and therefore satisfy the Bell inequalities [22]. Peres [23] derives a simple algebraic test—the *positive partial transpose* (PPT) criterion—as a necessary condition for a state ϱ to allow such a decomposition and hence to be separable. The test is used in the negative: if the partial transpose of a density matrix ϱ is not positive, then the system thus prepared is nonseparable or (quantum) entangled.

It is an interesting and widely investigated problem to quantify the degree of entanglement of a state $\varrho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$; this can be done geometrically in terms of the trace norm distance $\|\varrho - \varrho_1 \otimes \varrho_2\|_1$ of ϱ from the product $\varrho_1 \otimes \varrho_2$ of the subsystem states $\varrho_1 = \text{tr}_{\mathcal{H}_2}[\varrho]$ and $\varrho_2 = \text{tr}_{\mathcal{H}_1}[\varrho]$.

Another, perhaps more natural way of assessing the difference between the states ϱ and $\varrho_1 \otimes \varrho_2$ is by way of comparing the measurement outcome probabilities of the

observables of product form $E_1 \otimes E_2$, that is, the deviation of the joint probability $p_\varrho^{E_1 \otimes E_2}$ from the product probability $p_{\varrho_1}^{E_1} \times p_{\varrho_2}^{E_2}$. This could be quantified in terms of the total variation norm of the probability measures, but this is neither practical nor very informative, except in some simple cases (see exercises).

Another simple quantification of the difference between ϱ and the associated $\varrho_1 \otimes \varrho_2$ is given by the normalised correlation coefficient in the case where the observables E_1 and E_2 are chosen to have value spaces $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

To formulate this notion, let $\mu : \mathcal{B}(\mathbb{R}^2) \rightarrow [0, 1]$ be an arbitrary probability measure with the margins μ_1 and μ_2 . Let $\epsilon_{12} = \int_{\mathbb{R}^2} xy d\mu$ be the $(1, 1)$ -moment of μ , and let $\epsilon_i = \int_{\mathbb{R}} x d\mu_i(x)$ and $\Delta_i = \left(\int_{\mathbb{R}} (\epsilon_i - x)^2 d\mu_i \right)^{1/2}$ denote the expectation and standard deviation of μ_i , $i = 1, 2$. Recalling that $\Delta_i = 0$ exactly when μ_i is the Dirac measure at ϵ_i , we assume that $\Delta_i \neq 0$ for both $i = 1, 2$. The normalised correlation coefficient of μ is defined as

$$\text{cor}(\mu) = \frac{\epsilon_{12} - \epsilon_1 \epsilon_2}{\Delta_1 \Delta_2}; \quad (9.34)$$

the Cauchy–Schwarz inequality gives that $-1 \leq \text{cor}(\mu) \leq 1$. The case of *strong correlation* $\text{cor}(\mu) = \pm 1$ occurs exactly when the marginal measures μ_i are linearly dependent, in the sense that the coordinate projections $\pi_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ which induce μ_i are linearly dependent through $\pi_1 = \pm \frac{\Delta_1}{\Delta_2}(\pi_2 - \epsilon_2) + \epsilon_1$, that is, for all $(x, y) \in \mathbb{R}^2$, $x = \pm \frac{\Delta_1}{\Delta_2}(y - \epsilon_2) + \epsilon_1$. Clearly, $\text{cor}(\mu) = 0$ if and only if $\epsilon_{12} = \epsilon_1 \epsilon_2$. This is the case when $\mu = \mu_1 \times \mu_2$.

The degree of entanglement of a state $\varrho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ can now be measured by the correlations $\text{cor}(p_\varrho^{E_1 \otimes E_2})$ it produces for the (real) observables E_i of the subsystems $\mathcal{S}_i$, $i = 1, 2$. It is a remarkable nonclassical feature that any pure state of a compound system which is not of the product form gives strong correlations for some pair of local observables.

The entanglement in a state $\varrho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ manifests itself in specific features of the joint probabilities of the form $p_\varrho^{E_1 \otimes E_2}$, where E_1 and E_2 are any two subsystem observables which by their mutual commutativity (more precisely, that of $h_1(E_1)$ and $h_2(E_2)$) can always be measured jointly on $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$, with joint observable $E_1 \otimes E_2$. Two commuting observables always possess a joint observable. Therefore, one may ask if the structure of $\mathcal{S}$ as consisting of two parts $(\mathcal{S}_1, \mathcal{S}_2)$, that is, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, has a special bearing on the nature of the quantum correlations inherent in such joint probabilities. This question is known as Tsirelson’s problem. If the system $\mathcal{S}$ is a finite level system, the answer to this problem is negative but in general the solution is not known [24–26].

The Role of Complete Positivity

The notion of complete positivity has already occurred in several instances. In Sect. 7.7 (Theorem 7.14) it was shown that completely positive instruments admit a particularly suggestive unitary dilation, a fact which turns out to be the key for the main result of Chap. 10—the existence of a measurement model for an arbitrary observable. Moreover, as pointed out in Sect. 9.4, the evolution of a closed quan-

tum system $\mathcal{S}$ can be given as a group homomorphism $\mathbb{R} \ni t \mapsto s_t \in \text{Aut}(\mathcal{S}(\mathcal{H}))$. From the Wigner theorem as well as from the properties of the time group $(\mathbb{R}, +)$ it follows that the maps s_t can be implemented unitarily, that is, they are of the form $s_t(\varrho) = U_t \varrho U_t^*$ for a unique unitary U_t (up to a phase factor). Therefore state transformations s_t are completely positive. If $\mathcal{S}$ is a compound system, the induced subsystem dynamics are also completely positive, though not in general unitarily implementable.

There is a simple and general argument showing that complete positivity is a natural requirement for the dynamics of any physical system. Consider a system $\mathcal{S}$ whose evolution is given by a family of state transformations s_t . Let $\mathcal{S}'$ be another system which does not interact with $\mathcal{S}$. The total evolution is then necessarily of the product form, and the nature of the evolution of $\mathcal{S}$ does not depend on the nature of the evolution of $\mathcal{S}'$. Therefore, one could assume that $\mathcal{S}'$ does not change in time, so that its (constant) evolution is given by the unit operation ι' . The dynamics of the compound system $(\mathcal{S}, \mathcal{S}')$ is then given by $s_t \otimes \iota'$, which means that $s_t \otimes \iota'$ is a state transformation, and, in particular, positive. Since $\mathcal{S}'$ is an arbitrary system, this implies that s_t must be completely positive.

9.6 Exercises

1. Let $\mathcal{H} = \mathbb{C}^2$. Show that any selfadjoint operator can be written in the form

$$A = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma}, \quad (a_0, \mathbf{a}) \in \mathbb{R}^4,$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the *Pauli matrices*.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Determine the spectral decomposition of such an A .

2. Show that any state $\varrho \in \mathcal{S}(\mathbb{C}^2)$ is of the form

$$\varrho = \varrho_{\mathbf{n}} = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma}), \quad \mathbf{n} \in \mathbb{R}^3, \quad \|\mathbf{n}\| \leq 1,$$

and that this state is pure exactly when $\mathbf{n}$ is a unit vector.

3. Let $B(0, 1)$ be the unit ball of $\mathbb{R}^3$. Show that the map $\mathbf{n} \mapsto \varrho_{\mathbf{n}}$ is a bijection $B(0, 1) \rightarrow \mathcal{S}(\mathbb{C}^2)$ which preserves the convex structures of the two sets.
4. The spin quantity of a spin $-\frac{1}{2}$ object in direction $\mathbf{a} \in \mathbb{R}^3$, $\|\mathbf{a}\| = 1$, is given by the selfadjoint operator $S_{\mathbf{a}} = \frac{\hbar}{2} \mathbf{a} \cdot \boldsymbol{\sigma}$, with the spectral decomposition $S_{\mathbf{a}} = \frac{\hbar}{2} \varrho_{\mathbf{a}} - \frac{\hbar}{2} \varrho_{-\mathbf{a}}$ (Exercises 1 and 2). Compute the measurement outcome probabilities for $S_{\mathbf{a}}$ in a spin state $\varrho_{\mathbf{n}}$. Show also that $\text{Exp}(S_{\mathbf{a}}, \varrho_{\mathbf{n}}) = \frac{\hbar}{2} \mathbf{a} \cdot \mathbf{n}$ and $\text{Var}(S_{\mathbf{a}}, \varrho_{\mathbf{n}}) = (\frac{\hbar}{2})^2 [1 - (\mathbf{a} \cdot \mathbf{n})^2]$.

5. Prove the result of Hadjisavvas [5]: a vector state φ is a component of a state ϱ if and only if $\varphi \in \varrho^{1/2}(\mathcal{H})$.
6. Show that the family $(P[\eta_i], w_i)_{i \in I}$ of Theorem 9.2 is a decomposition of the state ϱ .
7. Show that the map $A \mapsto A^\perp$ is not an orthocomplementation in $E(\mathcal{H})$. Verify also the result (9.5), which shows that the restriction of this map to $\mathbf{P}(\mathcal{H})$ is an orthocomplementation (see Remark 2.3).
8. Show that an effect $A \in E(\mathcal{H})$ is a projection if and only if its set of upper bounds u.b. $\{A, A^\perp\}$ in the order given by the sum contains only the identity operator I .
9. Assume that (9.8) holds for any state ϱ and its possible decompositions. Show that the map (9.7) extends to a positive linear functional $\Phi_X^m : \mathcal{T}_s(\mathcal{H}) \rightarrow \mathbb{R}$ such that $\Phi_X^m(\varrho) = p_\varrho^m(X)$ for any state ϱ .
10. Prove the set inclusion (9.18).
11. Show that a mapping $\tilde{m}$ with the properties (9.21) and (9.22) defines a unique operation.
12. Show that any pure state automorphism is the restriction of a unique operation.
13. Show that the square of an antiunitary operator is linear and unitary.
14. The choice of the unitary operator U_t representing s_t is unique up to a phase factor. Show that this implies that the selfadjoint operator H which generates the evolution is unique up to a constant, that is, if H' is another selfadjoint operator generating the same evolution, then $H' = H + hI$ for some real number $h \in \mathbb{R}$.
15. Show that the map (9.23) is a probability measure with the property (9.24).
16. Show that the condition $\text{tr}[\varrho_1 P_1] = \text{tr}[\varrho P_1 \otimes I_2]$ for any $P_1 \in \mathcal{P}(\mathcal{H}_1)$ is equivalent to the condition (9.31).
17. Prove that the normalised correlation coefficient (9.34) satisfies the inequality $-1 \leq \text{cor}(\mu) \leq 1$.
18. Show that for two pure states ϱ and ϱ' one gets $\|\varrho - \varrho'\|_1 = 2\sqrt{1 - \text{tr}[\varrho\varrho']}$.
19. Consider a composition of two two-level systems with the Hilbert spaces $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$, and assume that the compound system is in the Einstein–Podolsky–Rosen–Bohm, or in short, Bell state $\Phi = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle)$. Here $|\pm\rangle$ denote the eigenvectors of S_z and, for instance, $|+-\rangle = |+\rangle \otimes |-\rangle$. Show that Φ is rotation invariant. Compute the distance $\|P[\Phi] - \varrho_1 \otimes \varrho_2\|_1$. Consider any two spin quantities $S_a = \frac{1}{2}\mathbf{a} \cdot \boldsymbol{\sigma}$ and $S_b = \frac{1}{2}\mathbf{b} \cdot \boldsymbol{\sigma}$ of the respective subsystems ($\hbar = 1$). Compute the total variation norm of the difference of the probability measures $p_{P[\Phi]}^{S_a \otimes S_b}$ and $p_{\varrho_1}^{S_a} \times p_{\varrho_2}^{S_b}$ as well as the normalised correlation of the joint probability measure $p_{P[\Phi]}^{S_a \otimes S_b}$.
20. Let $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ be a nondecomposable vector state. Construct subsystem observables E_1 and E_2 which are strongly correlated in this state.
21. Assume that E_1 and E_2 are the spectral measures of the bounded selfadjoint operators A_1 and A_2 , respectively. Show that

$$\text{cor}(p_{\varrho}^{A_1 \otimes A_2}) = \frac{\text{tr}[\varrho A_1 \otimes A_2] - \text{tr}[\varrho_1 A_1] \text{tr}[\varrho_2 A_2]}{\Delta_1 \Delta_2}$$

where $\varrho_i = \text{tr}_{\mathcal{H}_i} \varrho$ and $\Delta_i^2 = \text{tr}[\varrho_i A_i^2] - \text{tr}[\varrho_i A_i]^2 \neq 0$.

22. Let $\mathcal{H}_1 = \mathcal{H}_2 = L^2(\mathbb{R})$ and a_k be the lowering operator of $\mathcal{H}_k$, that is, $a_k = \frac{1}{\sqrt{2}}(Q_k + iP_k)$, where Q_k, P_k are the position and momentum operators of $\mathcal{H}_k$, $k = 1, 2$. Recall that the Hermite functions $h_n \equiv |n\rangle$ form an orthonormal basis of $L^2(\mathbb{R})$ and, for instance $a|0\rangle = 0$ and $a|n\rangle = \sqrt{n}|n-1\rangle$, $n \geq 1$. We associate the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ with the system of two dynamically independent particles moving on the line $\mathbb{R}$. Define the two mode squeezing operator

$$S(\zeta) = \exp(\bar{\zeta} a_1^* \otimes a_2^* - \zeta a_1 \otimes a_2)$$

where $\zeta = re^{-i\theta} \in \mathbb{C}$, $r = |\zeta|$, is the squeezing parameter. Show that in the two mode squeezed vacuum state

$$\psi_{\zeta} = S(\zeta)|0\rangle \otimes |0\rangle = (\cosh \theta)^{-1} \sum_{n=0}^{\infty} \beta^n e^{in\theta} |n\rangle \otimes |n\rangle$$

where we have $\beta = \tanh r$, the value of the correlation is

$$\text{cor}(p_{\psi_{\zeta}}^{Q_1 \otimes Q_2}) = \frac{2\beta}{1 + \beta^2} \cos \theta.$$

What happens in the limit $r \rightarrow \infty$ (especially when $\theta = 0$)?

References

1. von Neumann, J.: Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1927**, 245–272 (1927)
2. Dirac, P.: The Principles of Quantum Mechanics, 4th edn. Oxford, Clarendon Press (1981)
3. von Neumann, J.: Mathematische Grundlagen der Quantenmechanik. Die Grundlehren der mathematischen Wissenschaften, Band 38. Springer-Verlag, Berlin (1968). (Reprint of the 1932 original)
4. Hughston, L.P., Jozsa, R., Wootters, W.K.: A complete classification of quantum ensembles having a given density matrix. Phys. Lett. A **183**(1), 14–18 (1993)
5. Hadjisavvas, N.: Properties of mixtures on non-orthogonal states. Lett. Math. Phys. **5**, 327–332 (1981). July
6. Cassinelli, G., De Vito, E., Leviero, A.: On the decompositions of a quantum state. J. Math. Anal. Appl. **210**(2), 472–483 (1997)
7. Davies, E.B.: Quantum Theory of Open Systems. Academic Press, New York (1976)
8. Lahti, P., Mączyński, M.: Partial order of quantum effects. J. Math. Phys. **36**(4), 1673–1680 (1995)
9. Gleason, A.M.: Measures on the closed subspaces of a Hilbert space. J. Math. Mech. **6**, 885–893 (1957)

10. R. S. Ingarden. Information theory and thermodynamics of light. I. Foundations of information theory. *Fortschr. Physik*, **12**, 567–594 (1964)
11. Holevo, A.S.: *Statistical Structure of Quantum Theory*. Lecture Notes in Physics, vol. 67. Monographs. Springer, Berlin (2001)
12. Heinosaari, T., Pellonpää, J.-P.: Generalized coherent states and extremal positive operator valued measures. *J. Phys. A* **45**(24), 244019 (2012)
13. Cassinelli, G., De Vito, E., Lahti, P., Leviero, A.: *The Theory of Symmetry Actions in Quantum Mechanics—With an application to the Galilei group*. Lecture Notes in Physics, vol. 654. Springer-Verlag, Berlin (2004)
14. Molnár, L.: *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces*, Lecture Notes in Mathematics, vol. 1895. Springer, Heidelberg (2007)
15. Varadarajan, V.S.: *Geometry of Quantum Theory*, 2nd edn. Springer-Verlag, New York (1985)
16. Mackey, G.W.: *Unitary Group Representations in Physics, Probability, and Number Theory*, 2nd edn. Advanced Book Classics. Addison-Wesley Publishing Company Advanced Book Program, Redwood City (1989)
17. Mackey, G.W.: Axiomatics of particle interactions. In *Proceedings of the International Quantum Structures Association, Part I (Castiglione, 1992)*, vol. 32, pp 1643–1659 (1993)
18. Cassinelli, G., Zanghì, N.: Conditional probabilities in quantum mechanics. I. Conditioning with respect to a single event. *Nuovo Cim. B* (11), **73**(2), 237–245 (1983)
19. Cassinelli, G., Zanghì, N.: Conditional probabilities in quantum mechanics. II. Additive conditional probabilities. *Nuovo Cim. B* (11), **79**(2), 141–154 (1984)
20. Aerts, D., Daubechies, I.: Physical justification for using the tensor product to describe two quantum systems as one joint system. *Helv. Phys. Acta*, **51**(5-6), 661–675 (1979), (1978)
21. Pulmannová, S.: Tensor products of Hilbert space effect algebras. *Rep. Math. Phys.* **53**(2), 301–316 (2004)
22. Werner, R.: Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40**, 4277–4281 (1989)
23. Peres, A.: Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413–1415 (1996)
24. Cirel'son, B.S.: Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4**(2), 93–100 (1980)
25. Khalfin, L.A., Tsirelson, B.S.: Quantum and quasiclassical analogs of Bell inequalities. In: *Symposium on the Foundations of Modern Physics (Joensuu, 1985)*, pp. 441–460. World Scientific Publishing, Singapore (1985)
26. Scholz, V.B., Werner, R.F.: Tsirelson's problem (2008). [arXiv:0812.4305](https://arxiv.org/abs/0812.4305)

Chapter 10

Measurement

The purpose of a measurement is the determination of properties of the physical system under investigation. In this sense the general conception of measurement is that of an unambiguous comparison: the object system $\mathcal{S}$, *prepared* in a state ϱ , is brought into a suitable contact—a *measurement coupling*—with another, independently prepared system, the *measuring apparatus*, from which the *result* related to the measured observable is *determined* by *reading* the value of the *pointer*. It is the task of the quantum theory of measurement to investigate measuring processes as physical processes subject to quantum mechanics, and to describe how the stages and steps described above are realised.

In treating the measuring apparatus as a quantum mechanical system, measurement theory is part of the quantum theory of compound systems. As a reaction to the unsolved quantum measurement problem, there is a widespread view that at least part of the macroscopic aspects of an apparatus must be described in terms of classical physics. In order to avoid any interpretational commitment beyond the minimal interpretation, the considerations of this chapter will be based on the assumption that at least part of the apparatus is microscopic and can be described in terms of quantum mechanics. This part will be called the *probe*, and for the purpose of *modelling* the measurement process as a transfer of information from the object system to the probe, we will refer to the probe observable into which the information about the measured observable is to be transcribed as the *pointer observable*. We will not discuss here the subsequent processes of amplification and recording of the value of the probe observable into some macroscopic pointer observable; but we return to the measurement problem in Chap. 22.

10.1 Measurement Schemes

General Concept

The elements required for modelling a measurement of an observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ of the system $\mathcal{S}$ are the following: a *probe* $\mathcal{P}$, with its Hilbert space $\mathcal{K}$, prepared initially in a state σ , a *pointer observable* $Z \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{K})$ and a unitary map $U : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{K}$ which models the *measurement coupling* between the system and the probe. If $\varrho \in \mathcal{S}(\mathcal{H})$ is the initial state of the system then $\varrho \otimes \sigma$ is the initial state of $\mathcal{S} + \mathcal{P}$, under the assumption that before the measurement the system and the probe are dynamically and probabilistically independent of each other. The final state of $\mathcal{S} + \mathcal{P}$, that is, the state of $\mathcal{S} + \mathcal{P}$ after the measurement is then $U(\varrho \otimes \sigma)U^*$. The description of a measurement will thus be based on a quadruple $(\mathcal{K}, Z, \sigma, U)$, also denoted as $\mathcal{M}$, and we refer to it as a *measurement scheme* for the system $\mathcal{S}$.

The final state $U(\varrho \otimes \sigma)U^*$ of $\mathcal{S} + \mathcal{P}$ determines uniquely the final states of its subsystems $\mathcal{S}$ and $\mathcal{P}$ as the reduced states,

$$\begin{aligned}\varrho^f &= \text{tr}_{\mathcal{K}}[U(\varrho \otimes \sigma)U^*], \\ \sigma^f &= \text{tr}_{\mathcal{H}}[U(\varrho \otimes \sigma)U^*].\end{aligned}$$

It follows, in particular, that each initial state ϱ of the measured system determines uniquely the probability distribution of the pointer observable Z in the final probe state σ^f . This, in turn, entails that every measurement scheme defines a unique observable of the measured system.

Proposition 10.1 *Given a measurement scheme $(\mathcal{K}, Z, \sigma, U)$, the mapping*

$$\mathcal{A} \ni X \mapsto Z_{\sigma^f}(X) = \text{tr}[\sigma^f Z(X)] = \text{tr}[U(\varrho \otimes \sigma)U^*(I \otimes Z(X))] \in [0, 1] \quad (10.1)$$

is a probability measure for all states $\varrho \in \mathcal{S}(\mathcal{H})$ and it determines a unique observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ of $\mathcal{S}$ such that for all $\varrho \in \mathcal{S}(\mathcal{H})$, $X \in \mathcal{A}$,

$$E_{\varrho}(X) = Z_{\sigma^f}(X). \quad (10.2)$$

This result confirms the fact that a measurement scheme $\mathcal{M}$ defines a unique observable E of $\mathcal{S}$ which satisfies the *probability reproducibility condition* (10.2) with respect to $\mathcal{M}$. We say that E is the observable measured by $\mathcal{M}$ and refer to $\mathcal{M}$ as a measurement of E . To emphasise the role of $\mathcal{M}$, we occasionally write $E^{\mathcal{M}}$ for the measured observable defined by (10.2).

Corollary 7.1 on the Naimark dilation of a semispectral measure entails that every observable of a quantum system can be measured. This result is the first fundamental theorem of the quantum theory of measurement.

Theorem 10.1 *For any observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ of a quantum system S there is a measurement scheme $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$, such that E is the observable measured by this scheme, that is, $E = E^{\mathcal{M}}$. Moreover, the pointer observable $\mathbf{Z}$ can be chosen to be sharp and the initial probe state σ to be a pure state $P[\phi]$; in this case E assumes the explicit form*

$$E(X) = V_{\phi}^* U^* (I \otimes Z(X)) U V_{\phi},$$

where V_{ϕ} is the embedding $\varphi \mapsto \varphi \otimes \phi$.

The probability reproducibility condition (10.2) is considered here to be the first condition for $\mathcal{M}$ to serve as a measurement of E . It only guarantees that the measurement outcome probabilities E_{ϱ} for E in an initial object state ϱ are reproduced in the distribution Z_{σ^f} of the pointer values in the final probe state σ^f . This statistical level of description is sufficient in some applications, such as the problem of state determination that is discussed in Chap. 18. However, one often wishes to get more information on the measured system S , for instance, the value of the measured observable E after the measurement; then one needs to impose further requirements on the measurement scheme $\mathcal{M}$. These additional features of $\mathcal{M}$ typically include various repeatability, ideality and correlation properties, some of which will be studied in Sects. 10.5 and 10.6.

Finally, the quantum mechanical description of measurements should account for the fact that the—generally indeterministic—measurement outcomes are realised in the form of definite “pointer positions” indicating the actual value of the pointer observable, and thus possibly also the value of the measured observable after the measurement. The question of how to explain the occurrence of definite measurement outcomes in view of the entanglement created by the measurement dynamics between object system, apparatus, and possibly the environment of both, constitutes the so-called *quantum measurement problem*, or *objectification problem*; it will be presented (albeit not solved) in Chap. 22.

Remark 10.1 It is sometimes useful to allow pointers whose value spaces differ from that of E . Then one needs a *pointer function* to relate the possible values of E to those of $\mathbf{Z}$. If $(\mathcal{E}, \mathcal{B})$ and $(\Omega, \mathcal{A})$ are the value spaces of $\mathbf{Z}$ and E , respectively, the pointer function is a measurable map $g : \mathcal{E} \rightarrow \Omega$, and the measurement scheme is a 5-tuple $(\mathcal{K}, \mathbf{Z}, \sigma, U, g)$ satisfying the condition $E_{\varrho}(X) = Z_{\sigma^f}(g^{-1}(X))$ for all $\varrho \in \mathcal{S}(\mathcal{H})$ and $X \in \mathcal{A}$. Denoting by $\mathbf{M}$ the observable on $(\mathcal{E}, \mathcal{B})$ defined by the measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U)$, then, equivalently, the measured observable E is a function of $\mathbf{M}$ in the sense that $E = \mathbf{M} \circ g^{-1} \equiv \mathbf{M}^g$. $\triangleleft$

Remark 10.2 Occasionally it is convenient to *discretise* measurements of a continuous observable. Let E be the observable on $(\Omega, \mathcal{A})$ defined by a measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U)$. We define a *reading scale* $\mathcal{R}$ to be a partition of Ω into mutually disjoint sets $X_i \in \mathcal{A}, i \in \mathbb{I} \subset \mathbb{N}$. Such a partition induces the measurable function $r : \Omega \rightarrow \mathbb{I}, X_i \ni x \mapsto r(x) = i$, so that $r^{-1}(\{i\}) = X_i$. The 5-tuple $(\mathcal{K}, \mathbf{Z}, \sigma, U, r)$ then defines

the discrete observable E' , which has the property $E'(\{i\}) = E(X_i) \equiv E_i$. The reading scale $\mathcal{R}$ also induces the discrete pointer observable $i \mapsto Z(X_i) \equiv Z_i = Z'(\{i\})$, so that the measurement schemes $(\mathcal{K}, Z, \sigma, U, r)$ and $(\mathcal{K}, Z', \sigma, U)$ are equivalent in the sense that they both induce the same discrete observable $i \mapsto E_i = \bar{E}(X_i)$. $\triangleleft$

Calibration Condition

A minimal demand a measurement should fulfil is that it should exhibit unequivocally what is the case; that is, if an observable E has a definite value, a measurement of that observable ought to determine its value with (probabilistic) certainty:

$$E_\varrho(X) = 1 \implies Z_{\sigma^f}(X) = 1. \quad (10.3)$$

The condition $E_\varrho(X) = 1$ is equivalent to $E(X)\varrho = \varrho$, and similarly for the pointer probability. Therefore, the calibration condition (10.3) can equivalently be stated in a non-probabilistic way.

Definition 10.1 A measurement scheme $\mathcal{M}$ satisfies the *calibration condition* with respect to an observable E if for any $X \in \mathcal{A}$ and for any $\varrho \in \mathcal{S}(\mathcal{H})$,

$$E(X)\varrho = \varrho \implies Z(X)\sigma^f = \sigma^f. \quad (10.4)$$

In general, the calibration condition will be nontrivial only if the observable can assume definite values; that is, for each $X \in \mathcal{A}$ the effect $E(X)$ should have eigenvalue 1. This is fulfilled for sharp observables, and in this case the calibration condition and probability reproducibility condition are equivalent.

Theorem 10.2 Let E be a sharp observable of the system $\mathcal{S}$. A measurement scheme $(\mathcal{K}, Z, \sigma, U)$ is a measurement of E if and only if it satisfies the calibration condition with respect to E .

Proof We only need to show the nontrivial part, that the calibration condition implies the probability reproducibility condition. Assume that $(\mathcal{K}, Z, \sigma, U)$ fulfils the calibration condition with respect to a sharp observable E . Consider first a vector state $\varrho = P[\varphi]$ of $\mathcal{S}$. If $E(X)P[\varphi] = P[\varphi]$, then $\text{tr}[\sigma^f Z(X)] = 1$. Next, if $E(X)P[\varphi] = 0$, then, by applying (10.3) to the complement set X^c of X , one also has that $\text{tr}[\sigma^f Z(X)] = 0$. Assume now that $P[\varphi]$ is such that $0 \neq E_\varphi(X) \neq 1$. We will use the identity

$$\begin{aligned} P[\varphi] &= E(X)P[\varphi]E(X) + E(X^c)P[\varphi]E(X^c) \\ &\quad + E(X)P[\varphi]E(X^c) + E(X^c)P[\varphi]E(X) \\ &= E_\varphi(X)\varrho_X + E_\varphi(X^c)\varrho_{X^c} + \varrho_{X X^c} + \varrho_{X^c X}, \end{aligned}$$

where

$$\begin{aligned}\varrho_X &= \mathbf{E}_\varphi(X)^{-1} \mathbf{E}(X) P[\varphi] \mathbf{E}(X), \\ \varrho_{X^c} &= \mathbf{E}_\varphi(X^c)^{-1} \mathbf{E}(X^c) P[\varphi] \mathbf{E}(X^c), \\ \varrho_{XX^c} &= \mathbf{E}(X) P[\varphi] \mathbf{E}(X^c).\end{aligned}$$

We compute:

$$\begin{aligned}\mathrm{tr}[\sigma^f \mathbf{Z}(X)] &= \mathrm{tr}[U(P[\varphi] \otimes \sigma) U^* I \otimes \mathbf{Z}(X)] \\ &= \mathbf{E}_\varphi(X) + 2\mathrm{Re}\{\mathrm{tr}[U(\varrho_{XX^c} \otimes \sigma) U^* I \otimes \mathbf{Z}(X)]\}\end{aligned}\quad (10.5)$$

where in the last equality we have applied the calibration condition to ϱ_X and ϱ_{X^c} , e.g.,

$$1 = \mathrm{tr}[U(\varrho_X \otimes \sigma) U^* I \otimes \mathbf{Z}(X)]. \quad (10.6)$$

We show that the last term of (10.5) vanishes identically, so that the probability reproducibility condition is obtained. Let $\sigma = \sum w_i P[\phi_i]$ be a decomposition of σ . It follows from (10.6) that for any $P[\phi_i]$ one also has

$$\mathrm{tr}[U(\varrho_X \otimes P[\phi_i]) U^* I \otimes \mathbf{Z}(X)] = 1,$$

and therefore

$$U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) \otimes P[\phi_i](\varphi \otimes \psi) = \mathbf{E}(X) \otimes P[\phi_i](\varphi \otimes \psi)$$

for all $\varphi \in \mathcal{H}$, $\psi \in \mathcal{K}$. Thus,

$$U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) \otimes P[\phi_i] = \mathbf{E}(X) \otimes P[\phi_i]$$

and so

$$\mathbf{E}(X^c) \otimes P[\phi_i] U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) \otimes P[\phi_i] = 0.$$

Taking the expectation with respect to $\varphi \otimes \phi_i$ yields

$$\mathrm{tr}[U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) P[\varphi] \mathbf{E}(X^c) \otimes P[\phi_i]] = 0.$$

But this gives

$$\begin{aligned}\mathrm{tr}[I \otimes \mathbf{Z}(X) U(\varrho_{XX^c} \otimes \sigma) U^*] \\ &= \mathrm{tr}[U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) P[\varphi] \mathbf{E}(X^c) \otimes \sigma] \\ &= \sum w_i \mathrm{tr}[U^*(I \otimes \mathbf{Z}(X)) U \mathbf{E}(X) P[\varphi] \mathbf{E}(X^c) \otimes P[\phi_i]] = 0.\end{aligned}$$

Finally, if $\varrho = \sum r_i P[\varphi_i]$, then one has

$$\mathbf{E}_{\varrho}(X) = \sum r_i \mathbf{E}_{\varphi_i}(X) = \sum r_i \text{tr}[U(P[\varphi_i] \otimes \sigma)U^* I \otimes \mathbf{Z}(X)] = \text{tr}[\sigma^f \mathbf{Z}(X)],$$

which completes the proof. $\square$

10.2 Instruments

The effect of a measurement $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$ on the system is described by a (completely positive) instrument $\mathfrak{I}_{\mathcal{M}} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{H}))$, defined by $\mathcal{M}$ through the relation

$$\mathfrak{I}_{\mathcal{M}}(X)(\varrho) = \text{tr}_{\mathcal{K}}[U(\varrho \otimes \sigma)U^* I \otimes \mathbf{Z}(X)] \quad (10.7)$$

for all $X \in \mathcal{A}$, $\varrho \in \mathcal{S}(\mathcal{H})$. We will denote this relation between measurement schemes and instruments as $\mathcal{M} \hookrightarrow \mathfrak{I}_{\mathcal{M}}$ and call $\mathfrak{I}_{\mathcal{M}}$ the instrument induced by $\mathcal{M}$. The induced instrument $\mathfrak{I}_{\mathcal{M}}$ reproduces the measured observable $\mathbf{E}$ via the equation

$$\text{tr}[\varrho \mathbf{E}(X)] = \text{tr}[\mathfrak{I}_{\mathcal{M}}(X)(\varrho)], \quad X \in \mathcal{A}, \quad \varrho \in \mathcal{S}(\mathcal{H}), \quad (10.8)$$

or equivalently, using the Heisenberg representation of $\mathfrak{I}_{\mathcal{M}}$,

$$\mathbf{E}(X) = \mathfrak{I}_{\mathcal{M}}(X)^*(I), \quad X \in \mathcal{A}. \quad (10.9)$$

Occasionally, we simply write $\mathfrak{I}$ for the instrument $\mathfrak{I}_{\mathcal{M}}$.

The interpretation of $\mathfrak{I}(X)(\varrho)$ as the nonnormalised final state of S on the condition that the probe observable has a value in X after the measurement can be justified as follows. Assume that an observable $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ of the system is to be measured after the measurement of $\mathbf{E}$. The observables $\mathbf{E}_1$ and $\mathbf{Z}$ have a joint probability in each state of $\mathcal{S} + \mathcal{P}$, particularly in the final system-probe state $U(\varrho \otimes \sigma)U^*$. Let μ denote this joint probability, so that $\mu(Y \times X) = \text{tr}[U(\varrho \otimes \sigma)U^* \mathbf{E}_1(Y) \otimes \mathbf{Z}(X)]$ for any $Y \in \mathcal{A}_1$, $X \in \mathcal{A}$. The (classical) conditional probability for $Y \simeq Y \times \Omega$ given $X \simeq \Omega_1 \times X$ is

$$\mu(Y|X) = \frac{\mu(Y \times X)}{\mu(\Omega_1 \times X)} = \frac{\text{tr}[U(\varrho \otimes \sigma)U^* \mathbf{E}_1(Y) \otimes \mathbf{Z}(X)]}{\text{tr}[U(\varrho \otimes \sigma)U^* I \otimes \mathbf{Z}(X)]} = \frac{\text{tr}[\mathfrak{I}(X)(\varrho) \mathbf{E}_1(Y)]}{\text{tr}[\mathfrak{I}(X)(\varrho)]}.$$

Since $\mathbf{E}_1$ is an arbitrary observable, the normalised state

$$\varrho^f(X) = \mathfrak{I}(X)(\varrho) / \text{tr}[\mathfrak{I}(X)(\varrho)]$$

is uniquely determined by the conditioning event $I \otimes \mathbf{Z}(X)$ as the state that gives the probabilities for any subsequent measurement on S . In particular,

$$\varrho^f(\Omega) = \mathfrak{I}(\Omega)(\varrho) = \text{tr}_{\mathcal{K}}[U(\varrho \otimes \sigma)U^*] = \varrho^f$$

is the state of $\mathcal{S}$ after the measurement with the plain condition that the measurement has been performed. Also, for any $X \in \mathcal{A}$, if $0 \neq \mathbf{E}_{\varrho}(X) \neq 1$, then

$$\varrho^f(\Omega) = \mathbf{E}_{\varrho}(X)\varrho^f(X) + \mathbf{E}_{\varrho}(X^c)\varrho^f(X^c). \quad (10.10)$$

It must be emphasised that the state $\varrho^f(X)$ cannot, in general, be interpreted conditionally with respect to the values of the measured observable $\mathbf{E}$. However, as we will see in Sect. 10.4, this interpretation is in fact tenable for the important class of Lüders measurements of discrete sharp observables. In this case, the (nonnormalised) state $\mathfrak{I}_{\mathcal{M}}(X)(\varrho)$ takes the form $\mathbf{E}(X)\varrho\mathbf{E}(X)$, which is the $\mathbf{E}(X)$ -conditioned state in the sense described in Sect. 9.4.

Similarly, for any $X \in \mathcal{A}$ such that $\mathbf{E}_{\varrho}(X) \neq 0$ one may define the final probe state

$$\sigma^f(X) = \mathbf{E}_{\varrho}(X)^{-1} \text{tr}_{\mathcal{H}}[I \otimes \mathbf{Z}(X)^{1/2} U(\varrho \otimes \sigma) U^* I \otimes \mathbf{Z}(X)^{1/2}].$$

If the pointer observable is sharp, and thus commutative, this state is indeed the final state of the probe conditional on the pointer condition “ $\mathbf{Z}X$ ” in the sense of Sect. 9.4.

The interpretation of the instrument $\mathfrak{I}$ given here presupposes that the occurrence of definite values of the probe observable can be justified. Such a justification will depend on the interpretation of quantum mechanics that one adopts (see Chap. 22).

Remark 10.3 The notion of an instrument according to Definition 7.5 is more general than that obtained from a measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U)$, in that the latter entails complete positivity. Conversely, according to Theorem 7.14, it is exactly the class of completely positive instruments that are induced by some measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U)$. $\triangleleft$

The relation between measurement schemes and induced instruments is many to one. This is a reflection of the fact that different apparatuses may realise the ‘same’ measurement, or that for one apparatus there may be different ways of carrying out the pointer reading. Therefore it is convenient to identify the class of measurement schemes yielding the same state changes for the object: two measurement schemes $\mathcal{M}_1$ and $\mathcal{M}_2$ are called (*operationally*) *equivalent* if their induced instruments are the same, that is, $\mathcal{M}_1 \hookrightarrow \mathfrak{I}$ and $\mathcal{M}_2 \hookrightarrow \mathfrak{I}$. Note that equivalent measurement schemes define the same observable $\mathbf{E}$.

The theory of measurement operates on a hierarchy of three levels of description: measurement schemes $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$, completely positive instruments $\mathfrak{I}$, and observables $\mathbf{E}$, which are linked through the equivalence relations introduced above and their associated classes (denoted by $[\cdot]_{\mathfrak{I}}$):

$$\mathbf{E} \rightarrow [\mathfrak{I}]_{\mathbf{E}}, \quad \mathfrak{I} \rightarrow [\mathcal{M}]_{\mathfrak{I}}.$$

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. At each level the basic outcome probabilities are recovered: for any $\varrho \in \mathcal{S}(\mathcal{H})$ and for any $X \in \mathcal{A}$,

$$\mathrm{tr}[\varrho \mathbf{E}(X)] = \mathrm{tr}[\mathcal{I}(X)(\varrho)] = \mathrm{tr}[U(\varrho \otimes \sigma)U^* I \otimes \mathbf{Z}(X)]. \quad (10.11)$$

Remark 10.4 Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be an observable with a minimal Naimark dilation $(\mathcal{K}, \mathbf{F}, V)$ and $\mathcal{I} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{H}))$ a *completely positive* instrument. Then, by Theorem 7.12 (see also Remark 8.3),

$$\mathbf{E}(X) = \mathcal{I}(X)^*(I)$$

for all $X \in \mathcal{A}$ if and only if, there exists a (unique) completely positive unital linear map (i.e. a channel) $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ such that, for all $X \in \mathcal{A}$ and $B \in \mathcal{L}(\mathcal{H})$,

$$\mathbf{F}(X)\Phi(B) = \Phi(B)\mathbf{F}(X)$$

and

$$\mathcal{I}(X)^*(B) = V^*\mathbf{F}(X)\Phi(B)V.$$

Hence, the class $[\mathcal{I}]_{\mathbf{E}}$ can be identified with the convex set of channels Φ commuting with the minimal Naimark dilation $\mathbf{F}$ of $\mathbf{E}$. $\triangleleft$

10.3 Sequential, Joint and Mixed Measurements

In this section we consider three ways of “combining” several measurement schemes, which serve as model realisations of sequential, joint and mixed measurements. As measurement schemes are here assumed to be applicable to arbitrary states, it is possible for any two (or more) measurements to be performed sequentially or to be mixed. In contrast, the idea of joint measurement concerns the application of several measurement schemes simultaneously to the same system; this cannot be expected to be feasible without qualification.

Sequential Measurements

Consider any two measurement schemes $\mathcal{M}_i = (\mathcal{H}_i, \mathbf{Z}_i, \sigma_i, U_i)$ of the observables $\mathbf{E}_i \in \mathcal{O}(\Omega_i, \mathcal{A}_i, \mathcal{H})$, $i = 1, 2$, of a system $\mathcal{S}$. These schemes can be performed on $\mathcal{S}$ successively in either order: first $\mathcal{M}_1$ and then $\mathcal{M}_2$, or vice versa. To formulate this idea, we introduce some obvious notations. Let $\mathcal{A}_1 \times \mathcal{A}_2$ denote the Cartesian product of $\mathcal{A}_1$ and $\mathcal{A}_2$, and $\mathcal{A}_1 \otimes \mathcal{A}_2$ the σ -algebra of subsets of $\Omega_1 \times \Omega_2$ generated by the sets $X \times Y$, $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. Further, let I_{12} be the isometric isomorphism $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H} \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$ that switches the positions of $\mathcal{H}_1$ and $\mathcal{H}_2$ in the tensor product $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2$. Then $I_{12}^{-1} = I_{21}$ and we put

$$\begin{aligned}\widetilde{U}_1 &= U_1 \otimes I_2, & \widetilde{U}_2 &= I_{21} U_2 \otimes I_1 I_{12}, \\ \widehat{U}_2 &= U_2 \otimes I_1, & \widehat{U}_1 &= I_{12} U_1 \otimes I_2 I_{21}.\end{aligned}$$

One may now combine the measurements $\mathcal{M}_1$ and $\mathcal{M}_2$ into *sequential measurements* either as $\mathcal{M}_{12} = (\mathcal{H}_1 \otimes \mathcal{H}_2, \mathbf{Z}_1 \otimes \mathbf{Z}_2, \sigma_1 \otimes \sigma_2, \widetilde{U}_2 \widetilde{U}_1)$ or as $\mathcal{M}_{21} = (\mathcal{H}_2 \otimes \mathcal{H}_1, \mathbf{Z}_2 \otimes \mathbf{Z}_1, \sigma_2 \otimes \sigma_1, \widehat{U}_1 \widehat{U}_2)$. The observables $\mathbf{E}_{12} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{E}_{21} : \mathcal{A}_2 \otimes \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ measured by the schemes $\mathcal{M}_{12}$ and $\mathcal{M}_{21}$ are determined as follows:

$$\begin{aligned}\mathfrak{J}_{12}(X \times Y)(\varrho) &= \text{tr}_{12}[\widetilde{U}_2 \widetilde{U}_1(\varrho \otimes \sigma_1 \otimes \sigma_2) \widetilde{U}_1^* \widetilde{U}_2^* I \otimes \mathbf{Z}_1(X) \otimes \mathbf{Z}_2(Y)] \\ &= \mathfrak{J}_2(Y)(\mathfrak{J}_1(X)(\varrho)) = \mathfrak{J}_2(Y) \circ \mathfrak{J}_1(X)(\varrho),\end{aligned}\quad (10.12)$$

for all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$, and for any vector state $\varrho \in \mathcal{S}(\mathcal{H})$, and similarly

$$\mathfrak{J}_{21}(Y \times X)(\varrho) = \mathfrak{J}_1(X) \circ \mathfrak{J}_2(Y)(\varrho) \quad (10.13)$$

for all $Y \in \mathcal{A}_2$, $X \in \mathcal{A}_1$, and $\varrho \in \mathcal{S}(\mathcal{H})$. The observables $\mathbf{E}_{12}$ and $\mathbf{E}_{21}$ measured by the sequential measurements $\mathcal{M}_{12}$ and $\mathcal{M}_{21}$ are thus of the form

$$\mathbf{E}_{12}(X \times Y) = \mathfrak{J}_{12}^*(X \times Y)(I) = \mathfrak{J}_1^*(X)(\mathfrak{J}_2^*(Y)(I)) = \mathfrak{J}_1^*(X)(\mathbf{E}_2(Y)) \quad (10.14)$$

$$\mathbf{E}_{21}(Y \times X) = \mathfrak{J}_{21}^*(Y \times X)(I) = \mathfrak{J}_2^*(Y)(\mathfrak{J}_1^*(X)(I)) = \mathfrak{J}_2^*(Y)(\mathbf{E}_1(X)). \quad (10.15)$$

We call them *sequential observables*. Their margins are

$$\mathbf{E}_{12}(X \times \Omega_2) = \mathbf{E}_1(X), \quad \mathbf{E}_{12}(\Omega_1 \times Y) = \mathfrak{J}_1^*(\Omega_1)(\mathbf{E}_2(Y)) \quad (10.16)$$

$$\mathbf{E}_{21}(Y \times \Omega_1) = \mathbf{E}_2(Y), \quad \mathbf{E}_{21}(\Omega_2 \times X) = \mathfrak{J}_2^*(\Omega_2)(\mathbf{E}_1(X)). \quad (10.17)$$

The first of these is the observable measured first, $\mathbf{E}_1$, whereas the second is a kind of ‘distorted’ version, denoted $\widetilde{\mathbf{E}}_2$, of the observable measured secondly, $\mathbf{E}_2$, the distortion depending on the first measurement.

The instruments $\mathfrak{J}_{12}$ and $\mathfrak{J}_{21}$ of the sequential measurements $\mathcal{M}_{12}$ and $\mathcal{M}_{21}$ are compositions of the instruments $\mathfrak{J}_1$ and $\mathfrak{J}_2$ of the measurements $\mathcal{M}_1$ and $\mathcal{M}_2$, that is, $\mathfrak{J}_{12} = \mathfrak{J}_2 \circ \mathfrak{J}_1$ and $\mathfrak{J}_{21} = \mathfrak{J}_1 \circ \mathfrak{J}_2$ in the sense of Eqs. (10.12) and (10.13). Conversely, consider any two completely positive instruments $\mathfrak{J}_i : \mathcal{A}_i \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{H}))$, $i = 1, 2$. Then, for instance, for each state $\varrho \in \mathcal{S}(\mathcal{H})$ the function

$$\mathcal{A}_1 \times \mathcal{A}_2 \ni (X, Y) \mapsto \mu_\varrho(X, Y) = \text{tr}[\mathfrak{J}_2(Y)(\mathfrak{J}_1(X)(\varrho))] \in [0, 1]$$

is a *probability bimeasure*, that is, a separately σ -additive function with $\mu_\varrho(\Omega_1, \Omega_2) = 1$. By the duality $\mathcal{T}(\mathcal{H})^* \cong \mathcal{L}(\mathcal{H})$, the bimeasures μ_ϱ , $\varrho \in \mathcal{S}(\mathcal{H})$, define a *positive operator bimeasure*, a *biobservable*, $\mathbf{B}_{12} : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ such that

$$\mathrm{tr}[\varrho \mathbf{B}_{12}(X, Y)] = \mu_\varrho(X, Y),$$

for all $\varrho \in \mathcal{S}(\mathcal{H})$, $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. (These notions are discussed in detail in Chap. 11.) The marginal observables $X \mapsto \mathbf{B}_{12}(X, \Omega_2)$ and $Y \mapsto \mathbf{B}_{12}(\Omega_1, Y)$ are easily seen to be identical to the observables $X \mapsto \mathcal{I}_1(X)^*(I) = \mathbf{E}_1(X)$ and $Y \mapsto \mathcal{I}_1^*(\Omega_1)(\mathcal{I}_2^*(Y)(I)) = \mathcal{I}_1^*(\Omega_1)(\mathbf{E}_2(Y))$, with $\mathcal{I}_1 \hookrightarrow \mathbf{E}_1$ and $\mathcal{I}_2 \hookrightarrow \mathbf{E}_2$, as in Eq. (10.14). The fact that the biobservable $\mathbf{B}_{12}$, resp. $\mathbf{B}_{21}$, extends to the sequential observable $\mathbf{E}_{12}$, resp. $\mathbf{E}_{21}$, follows in this context from the measurement representation (Theorem 7.14) of a completely positive instrument together with the tensor product structure, cf. Sect. 9.5. Therefore, no topological assumptions are needed here.

The sequential measurements $\mathcal{M}_{12}$ and $\mathcal{M}_{21}$ define sequential instruments $\mathcal{I}_{12}$ and $\mathcal{I}_{21}$, which, in turn, define sequential observables $\mathbf{E}_{12}$ and $\mathbf{E}_{21}$. As is evident from the construction they are generally order dependent, as manifested by the relevant—generally different—marginal observables (10.16) and (10.17). In Chap. 11 we investigate the case of order-independent sequential observables.

Joint Measurements

We now turn to the question of the simultaneous measurability of two observables $\mathbf{E}_1$ and $\mathbf{E}_2$ of a system $\mathcal{S}$. We say that $\mathbf{E}_1$ and $\mathbf{E}_2$ can be *measured jointly* if there is a measurement scheme $\mathcal{M}$ with pointer functions f_1 and f_2 such that

$$\begin{aligned} \mathbf{E}_1(X) &= E^{\mathcal{M}}(f_1^{-1}(X)), \quad X \in \mathcal{A}_1, \\ \mathbf{E}_2(Y) &= E^{\mathcal{M}}(f_2^{-1}(Y)), \quad Y \in \mathcal{A}_2. \end{aligned}$$

In conjunction with Theorem 10.1 we find that any two observables can be measured jointly exactly when they are functions of a third observable. In Chap. 11 we show (Theorem 11.1) that, subject to some topological requirements, the joint measurability of $\mathbf{E}_1$ and $\mathbf{E}_2$ is equivalent to the existence of a joint observable $\mathbf{E} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ with the property $\mathbf{E}(X \times Y) = E^{\mathcal{M}}(f_1^{-1}(X) \cap f_2^{-1}(Y))$.

The sequential observable $\mathbf{E}_{12}$ of Eq. (10.14) is a joint observable of its margins, which are $\mathbf{E}_1$ and the distorted version of $\mathbf{E}_2$, $\tilde{\mathbf{E}}_2 = \mathcal{I}_1^*(\Omega_1) \circ \mathbf{E}_2$; in this case, $\mathbf{E}_{12}(X \times Y) = \mathcal{I}_1^*(X)(\mathbf{E}_2(Y))$. Since generally $\tilde{\mathbf{E}}_2 \neq \mathbf{E}_2$, it is evident that the naive idea of measuring $\mathbf{E}_1$ and $\mathbf{E}_2$ jointly by measuring them in immediate succession is bound to fail.

A more sophisticated attempt would be to find a suitable observable $\mathbf{F}_2$ that, when measured immediately after $\mathbf{E}_1$, is distorted into $\tilde{\mathbf{F}}_2 = \mathbf{E}_2$. Again, it turns out that this does not work if $\mathbf{E}_1$, $\mathbf{E}_2$ are not jointly measurable. The best one can hope to achieve is to find an observable $\mathbf{F}_2$ for which the difference between $\mathbf{E}_2$ and $\tilde{\mathbf{F}}_2$ becomes

minimal in some sense. This procedure constitutes an instance of an approximate joint measurement, a subject that is introduced and studied systematically in Chap. 13.

Remark 10.5 The classical idea of measurements not disturbing each other is not entirely lost in quantum mechanics. In Sect. 10.4 we present the class of Lüders measurements for discrete sharp observables along with measurement realisations in terms of the so-called standard model. These schemes are mutually nondisturbing when applied sequentially or simultaneously if the measured observables commute (see also Exercise 8).

Mixing Measurements

Finally we consider the possibility of *mixing* measurement schemes, which provide realisations of the procedure of mixing observables as described in Sect. 9.3.

Let $(\mathcal{K}, \mathbf{Z}, \sigma, U)$ be a measurement scheme whose probe state σ is a mixed state, with a decomposition $\sigma = \sum_i r_i P[\phi_i]$. Then the measured observable $\mathbf{E}$ is a mixture of the observables $\mathbf{E}_i$, defined by the measurement schemes $(\mathcal{K}, U, P[\phi_i], \mathbf{Z})$, that is, $\mathbf{E} = \sum r_i \mathbf{E}_i$. If $\mathbf{E}$ is an extremal observable (for instance, a sharp observable), then all the measurement schemes $(\mathcal{K}, U, P[\phi_i], \mathbf{Z})$ defined by the different decompositions of σ are equivalent.

In this decomposition of an observable into a mixture, all component observables are defined on the same outcome space. This procedure can be considered as a model for measurement imperfections arising from the fact that the initial probe state may be subject to fluctuations or otherwise only incompletely known, resulting in additional uncertainty over which observable is actually being measured in each run of the experiment.

There is also the possibility of mixing measurements with not necessarily identical outcome spaces. Consider measurement schemes $\mathcal{M}_1, \mathcal{M}_2$ for observables $\mathbf{E}_1, \mathbf{E}_2$. Let $\mathcal{H}_3$ be an auxiliary Hilbert space, and with two nonzero projections $C_1, C_2 = I - C_1$ in $\mathcal{H}_3$, we define a *choice observable* $i \mapsto C_i, i = 1, 2$. Fix a $\sigma_3 \in \mathcal{S}(\mathcal{H}_3)$ such that the probabilities $p_i = \text{tr}[\sigma_3 C_i], i = 1, 2$, are nonzero. Denoting by U_1 and U_2 the measurement couplings and their natural extensions to the Hilbert space $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2$, we define the unitary operator $U = U_1 \otimes C_1 + U_2 \otimes C_2$ on $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. Put $\Omega = \Omega_1 \times \Omega_2 \times \{1, 2\}$ and let $\mathcal{A}$ be the σ -algebra of subsets of Ω of the form $X_1 \times \Omega_2 \times \{1\} \cup \Omega_1 \times X_2 \times \{2\}, X_i \in \mathcal{A}_i$. Finally, define the spectral measure $\mathbf{P}$ on $\mathcal{A}$ by

$$\mathbf{P}(X_1 \times \Omega_2 \times \{1\} \cup \Omega_1 \times X_2 \times \{2\}) = \mathbf{Z}_1(X_1) \otimes I_2 \otimes C_1 + I_1 \otimes \mathbf{Z}_2(X_2) \otimes C_2.$$

The observable $\mathbf{E}^{\mathcal{M}}$ induced by the measurement scheme

$$\mathcal{M} = (\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3, \mathbf{P}, \sigma_1 \otimes \sigma_2 \otimes \sigma_3, U)$$

is now

$$\mathbf{E}^{\mathcal{M}}(X_1 \times \Omega_2 \times \{1\} \cup \Omega_1 \times X_2 \times \{2\}) = p_1 \mathbf{E}_1(X_1) + p_2 \mathbf{E}_2(X_2).$$

Since one can keep track (via the values of the choice observable) of which measurement outcome belongs to either E_1 or E_2 , this procedure yields a determination of the statistics of both of these observables, whether they are jointly measurable or not. Clearly, this is not a joint measurement as each copy of the object system from the measured ensemble is exposed only to one of the measurements, either $\mathcal{M}_1$ or $\mathcal{M}_2$.

This method of randomising measurements is important for experimental tests of Bell inequalities, where it can serve to exclude any chance of the two separated measurement setups communicating with each other.

10.4 Examples of Measurement Schemes

Throughout this section, we let $(\varphi_n)_{n=1}^{\dim \mathcal{H}}$ be an orthonormal basis of $\mathcal{H}$ and V the linear span of the vectors φ_n , that is, the subspace containing all the finite linear combinations of the vectors φ_n .

Minimal Measurements of Discrete Observables

A discrete observable E can be viewed as a collection of positive operators $E_i \in \mathcal{L}(\mathcal{H}) \setminus \{0\}$, $1 \leq i \leq N$, such that $\sum_{i=1}^N E_i = I_{\mathcal{H}}$. Now a minimal Kolmogorov decomposition of E is of the form

$$E_i = \sum_{n,m=1}^{\dim \mathcal{H}} \langle \eta_n(i) | \eta_m(i) \rangle |\varphi_n\rangle \langle \varphi_m|$$

where the vectors $\eta_n(i) \in \mathcal{H}_{m(i)}$, $1 \leq n \leq \dim \mathcal{H}$, $\text{span } \mathcal{H}_{m(i)} \subset \mathcal{H}$ (see Sect. 8.6). Note that the corresponding direct integral Hilbert space $\mathcal{H}_{\oplus}$ reduces to a direct sum $\bigoplus_{i=1}^N \mathcal{H}_{m(i)}$.

By Remark 8.3, any minimal measurement dilation of E is given by isometries $C(i) : \mathcal{H}_{m(i)} \rightarrow \mathcal{H}$ and channels $\Phi(i) : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}_{m(i)})$,

$$[\Phi(i)](B) = C(i)^* B C(i), \quad B \in \mathcal{L}(\mathcal{H}),$$

which defines a Heisenberg instrument $\mathfrak{I}_C^*$,

$$\mathfrak{I}_C^*(i)(B) = \sum_{n,m=1}^{\dim \mathcal{H}} \langle \eta_n(i) | C(i)^* B C(i) \eta_m(i) \rangle |\varphi_n\rangle \langle \varphi_m|$$

whose Schrödinger form is

$$\mathfrak{I}_C(i)(\varrho) = C(i) \left(\sum_{n,m=1}^{\dim \mathcal{H}} \langle \varphi_m | \varrho \varphi_n \rangle |\eta_m(i)\rangle \langle \eta_n(i)| \right) C(i)^*, \quad \varrho \in \mathcal{T}(\mathcal{H}).$$

The (minimal) probe Hilbert space $\mathcal{K}$ is N -dimensional and a (discrete) pointer observable $\mathbf{M}$ consists of the projections $\mathbf{M}_i = P[\phi_i] = |\phi_i\rangle\langle\phi_i|$, $1 \leq i \leq N$, where $(\phi_i)_{i=1}^N$ is some orthonormal basis of $\mathcal{K}$. By fixing a unit vector $\phi \in \mathcal{K}$, the initial vector state of the probe, one can define a unitary measurement interaction $U_\phi \in \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ by extending¹

$$U_\phi(\varphi_m \otimes \phi) = \sum_{i=1}^N C(i)\eta_m(i) \otimes \phi_i.$$

Since $\mathcal{I}_C^*(i)(B) = V_\phi^* U_\phi^* [B \otimes \mathbf{M}_i] U_\phi V_\phi$, with $V_\phi(\varphi) = \varphi \otimes \phi$, then $(\mathcal{K}, \phi, \mathbf{M}, U_\phi)$ is a measurement of $\mathbf{E}$. We call it a *minimal measurement* of $\mathbf{E}$ and denote it by $\mathcal{M}^m$.

Note that, by defining new structure vectors $\tilde{\eta}_m(i) = C(i)\eta_m(i)$ one gets

$$E_i = \sum_{n,m=1}^{\dim \mathcal{H}} \langle \tilde{\eta}_n(i) | \tilde{\eta}_m(i) \rangle |\varphi_n\rangle\langle\varphi_m|,$$

that is, a new minimal Kolmogorov decomposition of $\mathbf{E}$. Hence, it is not very restrictive to assume that $C(i) = \sum_{n=1}^{m(i)} |\varphi_n\rangle\langle\varphi_n|$ so that the operators $C(i)$ disappear from the above formulas.²

If $\mathbf{E}$ is of rank 1 then the isometries $C(i) : \mathcal{H}_1 \rightarrow \mathcal{H}$ can be identified with unit vectors $\psi_i \in \mathcal{H}$ and one gets

$$\mathcal{I}_C^*(i)(B) = \sum_{n,m=1}^{\dim \mathcal{H}} \langle \psi_i | B \psi_i \rangle \langle \eta_n(i) | \eta_m(i) \rangle |\varphi_n\rangle\langle\varphi_m| = \langle \psi_i | B \psi_i \rangle E_i$$

or

$$\mathcal{I}_C(i)(\varrho) = |\psi_i\rangle\langle\psi_i| \text{tr}[\varrho E_i],$$

that is, $\mathcal{I}_C$ produces a vector state ψ_i if the i th outcome ω_i is obtained (regardless of the input state ϱ).

Example 10.1 If $\mathbf{E}$ is projection valued then $\mathcal{H} \simeq \bigoplus_{i=1}^N \mathcal{H}_{m(i)}$. Let then the vectors φ_{ij} , $i \leq N$, $j \leq m(i)$, constitute an orthonormal basis of $\mathcal{H}$ such that

$$E_i = \sum_{j=1}^{m(i)} |\varphi_{ij}\rangle\langle\varphi_{ij}|.$$

¹Recall that, if for some i , $m_i < \infty$ (i.e. $\mathbf{E}$ is not rank- ∞) then isometries $C(i)$ can be arbitrary [1]. In the rank- ∞ case, $m(i) = \infty$ for all i (implying $\dim \mathcal{H} = \infty$), there exist (infinitely many) isometries $C(i)$ such that U_ϕ extends to a unitary operator. For example, one may choose $C(i) = \sum_{n=1}^{\infty} |\varphi_{2n}\rangle\langle\varphi_n|$ for all i .

²In the rank- ∞ case, this choice for the isometries $C(i)$ does not give a unitary U_ϕ but then one can add one extra dimension to $\mathcal{K}$, i.e. define the ancillary space to be $\mathcal{K} \oplus \mathbb{C}$.

Now the vectors φ_n are replaced by the vectors φ_{ij} , that is, the index n is replaced by ij , so that $\eta_n(i')$ becomes $\eta_{ij}(i') = \varphi_{ij}\delta_{ii'}$. Hence,

$$\mathfrak{I}_C^*(i)(B) = \sum_{j,k=1}^{m(i)} \langle \varphi_{ij} | C(i)^* B C(i) \varphi_{ik} \rangle | \varphi_{ij} \rangle \langle \varphi_{ik} | = E_i C(i)^* B C(i) E_i,$$

with $C(i) : \mathcal{H}_{m(i)} \rightarrow \mathcal{H}$ an isometry. Note that $\mathcal{H}_{m(i)} = E_i(\mathcal{H})$ and

$$U_\phi(\varphi_{ij} \otimes \phi) = \sum_{i'=1}^N C(i') \eta_{ij}(i') \otimes \phi_{i'} = \psi_{ij} \otimes \phi_i$$

where the vectors $\psi_{ij} = C(i)\varphi_{ij}$ are such that $\langle \psi_{ij} | \psi_{ik} \rangle = \langle \varphi_{ij} | \varphi_{ik} \rangle = \delta_{jk}$ holds (and, in the rank- ∞ case, are chosen in such a way that U_ϕ extends to a unitary operator). $\triangleleft$

Example 10.2 To continue the above example, any minimal measurement of a discrete sharp observable $A = \sum_{i=1}^N a_i P_i$, with N (possibly infinite) distinct eigenvalues, is now of the form $\mathcal{M}^m = (\mathcal{K}, Z, \phi, U_\phi)$, where $\mathcal{K}$ is a Hilbert space with an orthonormal basis $(\phi_i)_{i=1}^N$, $Z = \sum_i a_i P[\phi_i]$, $\phi \in \mathcal{K}$ a fixed unit vector, and U_ϕ a unitary extension of the mapping

$$\varphi_{ij} \otimes \phi \mapsto \psi_{ij} \otimes \phi_i, \quad (10.18)$$

where the vectors φ_{ij} , $i \leq N$, $j \leq m(i)$, with $A\varphi_{ij} = a_i\varphi_{ij}$, form an orthonormal basis of $\mathcal{H}$, and $\{\psi_{ij}\}$ is any set of unit vectors in $\mathcal{H}$ satisfying the orthogonality conditions $\langle \psi_{ij} | \psi_{ik} \rangle = \delta_{jk}$ for all $j, k \leq m(i)$, for each $i \leq N$, and $\dim \{\varphi_{ij} \otimes \phi\}^\perp = \dim \{\psi_{ij} \otimes \phi_i\}^\perp$. This formulation is due to [2].

Let $\varphi = \sum c_{ij}\varphi_{ij}$ be the initial vector state of $\mathcal{S}$, so that

$$U_\phi(\varphi \otimes \phi) = \sum_{i,j} c_{ij} \psi_{ij} \otimes \phi_i \quad (10.19)$$

is the final state of $\mathcal{S} + \mathcal{P}$. The instrument $\mathfrak{I}_{U_\phi}$ takes the form

$$\mathfrak{I}_{U_\phi}(X)(\varrho) = \sum_{i: a_i \in X} K_i \varrho K_i^*, \quad (10.20)$$

with $K_i = \sum_l |\psi_{il}\rangle \langle \varphi_{il}| \in \mathcal{L}(\mathcal{H})$. If U and U' are two different unitary extensions of (10.18), then the resulting A -measurements are equivalent, that is, the instruments $\mathfrak{I}_U$ and $\mathfrak{I}_{U'}$ are the same.

The properties of the measurement $\mathcal{M}^m$ depend on the generating set $\{\psi_{ij}\}$. For instance, if this set is orthonormal, then the decomposition (10.19) is biorthogonal (for each $\varphi \in \mathcal{H}$) and the natural decompositions of the reduced states of $\mathcal{S}$ and $\mathcal{P}$

are the spectral decompositions. The most ideal case is obtained if the set $\{\psi_{ij}\}$ is chosen identical to $\{\varphi_{ij}\}$. In that case the unitary operator³ of $\mathcal{M}^m$ has the restriction

$$U_L(\varphi_{ij} \otimes \phi) = \varphi_{ij} \otimes \phi_i, \quad (10.21)$$

and the instrument $\mathfrak{I}_L$ reduces to

$$\mathfrak{I}_L(X)(\varrho) = \sum_{i:a_i \in X} P_i \varrho P_i. \quad (10.22)$$

We call this particular measurement of A its *Lüders measurement* and the induced instrument the *Lüders instrument*. $\triangleleft$

Example 10.3 There are other classes of unitary measurements of (a degenerate) A that can be obtained from $\mathcal{M}^m$ by means of the following method. Let B be any refinement of A in the sense that $A = f(B)$ for some function f , i.e. $\mathbf{E}^A(X) = \mathbf{E}^B(f^{-1}(X))$ for all $X \in \mathcal{B}(\mathbb{R})$. If $\mathcal{M}^m$ is a measurement of B , then $(\mathcal{K}, \mathbf{Z}, \phi, U, f)$ is a measurement of A (cf. Remark 10.1). In particular, if the measurement of B is a Lüders measurement, then the resulting A -measurement will be called a *von Neumann measurement*. Accordingly the induced instrument will be referred to as a *von Neumann instrument*. In this view Lüders measurements are but a special class of von Neumann measurements. They turn out to be the ideal ones (Sect. 10.6). Especially, if B is a maximal rank-1 refinement of A , then the resulting von Neumann instrument of A is

$$\begin{aligned} \mathfrak{I}_{\text{vN}}^A(X)(\varrho) &= \sum_{i:a_i \in X} \sum_j P[\varphi_{ij}] \varrho P[\varphi_{ij}] \\ &= \mathfrak{I}_L^B(f^{-1}(X))(\varrho) \\ &= \sum_{a_i \in X} \sum_{f(b_{ij})=a_i} \mathbf{E}^B(\{b_{ij}\}) \varrho \mathbf{E}^B(\{b_{ij}\}). \end{aligned} \quad (10.23)$$

For the sake of clarity, we have indicated here the observable associated with the instrument in question. It should be noted that in general the final state of the object system $\mathcal{S}$ after a von Neumann measurement of A is not the same as the state reached after a Lüders measurement of A . $\triangleleft$

Remark 10.6 The unitary mapping (10.21) in the measurement $(\mathcal{K}, \mathbf{Z}, \phi, U_L)$ of A was already discussed by von Neumann [3], whereas the instrument (10.22) induced by that measurement was studied in greater detail by Lüders [4]. In the literature this measurement is called a Lüders or von Neumann–Lüders measurement. $\triangleleft$

Minimal measurements of a collection of discrete sharp observables can be combined either sequentially or by mixing in a straightforward way; we leave it as an exercise for a reader to work out the details.

³Also, in the rank- ∞ case, (10.21) extends to a unitary operator U_L .

Complete Measurements of Observables

Here we generalise the structure exhibited in Example 10.3. Let $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be an observable and $E(X) = \int_X \sum_{k=1}^{m(\omega)} |v_k(\omega)\rangle \langle v_k(\omega)| d\mu(\omega)$ the minimal diagonalisation of E (see Sect. 8.7). Recall that both sides of the above equation are interpreted as sesquilinear forms $V \times V \rightarrow \mathbb{C}$. If one attaches multiplicities $k \leq m(\omega)$ to the outcome space Ω of E one can refine E as follows. Define $\Omega_1 = \{(\omega, k) \in \Omega \times \mathbb{N}_0 \mid k \leq m(\omega)\}$ and let $\mathcal{A}_1$ be the restriction of the product σ -algebra of $\mathcal{A}$ and $2^{\mathbb{N}_0}$ to Ω_1 . The (maximal) rank-1 refinement of E is an observable $E_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ defined by

$$E_1(X \times \{k\}) = \int_X |v_k(\omega)\rangle \langle v_k(\omega)| d\mu(\omega),$$

where $X \in \mathcal{A}$ and $k \leq m(\omega)$. It is easy to see that E is a function of E_1 . Indeed, $E(X) = E_1(f^{-1}(X))$, $X \in \mathcal{A}$, where $f : \Omega_1 \rightarrow \Omega$ is a (measurable) function, $f(\omega, k) = \omega$. Thus, any measurement of E_1 can be viewed as a refined measurement of E where one gets information on the multiplicities of E . From Remark 8.3 we see that any E_1 -compatible (completely positive) instrument $\mathcal{I} : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{T}(\mathcal{H}))$ is of the form $\mathcal{I}^*(X_1)(B) = \int_{X_1} \text{tr}[\sigma(\omega_1)B] dE_1(\omega_1)$, $X_1 \in \mathcal{A}_1$, $B \in \mathcal{L}(\mathcal{H})$, or equivalently,

$$\mathcal{I}(X_1)(\varrho) = \int_{X_1} \sigma(\omega_1) \text{tr}[\varrho E_1(d\omega_1)], \quad X_1 \in \mathcal{A}_1, \quad \varrho \in \mathcal{S}(\mathcal{H}),$$

where the map $\Omega_1 \ni \omega_1 \mapsto \sigma(\omega_1) \in \mathcal{S}(\mathcal{H})$ is weakly measurable.⁴ For any $X_1 \in \mathcal{A}_1$, the conditional final state $\varrho^f(X_1) = \mathcal{I}(X_1)(\varrho) / \text{tr}[\mathcal{I}(X_1)(\varrho)]$ is a generalised mixture of the states $\sigma(\omega_1)$ which does not depend on the initial state ϱ . Hence, the operations of $\mathcal{I}$ act as state preparations.

Especially, if the states $\sigma(\omega_1)$ are pure, i.e. (by denoting $\omega_1 = (\omega, k)$), one has unit vectors $\varphi(\omega, k) \in \mathcal{H}$ such that $\sigma(\omega_1) = |\varphi(\omega, k)\rangle \langle \varphi(\omega, k)|$, the unitary interaction $U \in \mathcal{L}(\mathcal{H} \otimes L^2(\mu_1))$ of the (minimal) measurement model satisfies

$$[U(\psi \otimes \xi)](\omega, k) = \langle d_k(\omega) | \psi \rangle \varphi(\omega, k) \otimes 1$$

for all $\psi \in V$ and μ_1 -almost all $(\omega, k) \in \Omega_1$, where μ_1 is a σ -finite measure mutually absolutely continuous with E_1 . As before, the pointer observable is the canonical spectral measure M of $L^2(\mu_1)$.

Any measurement of E_1 is a *complete measurement* of E in the sense that (a) it gives information on the multiplicities of the outcomes, (b) it is preparative, i.e. the output states $\sigma(\omega_1)$ do not depend on the input state ϱ , and (c) the measurement breaks a possible entanglement between the system and its environment [5]. Moreover, E_1 can be measured by performing a sequential measurement of E and some discrete

⁴If E is absolutely continuous with respect to the positive measure μ then E_1 is absolutely continuous with respect to (the restriction of) the product of μ and the counting measure of $\mathbb{N}_0$ (to $\mathcal{A}_1$). The maps $\omega_1 \mapsto \langle \varphi | \sigma(\omega_1) \psi \rangle$, $\varphi, \psi \in \mathcal{H}$, are measurable with respect to this (restricted) product measure.

‘multiplicity’ observable [6]. If one measures any observable after a measurement of E_1 then the joint observable associated with the sequential measurement is a smearing of E_1 [7].

The Standard Model

Suppose that one intends to measure a sharp observable A of the system S by coupling it to a sharp observable B of a probe $\mathcal{P}$ through an interaction

$$U = e^{i\lambda A \otimes B}, \quad (10.24)$$

where A and B are the selfadjoint operators corresponding to A and B and $\lambda (\neq 0)$ is a coupling constant. Recalling that the spectral measure of the selfadjoint operator $A \otimes B$ is obtained from the joint measure $X \times Y \mapsto A(X) \otimes B(Y)$ by the function $(x, y) \mapsto xy$, one may write U in the form

$$U = \int_{\mathbb{R}} A(da) \otimes e^{ia\lambda B}. \quad (10.25)$$

Then an initial vector state $\varphi \otimes \phi$ of the compound system is transformed into

$$U(\varphi \otimes \phi) = \int_{\mathbb{R}} A(da) \varphi \otimes e^{ia\lambda B} \phi = \int_{\mathbb{R}} A(da) \varphi \otimes \phi_{\lambda a}, \quad (10.26)$$

where we have defined $\phi_{\lambda a} = e^{ia\lambda B} \phi$. If the coupling U is to serve its purpose one needs to choose the initial probe state ϕ , the pointer observable Z and a pointer function g such that $\mathcal{M} = (\mathcal{K}, Z, \phi, U, g)$ constitutes a measurement of A . With any choice of ϕ , Z , and g , the probability reproducibility condition specifies the measured observable $E = E^{\mathcal{M}}$. To evaluate this condition, one first determines the final state of the apparatus for the initial state ϱ of S ,

$$\sigma^f = \int_{\mathbb{R}} A_{\varrho}(da) P[\phi_{\lambda a}]. \quad (10.27)$$

The presence of the coupling constant λ suggests a rescaling by means of the pointer function $g(x) = \lambda^{-1}x$. The measured observable then takes the form

$$E(X) = \int_{\mathbb{R}} Z_{\phi_{\lambda a}}(\lambda X) A(da). \quad (10.28)$$

The structure of the effects $E(X)$ shows that in general the actually measured observable E is not the observable A but a smeared version of it, that is, $E = p \star A$, with the Markov kernel $p(X, a) = Z_{\phi_{\lambda a}}(\lambda X)$, see Eq. (8.4). This raises the question whether $E = A$ can be achieved by a suitable choice of ϕ and Z . This is the case exactly when for (A -almost) all $a \in \mathbb{R}$, $Z_{\phi_{\lambda a}}(\lambda X) = \chi_X(a)$, with χ_X denoting the characteristic function of the set X . Since

$$\langle \phi_{\lambda a} | Z(\lambda X) \phi_{\lambda a} \rangle = \langle \phi | e^{-ia\lambda B} Z(\lambda X) e^{ia\lambda B} \phi \rangle,$$

one may expect that an ‘optimal reading’ can be obtained if it is possible to choose the pointer Z to be *conjugate* to B in the sense of covariance:

$$e^{-ia\lambda B} Z(\lambda X) e^{ia\lambda B} = Z(\lambda(X - a)). \quad (10.29)$$

With such a choice the measured observable $E = p \star A$ takes the form of a convolution $Z_\phi^g \star A$ of the spectral measure A with the probability measure Z_ϕ^g , Eq. (8.7), that is, for all X ,

$$E(X) = \int Z_\phi(\lambda(X - a)) A(da). \quad (10.30)$$

Example 10.4 As a simple example, consider a discrete sharp observable A , with $A = \sum_{k=1}^N a_k P_k$, and assume that the set of eigenvalues of A is closed. For the probe we take a particle moving in one-dimensional space, so that $\mathcal{K} = L^2(\mathbb{R})$, and couple A with the probe momentum P_p via $U = \exp(-i\lambda A \otimes P_p)$, with P_p being the momentum operator. Since the momentum generates translations on the position, it is natural to take the probe position Q_p as the pointer observable. From (10.30) one now gets

$$E(X) = (Z_\phi^g \star A)(X) = \sum_k Z_\phi^g(X - a_k) P_k.$$

Assuming that the spacing between the eigenvalues a_k is greater than $\frac{\delta}{\lambda}$ and that ϕ is supported in $(-\frac{\delta}{2}, \frac{\delta}{2})$, then the pointer states $\phi_{\lambda a_k}$ are supported in the mutually disjoint sets λI_k , where $I_k = (a_k - \frac{\delta}{2\lambda}, a_k + \frac{\delta}{2\lambda})$. We introduce a discretising pointer function $\tilde{g}$ with $\tilde{g}(I_k) = \{a_k\}$ for each k and $\tilde{g}((\bigcup_k I_k)^c) \subset \{a_k : k = 1, 2, \dots\}^c$; this gives

$$E(\{a_k\}) = \sum_i \langle \phi_{\lambda a_i} | Q_p(\lambda I_k) \phi_{\lambda a_i} \rangle P_i = P_k \quad (10.31)$$

for each k , so that the observable measured by this scheme is indeed A . The instrument induced by this measurement is the Lüders instrument. $\triangleleft$

Example 10.5 In the case of a continuous observable, such as position, the model of Example 10.4 realises a measurement not of the position but of a smeared version of it. In fact for $A = Q$ one obtains

$$E = \mu \star Q, \quad \mu = (Q_p)_{\phi(\lambda)}. \quad (10.32)$$

Since μ cannot be a Dirac measure, the measured observable $\mu \star Q$ is never the sharp position but an unsharp one. The ensuing instrument $\mathcal{I}^Q$ is

$$\mathcal{I}^Q(X)(\varrho) = \int_X K_q \varrho K_q^* dq, \quad K_q = \sqrt{\lambda} \phi(-\lambda(Q - q)). \quad (10.33)$$

Recall that K_q is the multiplicative operator acting as

$$(K_q \psi)(x) = \sqrt{\lambda} \phi(-\lambda(x - q)) \psi(x)$$

for all $x \in \mathbb{R}$, $\psi \in L^2(\mathbb{R})$, with the requirement that ϕ is such that K_q is bounded.

This scheme for a measurement of an unsharp position observable was already discussed informally in the book of von Neumann [3]. It is therefore called the von Neumann or *standard model of an approximate position measurement*. Note that the standard model is a minimal measurement realisation of $\mathbf{Q}^\mu$ by Remark 8.3. We shall use the term *standard model* to refer to any (approximate or sharp) measurement scheme for $\mathbf{A}$ whose coupling is of the form $e^{i\lambda A \otimes B}$. $\triangleleft$

Standard measurement models may be combined into sequential measurements. They also give rise to natural joint measurement schemes. In Chap. 15 these models will be used to study the possibility of measuring together position and momentum observables.

Generalised Standard Model

The standard measurement model can easily be extended to arbitrary observables. We follow a formulation given in [8].

Let $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be an observable one aims to measure, and consider its dilation $(\mathcal{H} \otimes \mathcal{H}_0, \mathbf{E}^A, V_\psi)$ into a spectral measure $\mathbf{E}^A : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H} \otimes \mathcal{H}_0)$

$$\mathbf{E}(X) = V_\psi^* \mathbf{E}^A(X) V_\psi$$

with $V_\psi : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}_0$ being the embedding $V_\psi(\varphi) = \varphi \otimes \psi$ given by a unit vector $\psi \in \mathcal{H}_0$.

Let $\mathcal{K} = L^2(\mathbb{R})$ be the Hilbert space of the probe. For each $\lambda > 0$ define the state transformation $\Phi^\lambda : \mathcal{T}(\mathcal{H} \otimes \mathcal{K}) \rightarrow \mathcal{T}(\mathcal{H} \otimes \mathcal{K})$ via

$$\Phi^\lambda(\varrho \otimes \sigma) = \text{tr}_{\mathcal{H}_0} \left[e^{-i\lambda A \otimes P} \varrho \otimes |\psi\rangle\langle\psi| \otimes \sigma e^{i\lambda A \otimes P} \right]$$

where P is the momentum operator of the probe. Monitoring the probe position $\mathbf{Q}$ and using the scaling $f^\lambda(x) = \lambda^{-1}x$, the 5-tuple $\mathcal{M}^\lambda = \langle \mathcal{K}, \sigma, \Phi^\lambda, \mathbf{Q}, f^\lambda \rangle$ constitutes a *generalised standard measurement model* for $\mathbf{E}$.

The instrument as well as the observable actually measured by the scheme $\mathcal{M}^\lambda$ are again easily computed. Since we consider their λ -dependence we explicitly parametrise them by λ . For notational simplicity, assume that the initial probe state σ is a vector state ϕ . The associated instrument and its dual are then

$$\begin{aligned} \mathcal{I}^\lambda(X)(\varrho) &= \int_X \text{tr}_{\mathcal{H}_0} [K_x V_\psi \varrho V_\psi^* K_x^*] dx, \quad \varrho \in \mathcal{T}(\mathcal{H}), \\ \mathcal{I}^\lambda(X)^*(B) &= \int_X V_\psi^* K_x^* (B \otimes I) K_x V_\psi dx, \quad B \in \mathcal{L}(\mathcal{H}), \end{aligned}$$

where

$$K_x = \sqrt{\lambda} \phi(-\lambda(A - x)) = \int \sqrt{\lambda} \phi(-\lambda(y - x)) d\mathbf{E}^A(y),$$

for all $x \in \mathbb{R}$. The observable measured by $\mathcal{M}^\lambda$ is a convolution of $\mathbf{E}$ by the probability $\mu^\lambda(X) = \langle \phi | \mathbf{Q}(\lambda X) \phi \rangle$,

$$(\mu^\lambda * \mathbf{E})(X) = \int \mu^\lambda(X - x) d\mathbf{E}(x), \quad X \in \mathcal{B}(\mathbb{R}).$$

Note that although the generalised standard measurement model as well as the associated instrument depend on the ancilla used, through the Naimark dilation of $\mathbf{E}$, the actually measured observable $\mu^\lambda * \mathbf{E}$ is independent of it.

Weak Measurements and Weak Values

The notion of weak value of a quantum observable, conceived in 1988 by Aharonov et al., has attracted a surge of interest in recent years, as witnessed in the exponential increase of citations of the source paper [9]. Despite its accepted experimental utility [10], this concept has remained somewhat vexing and even mysterious, to the point that its inventors, among others, keep revisiting its foundation (see, e.g., [11] and references therein). In this subsection we show that weak values can easily be understood in the conventional terms of the quantum theory of sequential measurements, with the first measurement being a (generalised) standard measurement of the form discussed in the preceding subsection.

The intuitive idea behind the weak value of an observable is that by letting the strength of the (standard) interaction between the object and probe become sufficiently weak, the disturbance caused by the first measurement on the system becomes negligible. The price to be paid is that the first measurement becomes very poor. In other words, the observable becomes more and more smeared. However, by a clever choice of the probe state it is possible to control the measurement so that the average value of the first measurement remains the same.

The original definition of a weak value of an observable can easily be generalised to cover arbitrary pairs of observables.⁵

Definition 10.2 Let $\mathbf{E}, \mathbf{F} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be any two observables, $\varphi \in \mathcal{D}(x, \mathbf{E})$, $\|\varphi\| = 1$, and let $Y \in \mathcal{B}(\mathbb{R})$ be such that $\mathbf{F}(Y)\varphi \neq 0$. The *weak value* of $\mathbf{E}$ in a vector state φ conditioned by $\mathbf{F}(Y)$ is

$$\mathbf{E}_w(\varphi, \mathbf{F}(Y)) = \frac{\langle \varphi | \mathbf{F}(Y) \mathbf{E}[1] \varphi \rangle}{\langle \varphi | \mathbf{F}(Y) \varphi \rangle}. \quad (10.34)$$

If $\mathbf{E}$ is of the form $\mu * \mathbf{E}^A$, with $\mu[1] = 0$, so that $\mathbf{E}[1] = A$, and if $\mathbf{F}$ is a discrete observable with $\mathbf{F}(Y) = |\eta\rangle\langle\eta|$ for some unit vector $\eta \in \mathcal{H}$, then the definition reduces to the expression

⁵This subsection follows [8].

$$\mathbf{E}_w(\varphi, \mathbf{F}(Y)) = \frac{\langle \eta | A \varphi \rangle}{\langle \eta | \varphi \rangle},$$

which is the original definition of Aharonov et al. [9].

We shall show that the above definition is operationally meaningful in the sense that its real and imaginary parts can be obtained as conditional averages in two different sequential measurement schemes in the limit of zero interaction strength.

Consider the generalised standard measurement scheme $\mathcal{M}^\lambda$ for $\mathbf{E}$, and suppose that after this measurement of $\mu^\lambda * \mathbf{E}$, one performs (an accurate) measurement of $\mathbf{F}$, thus obtaining the (sequential joint) observable $\mathbf{M}^\lambda : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ of the form (10.16). If one then postselects only the values (x, y) with $y \in Y$ (for a fixed Y for which $\mathbf{M}_2^\lambda(Y)\varphi \neq 0$) and normalises the probabilities, one obtains a conditional probability measure

$$X \mapsto \frac{\langle \varphi | \mathcal{J}^\lambda(X)^*(\mathbf{F}(Y))\varphi \rangle}{\langle \varphi | \mathcal{J}^\lambda(\mathbb{R})^*(\mathbf{F}(Y))\varphi \rangle} = \frac{\langle \varphi | \mathbf{M}^\lambda(X \times Y)\varphi \rangle}{\langle \varphi | \mathbf{M}_2^\lambda(Y)\varphi \rangle}.$$

The real part of the weak value is now obtained by evaluating the first moment of this measure and then taking the limit of vanishing interaction strength:

$$\text{Re}(\mathbf{E}_w(\varphi, \mathbf{F}(Y))) = \lim_{\lambda \rightarrow 0} \int x \frac{\langle \varphi | \mathcal{J}^\lambda(dx)^*(\mathbf{F}(Y))\varphi \rangle}{\langle \varphi | \mathcal{J}^\lambda(\mathbb{R})^*(\mathbf{F}(Y))\varphi \rangle}. \quad (10.35)$$

Now notice that $\lim_{\lambda \rightarrow 0} \langle \varphi | \mathcal{J}^\lambda(\mathbb{R})^*(\mathbf{F}(Y))\varphi \rangle = \langle \varphi | \mathbf{F}(Y)\varphi \rangle$, so it is sufficient to consider only the numerator, for which we write $\Lambda^\lambda[1]$, with $\Lambda^\lambda(X) = \langle \varphi | \mathcal{J}^\lambda(X)^*(\mathbf{F}(Y))\varphi \rangle$. Assuming that $\phi \in \mathcal{D}(Q)$ and $\varphi \in \widetilde{\mathcal{D}}(x, \mathbf{E})$ (the latter being equivalent to $\varphi \otimes \psi \in \mathcal{D}(A) = \mathcal{D}(x, \mathbf{E}^A)$), and using the translation covariance of position, we can show that $e^{-i\lambda A \otimes P} \varphi \otimes \psi \otimes \phi \in \mathcal{D}(x, I \otimes I \otimes Q)$ for all $\lambda \geq 0$. It follows that the required average value is given by

$$\Lambda^\lambda[1] = \frac{1}{\lambda} \langle e^{-i\lambda A \otimes P} \varphi \otimes \psi \otimes \phi | \mathbf{F}(Y) \otimes I \otimes Q e^{-i\lambda A \otimes P} \varphi \otimes \psi \otimes \phi \rangle$$

If we then make the assumptions that $\langle \phi | Q \phi \rangle = 0$ and $\phi \in \mathcal{D}(QP) \cap \mathcal{D}(PQ)$ we can calculate the limit $\lambda \rightarrow 0$. Using again the translation covariance of position, we can actually calculate the limit in a rigorous manner and obtain

$$\lim_{\lambda \rightarrow 0} \Lambda^\lambda[1] = i \langle \mathbf{E}[1]\varphi | \mathbf{F}(Y)\varphi \rangle \langle \phi | PQ\phi \rangle - i \langle \mathbf{F}(Y)\varphi | \mathbf{E}[1]\varphi \rangle \langle \phi | QP\phi \rangle.$$

Here we have used Theorem 7.10, according to which the operator identity $V_\psi^* A V_\psi = \mathbf{E}[1]$ holds on the square-integrability domain $\widetilde{\mathcal{D}}(x, \mathbf{E})$.

To reach the desired conclusion we pose the condition $\langle \phi | QP\phi \rangle = \frac{i}{2}$, which is satisfied for instance by a Gaussian $\phi(x) = (2\pi\Delta^2)^{-1/4} \exp(-x^2/4\Delta^2)$. Under these assumptions, Eq. (10.35) is clearly valid, which shows that at least the real part of

the (generalised) weak value is accessible via measurements. We summarise these considerations in the following proposition.

Proposition 10.2 *Let $\mathcal{I}^\lambda$ be the instrument defined by the measurement scheme $\mathcal{M}^\lambda = \langle \mathcal{K}, \phi, \Phi^\lambda, \mathbf{Q}, f^\lambda \rangle$, and assume that $\phi \in \mathcal{D}(\mathbf{Q}P) \cap \mathcal{D}(P\mathbf{Q})$ and ϕ satisfies $\langle \phi | \mathbf{Q} \phi \rangle = 0$ and $\langle \phi | \mathbf{Q} P \phi \rangle = \frac{i}{2}$. Then for all $\varphi \in \widehat{\mathcal{D}}(x, \mathbf{E})$,*

$$\lim_{\lambda \rightarrow 0} \int x \frac{\langle \varphi | \mathcal{I}^\lambda(dx)^*(\mathbf{F}(Y)) \varphi \rangle}{\langle \varphi | \mathcal{I}^\lambda(\mathbb{R})^*(\mathbf{F}(Y)) \varphi \rangle} = \operatorname{Re} \frac{\langle \varphi | \mathbf{F}(Y) \mathbf{E}[1] \varphi \rangle}{\langle \varphi | \mathbf{F}(Y) \varphi \rangle}.$$

In order to obtain the imaginary part we need another measurement setup. Thus, consider a measurement scheme $\mathcal{N}^\lambda = \langle \mathcal{K}, \phi, \Phi^\lambda, \mathbf{P}, f^\lambda \rangle$ that differs from $\mathcal{M}^\lambda$ by the pointer observable: instead of monitoring shifts in the probe's position we now observe the boosts the probe obtains. Note that at this point the initial probe state is arbitrary. This change has a significant effect on the scheme. In particular, the measured observable becomes a trivial one, $X \mapsto \langle \phi | \mathbf{P}(\lambda X) \phi \rangle I$, so that no information is gained about the object system. However, the instrument $\mathfrak{K}^\lambda$ associated with this scheme is nontrivial as can be seen from the dual form

$$\mathfrak{K}^\lambda(X)^*(B) = \int_X |\hat{\phi}(x)|^2 V_\psi^* e^{i\lambda x A} (B \otimes I) e^{-i\lambda x A} V_\psi dx,$$

for all $B \in \mathcal{L}(\mathcal{H})$.

Now suppose that we again perform a sequential measurement where the second observable is $\mathbf{F}$. Then the calculations for the conditional average value and the limit of zero interaction are performed as before. In particular, by denoting $\Gamma^\lambda(X) = \langle \varphi | \mathfrak{K}^\lambda(X)^*(\mathbf{F}(Y)) \varphi \rangle$ and by assuming that $\phi \in \mathcal{D}(P)$ with $\langle \phi | P \phi \rangle = 0$ we find that

$$\lim_{\lambda \rightarrow 0} \Gamma^\lambda[1] = 2 \langle \phi | P^2 \phi \rangle \cdot \operatorname{Im} \langle \mathbf{F}(Y) \varphi | \mathbf{E}[1] \varphi \rangle$$

so that by assuming that $\langle \phi | P^2 \phi \rangle = \frac{1}{2}$ we get the desired result. For the Gaussians we have $\langle \phi | P^2 \phi \rangle = 1/4\Delta^2$ so that the above condition is satisfied with the choice $\Delta^2 = \frac{1}{2}$ for the variance. Once again we summarise this as a proposition.

Proposition 10.3 *Let $\mathfrak{K}^\lambda$ be the instrument defined by the measurement scheme $\mathcal{N}^\lambda = \langle \mathcal{K}, \phi, \Phi^\lambda, \mathbf{P}, f^\lambda \rangle$ where $\phi \in \mathcal{D}(P^2)$ is such that $\langle \phi | P \phi \rangle = 0$ and $\langle \phi | P^2 \phi \rangle = \frac{1}{2}$. Then for all $\varphi \in \widehat{\mathcal{D}}(x, \mathbf{E})$,*

$$\lim_{\lambda \rightarrow 0} \int x \frac{\langle \varphi | \mathfrak{K}^\lambda(dx)^*(\mathbf{F}(Y)) \varphi \rangle}{\langle \varphi | \mathfrak{K}^\lambda(\mathbb{R})^*(\mathbf{F}(Y)) \varphi \rangle} = \operatorname{Im} \frac{\langle \varphi | \mathbf{F}(Y) \mathbf{E}[1] \varphi \rangle}{\langle \varphi | \mathbf{F}(Y) \varphi \rangle}.$$

10.5 Repeatable Measurements

Intuitively a measurement of an observable is repeatable if its repeated application produces no new result. We formulate this intuitive idea together with some useful weaker variants.

Definition 10.3 A measurement $\mathcal{M}$ of an observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ and the associated instrument $\mathfrak{I}_{\mathcal{M}}$ are called

(a) *repeatable* if

$$\mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(Y)(\mathfrak{I}_{\mathcal{M}}(X)\varrho)] = \mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(Y \cap X)\varrho] \quad \text{for all } X, Y \in \mathcal{A}, \varrho \in \mathcal{S}(\mathcal{H});$$

(b) *first kind* if

$$E_{\varrho}(X) = E_{\mathfrak{I}_{\mathcal{M}}(\Omega)(\varrho)}(X) \quad \text{for all } X \in \mathcal{A}, \varrho \in \mathcal{S}(\mathcal{H});$$

(c) *value reproducible* if

$$E_{\varrho}(X) = 1 \implies E_{\mathfrak{I}_{\mathcal{M}}(\Omega)(\varrho)}(X) = 1 \quad \text{for all } X \in \mathcal{A}, \varrho \in \mathcal{S}(\mathcal{H}).$$

A repeatable measurement is of the first kind and a first kind measurement is value reproducible, but neither of these implications can be reversed in general. However, in the class of sharp observables these concepts coincide.

Theorem 10.3 For a measurement $\mathcal{M}$ of a sharp observable $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:

- (i) $\mathcal{M}$ is value reproducible;
- (ii) $\mathcal{M}$ is first kind;
- (iii) $\mathcal{M}$ is repeatable.

Proof The implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) are immediate. Assume that (i) holds. Then for any $X \in \mathcal{A}$ and for any $\varrho \in \mathcal{S}(\mathcal{H})$

$$\mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(\Omega)(E(X)\varrho E(X))E(X)] = \mathrm{tr}[\varrho E(X)],$$

that is,

$$E(X)\mathfrak{I}_{\mathcal{M}}^*(\Omega)(E(X))E(X) = E(X)$$

for all $X \in \mathcal{A}$. This entails

$$E(X)\mathfrak{I}_{\mathcal{M}}^*(X)(E(X))E(X) + E(X)\mathfrak{I}_{\mathcal{M}}^*(X^c)(E(X))E(X) = E(X).$$

By the positivity and linearity of $\mathfrak{J}_{\mathcal{M}}^*(X^c)$ one then has

$$\begin{aligned} 0 &\leq \mathbf{E}(X)\mathfrak{J}_{\mathcal{M}}^*(X^c)(\mathbf{E}(X))\mathbf{E}(X) \\ &\leq \mathbf{E}(X)\mathfrak{J}_{\mathcal{M}}^*(X^c)(I)\mathbf{E}(X) = \mathbf{E}(X)\mathbf{E}(X^c)\mathbf{E}(X) = \mathbf{E}(X \cap X^c) = 0. \end{aligned}$$

Thus $\mathbf{E}(X)\mathfrak{J}_{\mathcal{M}}^*(X^c)(\mathbf{E}(X))\mathbf{E}(X) = 0$ and so

$$\mathbf{E}(X)\mathfrak{J}_{\mathcal{M}}^*(X)(\mathbf{E}(X))\mathbf{E}(X) = \mathbf{E}(X).$$

But one also has $0 \leq \mathfrak{J}_{\mathcal{M}}^*(X)(\mathbf{E}(X)) \leq \mathfrak{J}_{\mathcal{M}}^*(X)(I) = \mathbf{E}(X)$. Since $\mathbf{E}(X)$ is a projection operator this means that

$$\mathfrak{J}_{\mathcal{M}}^*(X)(\mathbf{E}(X)) = \mathbf{E}(X)\mathfrak{J}_{\mathcal{M}}^*(X)(\mathbf{E}(X))\mathbf{E}(X).$$

Hence, $\mathfrak{J}_{\mathcal{M}}^*(X)(\mathbf{E}(X)) = \mathbf{E}(X)$, which establishes the repeatability of $\mathcal{M}$. □

We say that an observable $\mathbf{E}$ admits a repeatable measurement if there is a measurement of $\mathbf{E}$ which is repeatable.

Theorem 10.4 *Let Ω be a locally compact Hausdorff space. If an observable $\mathbf{E} : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ admits a repeatable measurement, then it is discrete.*

The proof can be found in the Appendix (Sect. 10.8).

This theorem has an immediate but important consequence: *no ‘continuous’ observable admits a repeatable measurement.*

10.6 Ideal Measurements

The repeatability, value reproducibility, and first kind properties introduced in Sect. 10.5 entail a limited form of nondisturbance: they ensure that if a system is in an eigenstate of the measured observable, then the measurement does not force it out of the corresponding eigenspace. A more stringent kind of minimal disturbance is captured by the notion of ideality. There are various possible formulations of this concept; we consider here two: p-ideality and d-ideality.

A measurement is ideal if it alters the measured system only to the extent that is necessary for obtaining an outcome: all the properties which are real in the initial state of the object system (in the sense that they occur with probability one) and which can be measured together with the measured observable remain real also in the final state of the system.

In the case of sharp observables this intuitive idea is formalised into the probabilistic concept of p-ideality. The following definition makes use of the notion of compatibility, to be studied in Chap. 11; this refers to the possibility of measuring some observables together. For sharp observables, compatibility is equivalent to their commutativity (Corollary 11.3).

Definition 10.4 A measurement $\mathcal{M}$ of a sharp observable $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ and the induced instrument $\mathcal{I}_{\mathcal{M}}$ are called *p-ideal* if for any state $\varrho \in \mathcal{S}(\mathcal{H})$ and for any sharp property $P \in \mathcal{P}(\mathcal{H})$ compatible with E , the following implication holds:

$$\text{if } \text{tr}[\varrho P] = 1, \text{ then } \text{tr}[\mathcal{I}_{\mathcal{M}}(\Omega)(\varrho)P] = 1. \quad (10.36)$$

When applied to the projections $P = E(X)$, the p-ideality of $\mathcal{M}$ implies its value reproducibility: for any state ϱ and any value set X ,

$$\text{if } E_{\varrho}(X) = 1, \text{ then } E_{\mathcal{I}_{\mathcal{M}}(\Omega)(\varrho)}(X) = 1. \quad (10.37)$$

By Theorem 10.3, this condition is equivalent to the repeatability of $\mathcal{M}$ in the case of sharp observables. Thus we have the following result.

Theorem 10.5 *A p-ideal measurement of a sharp observable is repeatable.*

Remark 10.7 Let E be a sharp observable with a value space Ω that is a locally compact Hausdorff space. If E admits a p-ideal measurement then it is discrete. $\triangleleft$

The von Neumann measurements of a sharp discrete observable are always repeatable and thus satisfy (10.37). However these measurements are not p-ideal except when they are Lüders measurements.

A sharp observable A on $\mathbb{R}$ which admits a p-ideal measurement is discrete, $A = \sum_i a_i P_i$. Using the discreteness of A , the p-ideality of a measurement $\mathcal{M}$ of A implies the following condition: for all $i = 1, 2, \dots$ and all ϱ ,

$$\text{if } \text{tr}[\varrho P_i] = 1, \text{ then } \mathcal{I}_{\mathcal{M}}(\{a_i\})(\varrho) = \varrho. \quad (10.38)$$

Indeed, if φ is a vector state for which $\langle \varphi | P_i \varphi \rangle = 1$, then, by p-ideality, the probability $\text{tr}[\mathcal{I}_{\mathcal{M}}(\mathbb{R})(P[\varphi])P[\varphi]] = 1$. But in this case the support projection of the state $\mathcal{I}_{\mathcal{M}}(\mathbb{R})(P[\varphi])$ is contained in $P[\varphi]$, which is possible only if $\mathcal{I}_{\mathcal{M}}(\mathbb{R})(P[\varphi]) = P[\varphi]$. Since $P[\varphi]$ is a vector state with $\text{tr}[\mathcal{I}_{\mathcal{M}}(\{a_i\})(P[\varphi])] = 1$, we finally have $\mathcal{I}_{\mathcal{M}}(\{a_i\})(P[\varphi]) = P[\varphi]$. By linearity the argument extends to arbitrary states ϱ for which $\text{tr}[\varrho P_i] = 1$. We shall see below that condition (10.38) is in fact equivalent to the p-ideality of $\mathcal{I}_{\mathcal{M}}$.

Condition (10.38) admits an immediate generalisation to arbitrary discrete observables E . Since this generalisation will turn out to be important for the objectification problem, we formulate it as the definition of the d-ideality of a measurement, the prefix d referring to the assumption of discreteness.

Definition 10.5 Let E be a discrete observable with the outcome space $\Omega = \{\omega_1, \omega_2, \dots\}$. A measurement $\mathcal{M}$ of E and its induced instrument $\mathcal{I}_{\mathcal{M}}$ are called *d-ideal* if the measurement does not change the state of $\mathcal{S}$ whenever a particular result is certain from the outset:

$$\text{if } E_\varrho(\{\omega_i\}) = 1, \quad \text{then } \mathfrak{J}(\{\omega_i\})(\varrho) = \varrho \quad (10.39)$$

for all $i = 1, 2, \dots$ and any $\varrho \in \mathcal{S}(\mathcal{H})$.

In general, a d-ideal measurement need not be repeatable nor first kind. Indeed the Lüders instrument, $\varrho \mapsto E_i^{1/2} \varrho E_i^{1/2}$, of a discrete two-valued observable $\mathbf{E}$ (with the generating effects E_1 and $E_2 = I - E_1$) is d-ideal but never repeatable unless the E_i are projection operators. Suppose the E_i have eigenvalue 1 with associated spectral projections E_i^1 , and let $U : \mathcal{H} \rightarrow \mathcal{H}$ be a unitary mapping which acts as an identity on the eigenspaces $E_i^1(\mathcal{H})$. Then the $\mathbf{E}$ -compatible instrument $\varrho \mapsto U E_i^{1/2} \varrho E_i^{1/2} U^{-1}$ is still d-ideal but not first kind, unless U commutes with the E_i .

The question of the structure of ideal, repeatable measurements has been a major issue since von Neumann's work [12]. These properties are crucial for the realistic interpretation of quantum mechanics insofar as the existence of measurements with these properties ensures the interpretation of possible measurement results as potential properties of the system. The repeatability of a measurement $\mathcal{M}$ requires the measured observable $\mathbf{E}$ to be discrete (Theorem 10.4). The d-ideality or the repeatability of a measurement $\mathcal{M}$ of a discrete observable $\mathbf{E}$ do not imply that $\mathbf{E}$ is a sharp observable. In order to attain this conclusion, one needs to postulate an additional property of the measurement, its nondegeneracy [13].

Definition 10.6 A measurement $\mathcal{M}$ of an observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is *nondegenerate* if the set of all possible final states $\{\varrho^f : \varrho \in \mathcal{S}(\mathcal{H})\}$ separates the set of effects; that is, for any $B \in \mathbf{E}(\mathcal{H})$,

$$\text{if } \text{tr}[\varrho^f B] = 0 \text{ for all } \varrho \in \mathcal{S}(\mathcal{H}), \quad \text{then } B = 0. \quad (10.40)$$

From the positivity of the effects and the additivity of instruments it follows that if $\text{tr}[\varrho^f B] = 0$ for a given effect B in some state ϱ , then $\text{tr}[\varrho^f(X)B] = 0$ for all $X \in \mathcal{A}$ in that state.

Proposition 10.4 A Lüders measurement of a sharp observable $A = \sum_i a_i P_i$ is nondegenerate.

Proof Assume that $B \in \mathbf{E}(\mathcal{H})$ is such that $\sum_i \text{tr}[B \mathfrak{J}_L(\{a_i\})(\varrho)] = 0$ for all $\varrho \in \mathcal{S}(\mathcal{H})$. Thus, in particular, for any vector state $\varphi \in \mathcal{H}$ and for any i ,

$$\begin{aligned} \text{tr}[B \mathfrak{J}_L(\{a_i\})(P[\varphi])] &= \text{tr}[B P_i P[\varphi] P_i] = \text{tr}[B^{1/2} P_i P[\varphi] P_i B^{1/2}] \\ &= \|B^{1/2} P_i \varphi\|^2 = 0. \end{aligned}$$

Hence for all vector states $\varphi \in \mathcal{H}$, $\sum_i B^{1/2} P_i \varphi = B^{1/2} \varphi = 0$, so that $B^{1/2} = 0$ and thus $B = 0$. $\square$

Lemma 10.1 If an observable $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ admits a repeatable nondegenerate measurement, then $\mathbf{E}$ is a sharp discrete observable.

Proof Let $\mathcal{M}$ be a nondegenerate repeatable measurement of $\mathbf{E}$. Let P be the support projection of $\mathfrak{I}_{\mathcal{M}}^*(\Omega)$ so that $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(I - P) = 0$, that is, for all states $\varrho \in \mathcal{S}(\mathcal{H})$, $\text{tr}[(I - P)\mathfrak{I}_{\mathcal{M}}(\Omega)(\varrho)] = 0$. Since $\mathcal{M}$ is nondegenerate, this means that $I - P = 0$, that is, $P = I$. By Lemma 10.5 of Sect. 10.8 this means that $\mathbf{E}$ is projection valued. $\square$

Theorem 10.6 *An observable $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ admits a repeatable, nondegenerate, d-ideal measurement if and only if $\mathbf{E}$ is a discrete sharp observable. In that case the measurement $\mathcal{M}$ is equivalent to a Lüders measurement of $\mathbf{E}$; that is, the induced instrument $\mathfrak{I}_{\mathcal{M}}$ is of the form*

$$\mathfrak{I}_{\mathcal{M}}(X)(\varrho) = \sum_{\omega_i \in X} \mathbf{E}(\{\omega_i\})\varrho\mathbf{E}(\{\omega_i\}) \quad (10.41)$$

for all $\varrho \in \mathcal{S}(\mathcal{H})$, $X \in \mathcal{B}(\mathbb{R})$.

Proof A discrete sharp observable admits a Lüders measurement and such a measurement is repeatable, nondegenerate, and d-ideal. If an observable $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ admits a repeatable, nondegenerate, d-ideal measurement $\mathcal{M}$, then $\mathbf{E}$ is a discrete sharp observable, $\omega_i \mapsto \mathbf{E}(\{\omega_i\}) = P_i$ and it remains to be shown that the instrument $\mathfrak{I}_{\mathcal{M}}$ has the form (10.41). Again, it suffices to show that $\mathfrak{I}_{\mathcal{M}}(\{\omega_i\})(\varrho) = P_i\varrho P_i$ for all i for all $\varrho \in \mathcal{S}(\mathcal{H})$, that is, $\mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) = P_i A P_i$ for all $A \in \mathcal{L}(\mathcal{H})$.

Let $A \in \mathcal{E}(\mathcal{H})$. Then $0 \leq \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) \leq \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(I) = P_i$ which implies that

$$\mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) = P_i \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) P_i.$$

By the linearity of the map $\mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})$ the above equality extends to any $A \in \mathcal{L}(\mathcal{H})$. Thus for any $A \in \mathcal{L}(\mathcal{H})$ and $\varrho \in \mathcal{S}(\mathcal{H})$ we get

$$\text{tr}[\varrho \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A)] = \text{tr}[\varrho P_i \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) P_i] = \text{tr}[P_i \varrho P_i \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A)]. \quad (*)$$

For any $\varrho \in \mathcal{S}(\mathcal{H})$, if $\text{tr}[\varrho P_i] \neq 0$, then $\varrho_i = P_i \varrho P_i / \text{tr}[\varrho P_i] \in \mathcal{S}(\mathcal{H})$ and $\text{tr}[\varrho_i P_i] = 1$. The d-ideality of the measurement implies that $\mathfrak{I}_{\mathcal{M}}(\{\omega_i\})(\varrho_i) = \varrho_i$, that is,

$$\mathfrak{I}_{\mathcal{M}}(\{\omega_i\})(P_i \varrho P_i) = P_i \varrho P_i.$$

We may thus continue the chain (*) to get

$$\begin{aligned} \text{tr}[\varrho \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A)] &= \text{tr}[P_i \varrho P_i \mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A)] = \text{tr}[\mathfrak{I}_{\mathcal{M}}(\{\omega_i\})(P_i \varrho P_i) A] \\ &= \text{tr}[P_i \varrho P_i A] = \text{tr}[\varrho P_i A P_i], \end{aligned}$$

showing that $\mathfrak{I}_{\mathcal{M}}^*(\{\omega_i\})(A) = P_i A P_i$ whenever $\text{tr}[\varrho P_i] \neq 0$. If $\text{tr}[\varrho P_i] = 0$, then also both $\mathfrak{I}_{\mathcal{M}}(\{\omega_i\})(\varrho) = 0$ and $P_i \varrho P_i = 0$. $\square$

We leave it as an exercise to show that a d-ideal measurement of a discrete sharp observable is nondegenerate. Furthermore, a result similar to Theorem 10.5 shows

that also d-ideality implies repeatability. Hence the considerations of this Section give rise to the following statement.

Corollary 10.1 *The d-ideal measurements of a sharp discrete observable are exactly the Lüders measurements.*

A Lüders measurement of a sharp observable is p-ideal. Thus the concepts of p-ideality and d-ideality are equivalent in the case of sharp discrete observables.

10.7 Correlations, Disturbance and Entanglement

A measurement $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$ of an observable $\mathbf{E}$ reproduces the probability distributions, $\mathbf{E}_\varrho = \mathbf{Z}_{\sigma^\sharp}$; in order to achieve this, the measurement coupling must effect an information transfer from the object to the probe. Correspondingly, the state $U(\varrho \otimes \sigma)U^*$ reached by the total system after the coupling phase is typically entangled and displays correlations between the observables of the object and probe. This goes together with an inevitable change, or *disturbance*, of the state of the object, $\varrho \mapsto \mathfrak{I}_{\mathcal{M}}(\Omega)(\varrho) \neq \varrho$.

In Sects. 10.5 and 10.6, various repeatability and ideality conditions were introduced and found to characterise conditions under which the disturbance of the object state caused by the measurement may be kept minimal. As these investigations already suggest, there is no measurement that would leave the system unchanged in every initial state. This result is formalised in Proposition 10.5. Moreover, the measured observable $\mathbf{E}$ may be even strongly correlated with the pointer observable $\mathbf{Z}$ in the state $U(\varrho \otimes \sigma)U^*$, so that this state is not a product state. This happens, for instance, in a repeatable minimal measurement of a discrete, nondegenerate sharp observable (Exercise). While there is no equally stringent impossibility statement for non-entangling measurements, Proposition 10.7 shows the peculiar structure of such measurements.

Proposition 10.5 *Let $\mathcal{M}$ be a measurement of an observable $\mathbf{E} \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ with associated instrument $\mathfrak{I}_{\mathcal{M}}$. Assume that $\mathfrak{I}_{\mathcal{M}}(\Omega)(P[\varphi]) = P[\varphi]$ for all vector states $\varphi \in \mathcal{H}$. Then $\mathbf{E}$ is a trivial observable.*

Proof Assume that $\mathfrak{I}_{\mathcal{M}}(\Omega)(P[\varphi]) = P[\varphi]$. Then for any $X \in \mathcal{A}$,

$$\mathfrak{I}_{\mathcal{M}}(X)(P[\varphi]) + \mathfrak{I}_{\mathcal{M}}(X^c)(P[\varphi]) = P[\varphi],$$

which is a convex decomposition of the extreme state $P[\varphi]$. Therefore, for all X and all vector states $\varphi \in \mathcal{H}$, $\mathfrak{I}_{\mathcal{M}}(X)(P[\varphi]) = \mathbf{E}_\varphi(X)P[\varphi]$.

We next show that $\mathbf{E}_\varphi(X) = \mathbf{E}_\psi(X)$ for any vector states φ, ψ . First, consider an arbitrary normalised vector φ and let ψ be any normalised vector orthogonal to φ . These two vectors span a two-dimensional subspace. Let ξ be an arbitrary normalised superposition of φ and ψ not orthogonal to either of them, and η be a unit vector

orthogonal to ξ in the span of φ, ψ . Then $P[\varphi] + P[\psi] = P[\xi] + P[\eta]$, and applying the linear operation $\mathcal{I}_{\mathcal{M}}(X)$, one has

$$\mathbf{E}_{\varphi}(X)P[\varphi] + \mathbf{E}_{\psi}(X)P[\psi] = \mathbf{E}_{\xi}(X)P[\xi] + \mathbf{E}_{\eta}(X)P[\eta].$$

This is a selfadjoint operator of rank not greater than 2, presented in two versions of spectral decomposition with respect to two distinct orthonormal bases. From the uniqueness of the spectral decomposition it follows immediately that the spectrum must be degenerate and the operator a multiple of a rank 2 projection. Hence one must have $\mathbf{E}_{\varphi}(X) = \mathbf{E}_{\psi}(X) = \mathbf{E}_{\xi}(X) = \mathbf{E}_{\eta}(X)$.

This shows that given any vector state φ , for all vector states ψ orthogonal to φ , the value of $\mathbf{E}_{\psi}(X)$ is independent of ψ and indeed equal to $\mathbf{E}_{\varphi}(X)$. Furthermore, the same value is obtained for any vector state ξ not orthogonal to φ . Hence for any X , there is a constant $\lambda(X) \in [0, 1]$ such that $\mathbf{E}_{\varphi}(X) = \lambda(X)$ for all vector states φ . The measure properties of the map $X \mapsto \lambda(X)$ follows directly from the corresponding property of the instrument, which now is $\mathcal{I}_{\mathcal{M}}(X)(P[\varphi]) = \lambda(X)P[\varphi]$. $\square$

Note that this proof does not make use of the complete positivity of the instrument.

The value reproducibility of a repeatable measurement of a discrete observable means that states for which an outcome has probability equal to 1 will preserve this property. Ideality is the stronger feature that all sharp properties that commute with the measured observable and have probability 1 will retain probability 1 in the state after the measurement. Thus, a Lüders measurement leaves unchanged all eigenstates of the measured observable. Any non-eigenstate will be changed; in particular, a vector state is transformed into a mixture of eigenstates. This *state disturbance* $\varrho \mapsto \mathcal{I}_L^A(\mathbb{R})(\varrho)$ due to a Lüders measurement of a sharp observable A can be witnessed by comparing the statistics $\mathbf{B}_{\varrho^f}$ of some suitable observable $\mathbf{B}$ in the final state $\varrho^f = \mathcal{I}_L^A(\mathbb{R})(\varrho)$ with the *undisturbed* statistics $\mathbf{B}_{\varrho}$. As the following Theorem due to Lüders shows, not every observable can be used to detect the disturbance. At the same time, this gives an operational characterisation of commutativity.

Theorem 10.7 (Lüders [4]) *Let $A = \sum_i a_i P_i$ be a discrete sharp observable and $\mathcal{I}_{A,L}$ its Lüders instrument. Let $\mathbf{B}$ be a sharp observable on $(\Omega, \mathcal{A})$. Then*

$$[P_i, \mathbf{B}(X)] = 0 \text{ for all } i, X \iff \mathcal{I}_{A,L}^*(\mathbb{R}) \circ \mathbf{B} = \mathbf{B}. \quad (10.42)$$

Thus, the statistics of any observable $\mathbf{B}$ that commutes with A remains unchanged by a Lüders measurement of A .

Interestingly, commutativity may not be necessary for the statistics of an unsharp observable to remain unchanged under a Lüders measurement [14].

Proposition 10.6 *Let $\mathbf{E} = \{E_1, E_2, \dots, E_N\}$ be a finite-outcome observable, and let $\mathcal{I}_{\mathbf{E},L}$ be its generalised Lüders instrument, defined by*

$$\mathcal{I}_{\mathbf{E},L}(\{i\})(\varrho) = E_i^{1/2} \varrho E_i^{1/2}, \quad i = 1, 2, \dots, N. \quad (10.43)$$

Let $B = \sum_{k=1}^K b_k B_k$ be an effect with a strictly decreasing sequence of eigenvalues $(b_k) \subset [0, 1]$ and spectral projections B_k , $k \leq K$, $K \in \mathbb{N} \cup \{\infty\}$. Then

$$[E_i, B] = 0 \quad \forall i = 1, 2, \dots, N \quad \Longleftrightarrow \quad \mathfrak{I}_{E,L}^*(\mathbb{R})(B) = B. \quad (10.44)$$

If $\mathbf{E}$ is a two-outcome observable, $\mathbf{E} = \{E_1, E_2\}$, then (10.44) holds for any effect B .

It is known that $\mathfrak{I}_{E,L}^*(\mathbb{R})(B) = B$ may hold without $\mathbf{E}$, B commuting if the spectral condition on B is violated and $\mathbf{E}$ has more than two outcomes [15].

Proposition 10.7 *Let $\mathcal{M} = (\mathcal{K}, \mathcal{Z}, \sigma, U)$ be a measurement scheme with $\sigma = P[\phi_0]$ and assume that for all $\varphi \in \mathcal{H}$, $U(\varphi \otimes \phi_0) = \varphi' \otimes \phi'$ for some unit vectors $\varphi' \in \mathcal{H}$, $\phi' \in \mathcal{K}$. Then U acts in one of the following two ways:*

- (a) $U(\varphi \otimes \phi_0) = V(\varphi) \otimes \phi'_0$, where V is an isometry in $\mathcal{H}$ and ϕ'_0 is a fixed unit vector in $\mathcal{K}$.
- (b) $U(\varphi \otimes \phi_0) = \varphi'_0 \otimes W\varphi$, where W is an isometry from $\mathcal{H}$ to $\mathcal{K}$ and φ'_0 is a fixed unit vector in $\mathcal{H}$.

In the first case, the measured observable is trivial, $\mathbf{E}(X) = \mu(X) I$, with $\mu(X) = \langle \phi' | \mathbf{Z}(X) \phi' \rangle$, and the associated instrument is given by $\mathfrak{I}(X)(P[\varphi]) = \mu(X) P[V\varphi]$. In the second case, $\mathbf{E}$ is given by $\mathbf{E}(X) = W^{-1} \mathbf{Z}(X) W$, with the constant instrument $\mathfrak{I}(X)(P[\varphi]) = \langle \varphi | \mathbf{E}(X) \varphi \rangle P[\varphi']$.

Proof Let $\{\varphi_n : n = 1, 2, \dots\}$ be an orthonormal basis of $\mathcal{H}$. There are systems of unit vectors $\varphi'_n \in \mathcal{H}$, $\phi'_n \in \mathcal{K}$ such that $U \varphi_n \otimes \phi_0 = \varphi'_n \otimes \phi'_n$. Due to the unitarity of U , all the vectors $\varphi'_n \otimes \phi'_n$ are mutually orthogonal. We show that one of two cases (a), (b) must hold:

- (a) $\{\varphi'_n\}_{n \in \mathbb{N}}$ is an orthonormal system, all ϕ'_n are parallel to ϕ'_1 ;
- (b) $\{\phi'_n\}_{n \in \mathbb{N}}$ is an orthonormal system, all φ'_n are parallel to φ'_1 .

For two vectors ψ, ξ which are mutually orthogonal, $\langle \psi | \xi \rangle = 0$, we will write $\psi \perp \xi$. Since U is unitary, this map sends orthogonal vector pairs to orthogonal pairs. Hence from $\varphi_1 \perp \varphi_2$ it follows that $\varphi'_1 \perp \varphi'_2$ or $\phi'_1 \perp \phi'_2$. Consider the first case. Then

$$U\left(\frac{1}{\sqrt{2}}(\varphi_1 + \varphi_2) \otimes \phi_0\right) = \varphi'_{12} \otimes \phi'_{12} = \frac{1}{\sqrt{2}}\varphi'_1 \otimes \phi'_1 + \frac{1}{\sqrt{2}}\varphi'_2 \otimes \phi'_2,$$

where $\varphi'_{12} \in \mathcal{H}$, $\phi'_{12} \in \mathcal{K}$ are some unit vectors. Since $\varphi'_1 \perp \varphi'_2$, it follows that $\phi'_2 = c\phi'_1$ with some $c \in \mathbb{C}$, $|c| = 1$. So we have

$$U((\varphi_1 + \varphi_2) \otimes \phi_0) = (\varphi'_1 + c\varphi'_2) \otimes \phi'_1.$$

Still considering the case $\varphi'_1 \perp \varphi'_2$, the relation $\varphi_2 \perp \varphi_3$ implies that $\varphi'_2 \perp \varphi'_3$ or $\phi'_2 \perp \phi'_3$. Suppose the latter holds. We show that this leads to a contradiction. Indeed this assumption gives $\varphi'_3 = c'\varphi'_2$ and thus

$$\begin{aligned}
U((\varphi_1 + \varphi_2 + \varphi_3) \otimes \phi_0) &= \sqrt{3}\varphi'_{123} \otimes \phi'_{123} \\
&= \varphi'_1 \otimes \phi'_1 + \varphi'_2 \otimes \phi'_2 + \varphi'_3 \otimes \phi'_3 \\
&= (\varphi'_1 + c\varphi'_2) \otimes \phi'_1 + \varphi'_3 \otimes \phi'_3
\end{aligned}$$

where φ'_{123} and ϕ'_{123} are some unit vectors. Recalling that $\phi'_2 = c\phi'_1$ and, by assumption, $\phi'_2 \perp \phi'_3$, then $\phi'_1 \perp \phi'_3$, and we see that $\varphi'_1 + c\varphi'_2 = c''\varphi'_3$ for some $c'' \neq 0$. Upon taking the inner product of both sides with φ'_1 , we get (since $\varphi'_1 \perp \varphi'_2$) that $\langle \varphi'_1 | \varphi'_1 \rangle = c'' \langle \varphi'_1 | \varphi'_3 \rangle = 0$ (since $\varphi'_3 = c'\varphi'_2 \perp \varphi'_1$). Hence $\varphi'_1 = 0$ which is a contradiction.

Thus the assumption is false and we can only have $\varphi'_2 \perp \varphi'_3$. Continuing inductively, we obtain that $\{\varphi'_i : i \in \mathbb{N}\}$ is an orthonormal system and all $\phi'_i = c_i \phi'_1$. Therefore, we obtain possibility (a) in the present case. Linearity then entails that $U(\varphi \otimes \phi_0) = V(\varphi) \otimes \phi'_0$ for all $\varphi \in \mathcal{H}$ and some isometric map V .

An analogous consideration can be applied in the second case of $\phi'_1 \perp \phi'_2$, thus leading to the possibility (b) and

$$U(\varphi \otimes \phi_0) = \sum_i \langle \varphi_i | \varphi \rangle U(\varphi_i \otimes \phi_0) = \varphi'_0 \otimes \sum_i \langle \varphi_i | \varphi \rangle \phi'_i = \varphi'_0 \otimes W(\varphi)$$

for all $\varphi \in \mathcal{H}$ and some isometric map $W : \mathcal{H} \rightarrow \mathcal{K}$. □

Proposition 10.7 shows that non-entangling measurements are either trivial or based on a dynamics that swap states. An example of the latter case is a measurement scheme for which the probe is a copy of the object system and the unitary coupling is the swap map $U : \varphi \otimes \phi \mapsto \phi \otimes \varphi$. For such maps it has been shown that they cannot occur as elements of a continuous dynamical group $t \mapsto U_t$ [16].

10.8 Appendix

This Appendix gives a proof of Theorem 10.4.

The notion of a repeatable measurement, which dates back to von Neumann's 1932 book [3], raises the important question of the structure of observables that admit such measurements. In particular, one may ask whether these observables are necessarily discrete. This question was finally solved affirmatively by Ozawa [17] and Łuczak [18] building on important contributions by Stinespring [19], Davies and Lewis [20] and Davies [13]. In this book measurement schemes are assumed to be of the form $(\mathcal{K}, \mathbf{Z}, \sigma, U)$ so that their accompanying instruments are completely positive. Therefore, we may follow the proof of [17] which rests on that property of an instrument. In [18] the same result is obtained without using complete positivity. We split the proof of this result into a series of lemmas.

Lemma 10.2 *Let $\mathcal{M}$ be a repeatable measurement of an observable $\mathbf{E}$ and let $\mathfrak{I}_{\mathcal{M}}$ be the associated instrument. Then for any $X, Y \in \mathcal{A}$ and $A \in \mathcal{L}(\mathcal{H})$ the following properties hold:*

- (i) $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)^2) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X))$;
- (ii) $\mathfrak{I}_{\mathcal{M}}^*(X \cap Y)(A) = \mathfrak{I}_{\mathcal{M}}^*(Y)(A\mathbf{E}(X)) = \mathfrak{I}_{\mathcal{M}}^*(Y)(\mathbf{E}(X)A)$;
- (iii) $\mathfrak{I}_{\mathcal{M}}^*(X)(A) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(A\mathbf{E}(X)) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)A)$.

Proof Property (i) is an immediate consequence of the repeatability of the measurement and we leave it for the reader. To verify (ii), let $(\mathcal{K}, \mathbf{P}, \pi, V)$ constitute a Stinespring type representation of $\mathfrak{I}_{\mathcal{M}}$ so that for any $X \in \mathcal{A}$ and $A \in \mathcal{L}(\mathcal{H})$ we have $\mathfrak{I}_{\mathcal{M}}(X)^*(A) = V^*\mathbf{P}(X)\pi(A)V$. Using the repeatability of $\mathfrak{I}_{\mathcal{M}}$ one computes that

$$\begin{aligned} & [\pi(\mathbf{E}(X))V - \mathbf{P}(X)V]^* [\pi(\mathbf{E}(X))V - \mathbf{P}(X)V] \\ &= \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)^2) - \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)) = 0 \end{aligned}$$

so that $V^*\pi(\mathbf{E}(X)) = V^*\mathbf{P}(X)$ and $\pi(\mathbf{E}(X))V = \mathbf{P}(X)V$. Thus for any $X, Y \in \mathcal{A}$ and $A \in \mathcal{L}(\mathcal{H})$,

$$\begin{aligned} \mathfrak{I}_{\mathcal{M}}(X \cap Y)(A) &= V^*\mathbf{P}(X \cap Y)\pi(A)V = V^*\mathbf{P}(X)\mathbf{P}(Y)\pi(A)V \\ &= V^*\pi(\mathbf{E}(X))\mathbf{P}(Y)\pi(A)V = V^*\mathbf{P}(Y)\pi(\mathbf{E}(X))\pi(A)V \\ &= V^*\mathbf{P}(Y)\pi(\mathbf{E}(X)A)V = \mathfrak{I}_{\mathcal{M}}(Y)^*(\mathbf{E}(X)A), \end{aligned}$$

as well as

$$\begin{aligned} \mathfrak{I}_{\mathcal{M}}(X \cap Y)(A) &= V^*\mathbf{P}(X \cap Y)\pi(A)V = V^*\pi(A)\mathbf{P}(X \cap Y)V \\ &= V^*\pi(A)\mathbf{P}(Y)\mathbf{P}(X)V = V^*\pi(A)\mathbf{P}(Y)\pi(\mathbf{E}(X))V \\ &= V^*\mathbf{P}(Y)\pi(A\mathbf{E}(X))V = \mathfrak{I}_{\mathcal{M}}^*(Y)(A\mathbf{E}(X)). \end{aligned}$$

Property (iii) is just a special case of (ii) with $Y = \Omega$. □

Let $\mathcal{W}$ be the von Neumann algebra generated by the range of $\mathbf{E}$, that is, $\mathcal{W} = \{\mathbf{E}(X) \mid X \in \mathcal{A}\}''$, and let P be the support projection of the (normal completely positive) map $\mathfrak{I}_{\mathcal{M}}^*(\Omega)$ restricted to $\mathcal{W}$, that is, P is the least projection in $\mathcal{W}$ such that $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(P) = I$.

Lemma 10.3 *For any $A \in \mathcal{W}$,*

$$\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(AP) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(PA) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(PAP).$$

Proof Using the representation $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) = V^*\pi(A)V$ we get, for each $\varphi, \psi \in \mathcal{H}$,

$$\begin{aligned}
\left| \langle (\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) - \mathfrak{I}_{\mathcal{M}}^*(\Omega)(AP))\varphi | \psi \rangle \right| &= \left| \langle \mathfrak{I}_{\mathcal{M}}^*(\Omega)(A - AP)\varphi | \psi \rangle \right| \\
&= \left| \langle V^* \pi(A - AP)V\varphi | \psi \rangle \right| \\
&= \left| \langle \pi(I - P)V\varphi | \pi(A)V\psi \rangle \right| \\
&\leq \|\pi(I - P)V\varphi\| \|\pi(A)V\psi\| = 0,
\end{aligned}$$

as $\|\pi(I - P)V\varphi\|^2 = \langle V^* \pi(I - P)V\varphi | \varphi \rangle = \langle \mathfrak{I}_{\mathcal{M}}^*(\Omega)(I - P)\varphi | \varphi \rangle = 0$. Thus, $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(AP)$. The other two statements are obtained similarly. $\square$

Lemma 10.4 *For any $A \in \mathcal{W}$, $A \geq 0$, if $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) = 0$, then $PAP = 0$.*

Proof Let $A \in \mathcal{W}$, $A \geq 0$, be such that $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A) = 0$, and let $\mathbf{A}$ be the spectral measure of A . If $X \in \mathcal{B}([0, \|A\|])$, then $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{A}(X)) = 0$ and $P \leq I - \mathbf{A}(X)$, i.e., $\mathbf{A}(X) \leq I - P$. This implies $P\mathbf{A}(X)P = 0$ and finally $PAP = 0$. $\square$

Lemma 10.5 *Let $\mathfrak{I}_{\mathcal{M}}$ be a repeatable instrument of an observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ and let P be the support projection of $\mathfrak{I}_{\mathcal{M}}^*(\Omega)$. The map $X \mapsto \Pi(X) = P\mathbf{E}(X)P$ constitutes a projection measure such that $\Pi(X) = \mathbf{E}(X)P = P\mathbf{E}(X)$ for all $X \in \mathcal{A}$ (and $\Pi(\Omega) = P$).*

Proof From Lemma 10.3 we get

$$\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\Pi(X)) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(P\mathbf{E}(X)P) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)),$$

whereas Lemmas 10.2 and 10.3 give

$$\begin{aligned}
\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\Pi(X)^2) &= \mathfrak{I}_{\mathcal{M}}^*(\Omega)(P\mathbf{E}(X)P\mathbf{E}(X)P) \\
&= \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)P\mathbf{E}(X)) \\
&= \mathfrak{I}_{\mathcal{M}}^*(\Omega)(P\mathbf{E}(X)) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(P) \\
&= \mathfrak{I}_{\mathcal{M}}^*(\Omega)(P\mathbf{E}(X)) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\mathbf{E}(X)).
\end{aligned}$$

Therefore, $\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\Pi(X) - \Pi(X)^2) = 0$, so that by the nondegeneracy of $\mathfrak{I}_{\mathcal{M}}^*(\Omega)$ on $P\mathcal{W}P$ (Lemma 10.4) we have $\Pi(X) = \Pi(X)^2$ for all $X \in \mathcal{A}$, showing that Π is a projection measure. A direct computation gives

$$\mathfrak{I}_{\mathcal{M}}^*(\Omega)((\Pi(X) - \mathbf{E}(X)P)^*(\Pi(X) - \mathbf{E}(X)P)) = 0,$$

which, by Lemma 10.4, entails that $\Pi(X) = \mathbf{E}(X)P$, and thus also $\Pi(X) = \Pi(X)^* = P\mathbf{E}(X)$. $\square$

Theorem 10.8 *Let $(\Omega, \mathcal{B}(\Omega))$ be the Borel space associated with a locally compact Hausdorff space Ω . If an observable $\mathbf{E} : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ admits a repeatable measurement, then it is discrete.*

Proof Let $\mathcal{M}$ be a repeatable measurement of $\mathbf{E}$, P the support projection of $\mathfrak{I}_{\mathcal{M}}^*(\Omega)$, and let Π be the projection valued measure of Lemma 10.5 so that $\mathbf{E}(X) =$

$\mathfrak{I}_{\mathcal{M}}^*(X)(I) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(E(X)) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(PE(X)P) = \mathfrak{I}_{\mathcal{M}}^*(\Omega)(\Pi(X))$ for all $X \in \mathcal{A}$. Since Π is multiplicative and $\mathcal{H}$ is separable there is a countable set $X_o \subset \Omega$ such that the measure $\mathcal{B}(\Omega) \ni X \mapsto \Pi(X \cap X_o) \in \mathcal{L}(\Pi(X_o)(\mathcal{H}))$ is discrete and $\mathcal{B}(\Omega) \ni X \mapsto \Pi(X \cap X_o^c) \in \mathcal{L}(\Pi(X_o^c)(\mathcal{H}))$ is continuous.

Let $Q = \Pi(X_o^c) = P - \Pi(X_o)$ and define $\Psi_o(A) = Q\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A)Q$ for all $A \in \mathcal{L}(\mathcal{QH})$. Then $\Psi_o(Q) = Q$, showing that Ψ_o is a unital normal positive linear map $\mathcal{L}(\mathcal{QH}) \rightarrow \mathcal{L}(\mathcal{QH})$. Therefore, by Proposition 6.1, Ψ_o is the dual of a trace preserving positive linear map $\Phi_o : \mathcal{T}(\mathcal{QH}) \rightarrow \mathcal{T}(\mathcal{QH})$. Then for any $A \in \mathcal{L}(\mathcal{QH})$, and for all $X \in \mathcal{B}(\Omega)$, $T \in \mathcal{T}(\mathcal{QH})$,

$$\begin{aligned} \text{tr}[A\Pi(X \cap X_o^c)\Phi_o(\varrho)] &= \text{tr}[\Psi_o(A\Pi(X \cap X_o^c))\varrho] \\ &= \text{tr}[Q\mathfrak{I}_{\mathcal{M}}^*(\Omega)(A\Pi(X \cap X_o^c))Q\varrho] \\ &= \text{tr}[Q\mathfrak{I}_{\mathcal{M}}^*(\Omega)(\Pi(X \cap X_o^c)A)Q\varrho] \\ &= \text{tr}[\Psi_o(\Pi(X \cap X_o^c)A)\varrho] \\ &= \text{tr}[\Pi(X \cap X_o^c)A\Phi_o(\varrho)] \\ &= \text{tr}[A\Phi_o(\varrho)\Pi(X \cap X_o^c)], \end{aligned}$$

from which it follows that $\Pi(X \cap X_o^c)\Phi_o(\varrho) = \Phi_o(\varrho)\Pi(X \cap X_o^c)$ for all $X \in \mathcal{B}(\Omega)$ and $T \in \mathcal{T}(\mathcal{QH})$. Since the measure $X \mapsto \Pi(X \cap X_o^c)$ is continuous this implies that $\Phi_o(T) = 0$ for all $T \in \mathcal{T}(\mathcal{QH})$. That is, Φ_o , and hence Ψ_o , is the null map [13, Theorem 4.3.3]. Therefore, $Q = \Psi_o(Q) = 0$. $\square$

10.9 Exercises

1. Prove Proposition 10.1.
2. Verify Eq. (10.12).
3. Verify Eq. (10.20).
4. Verify Eq. (10.25).
5. Verify Eq. (10.31).
6. Verify Eqs. (10.32) and (10.33).
7. Determine the sequential observable obtained from a standard position measurement followed by a (sharp) momentum measurement.
8. Let A, B be commuting selfadjoint operators with closed discrete spectra, and let $\mathcal{M}_L^A, \mathcal{M}_L^B$ be standard schemes of Lüders measurements of the sharp observables corresponding to A, B . Show that the sequential application of these measurements is order independent and equivalent to their simultaneous application.
9. Consider a measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U)$ and assume that the measured observable E is a spectral measure. Show that for any $X \in \mathcal{A}$ the projections $I \otimes \mathbf{Z}(X)$ and $V_\phi V_\phi^*$ commute with each other.
10. Show that the repeatability condition of Definition 10.3 can be written equivalently in any of the following ways:

$$\mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(X)^2 \varrho] = \mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(X) P[\varphi]]; \quad (10.45)$$

$$\mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(X^c) \mathfrak{I}_{\mathcal{M}}(X) \varrho] = 0; \quad (10.46)$$

$$\mathrm{tr}[\mathfrak{I}_{\mathcal{M}}(Y) \mathfrak{I}_{\mathcal{M}}(X) \varrho] = 0 \quad \text{for all disjoint } X \text{ and } Y; \quad (10.47)$$

$$\mathrm{tr}[\varrho^f(X) \mathbf{E}(X)] = 1 \quad (\text{whenever } \mathbf{E}_{\varrho}(X) \neq 0); \quad (10.48)$$

$$\mathbf{E}(X) \varrho^f(X) = \varrho^f(X); \quad (10.49)$$

$$\mathbf{E}(Y \cap X) = \mathfrak{I}_{\mathcal{M}}^*(X)(\mathbf{E}(Y)); \quad (10.50)$$

$$\mathbf{E}(X) = \mathfrak{I}_{\mathcal{M}}^*(X)(\mathbf{E}(X)); \quad (10.51)$$

$$\mathfrak{I}_{\mathcal{M}}^*(X)(\mathbf{E}(X^c)) = 0; \quad (10.52)$$

each condition being valid for any $X, Y \in \mathcal{A}$ and $\varrho \in \mathcal{S}(\mathcal{H})$.

11. Verify the implication (10.37).
12. Show that a d-ideal measurement of a discrete sharp observable is nondegenerate.
13. Consider a minimal measurement $\mathcal{M}_m$ of a discrete sharp observable $A = \sum_i a_i P_i$ as given in Example 10.2. Compute the correlation coefficient $\mathrm{cor}(A, Z, P[U(\varphi \otimes \phi)])$ (as defined in Sect. 9.5). Show that if the generating vectors $\{\psi_{ij}\}$ form an eigenbasis of A , then $\mathrm{cor}(A, Z, P[U(\varphi \otimes \phi)]) = 1$ for all initial vector states φ for which the correlation coefficient is well defined.
14. Continue the previous exercise and compute the correlations between the final component states arising from this measurement scheme. Also compute the correlations between the values of A and Z , that is, the correlations between the sharp properties P_i and $P[\phi_i]$.
15. Show that the disturbed momentum $\tilde{\mathbf{P}}$ in the standard (approximate) position measurement, Eq. (10.32) of Sect. 10.4, is of the form $\tilde{\mathbf{P}} = \nu * \mathbf{P}$, with the probability measure $d\nu = \frac{1}{\lambda} |\hat{\phi}(\frac{p}{\lambda})|^2 dp$, where $\phi \in L^2(\mathbb{R})$, $\|\phi\| = 1$, is the initial probe state and λ is the coupling parameter.
16. Consider the von Neumann model of an approximate position measurement, Sect. 10.4. Compute the correlation $\mathrm{cor}(\mathbf{Q}^e, \mathbf{Q}_p, P[U(\varphi \otimes \phi)])$. Assume that the initial probe state is a Gaussian concentrated at the origin. Study the behaviour of the correlation coefficient as a function of the width of the Gaussian probe state ϕ .
17. Consider a vector state $\Psi \in \mathcal{H} \otimes \mathcal{K}$ that is not a product state. Use the polar decomposition of Ψ to construct observables which are strongly correlated in this state.

References

1. Pellonpää, J.-P., Tukiainen, M.: Minimal normal measurement models of quantum instruments. [arXiv:1509.08886](https://arxiv.org/abs/1509.08886)
2. Beltrametti, E., Cassinelli, G., Lahti, P.: Unitary measurements of discrete quantities in quantum mechanics. *J. Math. Phys.* **31**(1), 91–98 (1990)

3. 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)
4. Lüders, G.: Über die Zustandsänderung durch den Meßprozeß. *Ann. Phys. (Leipzig)*, **443**(5-8):322–328 (1950). English Translation by Kirkpatrick, K.A.: *Ann. Phys. (Leipzig)* **15**(9), 663–670 (2006)
5. Pellonpää, J.-P.: Complete quantum measurements break entanglement. *Phys. Lett. A* **376**(46), 3495–3498 (2012)
6. Pellonpää, J.-P.: Complete measurements of quantum observables. *Found. Phys.* **44**(1), 71–90 (2014)
7. Pellonpää, J.-P.: On coexistence and joint measurability of rank-1 quantum observables. *J. Phys. A* **47**(5), 052002 (2014)
8. Haapasalo, E., Lahti, P., Schultz, J.: Weak versus approximate values in quantum state determination. *Phys. Rev. A* **84**, 052107 (2011)
9. Aharonov, Y., Albert, D.Z., Vaidman, L.: How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.* **60**, 1351–1354 (1988)
10. Dressel, J., Malik, M., Miatto, F.M., Jordan, A.N., Boyd, R.W.: Colloquium: Understanding quantum weak values: basics and applications. *Rev. Mod. Phys.* **86**(1), 307–316 (2014)
11. Aharonov, Y., Botero, A.: Quantum averages of weak values. *Phys. Rev. A* **72**, 052111 (2005)
12. von Neumann, J.: Die Eindeutigkeit der Schrödingerschen Operatoren. *Math. Ann.* **104**(1), 570–578 (1931)
13. Davies, E.B.: *Quantum Theory of Open Systems*. Academic Press London, New York (1976)
14. Busch, P., Singh, J.: Lüders theorem for unsharp quantum measurements. *Phys. Lett. A* **249**, 10–12 (1998)
15. Arias, A., Gheondea, A., Gudder, S.: Fixed points of quantum operations. *J. Math. Phys.* **43**(12), 5872–5881 (2002)
16. Busch, P.: The role of entanglement in quantum measurement and information processing. *Int. J. Theor. Phys.* **42**(5), 937–941 (2003)
17. Ozawa, M.: Quantum measuring processes of continuous observables. *J. Math. Phys.* **25**(1), 79–87 (1984)
18. Łuczak, A.: *Instruments on von Neumann algebras*. Institute of Mathematics, Łódź University (1986)
19. Stinespring, W.F.: Positive functions on C^* -algebras. *Proc. Am. Math. Soc.* **6**, 211–216 (1955)
20. Davies, E.B., Lewis, J.T.: An operational approach to quantum probability. *Commun. Math. Phys.* **17**, 239–260 (1970)

Chapter 11

Joint Measurability

It is one of the key features of quantum mechanics that not all of the observables of this theory can be measured jointly; in other words, many pairs or larger families of observables are *incompatible*. We already observed that measurements necessarily disturb the system under investigation; here we study the phenomenon that measurements of incompatible observables also influence each other if one attempts to apply them jointly or sequentially. In order to quantify this irreducible disturbance and to identify possible ways of mitigating the resulting limitation of measurability, we first make precise minimal requirements of what constitutes a joint measurement of two or more observables. We have already encountered some formulations of this notion in Sects. 1.2 and 10.3. In Chap. 13 we proceed to present elements of a theory of approximate joint measurements of incompatible observables.

Insofar as the purpose of a measurement is to determine the values of observables and their distributions, a joint measurement of, say, two observables should be represented by a suitably defined observable that allows one to account for the joint occurrence of a pair of outcomes and their probability distribution for each state. There are several distinct ways of specifying this idea. Some of them are equivalent, some of different degrees of generality. We start with presenting those formulations which turn out to be equivalent and then proceed to indicate a few alternatives. We also explore some characterisations of pairs of observables that are jointly measurable.

11.1 Definitions and Basic Results

The subject of quantum incompatibility and the issue of joint measurability had a long history in the foundations of quantum mechanics. Accordingly, a variety of alternative terms were introduced besides joint or simultaneous measurability, notably compatibility, commensurability and coexistence. We will use the terms

joint measurability and compatibility interchangeably and refer to other notions, like coexistence, where appropriate. Pairs of observables that are not jointly measurable will be called *incompatible*.

In the following we consider observables E_1 and E_2 with value spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$. We recall that a positive operator bimeasure $B : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ is a *biobservable* if the marginal measures $X \mapsto B_1(X) = B(X, \Omega_2)$ and $Y \mapsto B_2(Y) = B(\Omega_1, Y)$ are observables. If G is an observable with the value space $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ then the effects $G_1(X) = G(X \times \Omega_2)$ and $G_2(Y) = G(\Omega_1 \times Y)$ constitute its marginal observables.

Definition 11.1 Let $E_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $E_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ be any two observables.

- (a) E_1 and E_2 are *functions* of an observable M , with the value space $(\mathcal{E}, \mathcal{B})$, if there are (measurable) functions $f_1 : \mathcal{E} \rightarrow \Omega_1$ and $f_2 : \mathcal{E} \rightarrow \Omega_2$ such that for each $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$,

$$E_1(X) = M(f_1^{-1}(X)), \quad E_2(Y) = M(f_2^{-1}(Y)).$$

- (b) E_1 and E_2 are *smearings* of an observable E , with the value space $(\Omega, \mathcal{A})$, if there are Markov kernels $p_1 : \mathcal{A}_1 \times \Omega \rightarrow [0, 1]$ and $p_2 : \mathcal{A}_2 \times \Omega \rightarrow [0, 1]$ such that for each $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$,

$$E_1(X) = \int_{\Omega} p_1(X, \omega) dE(\omega), \quad E_2(Y) = \int_{\Omega} p_2(Y, \omega) dE(\omega).$$

- (c) E_1 and E_2 have a *biobservable* if there is a biobservable $B : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ such that $B_1 = E_1$ and $B_2 = E_2$.
- (d) E_1 and E_2 have a *joint observable* if there is an observable $G : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ such that $G_1 = E_1$ and $G_2 = E_2$.

Any two observables that can be measured jointly in the sense of the measurement theory of Sect. 10.3 are functions of a third observable, and conversely. Clearly, if E_1 and E_2 are functions of E , then they are also smearings of E with the kernels $p_1(X, \omega) = \chi_X(f_1(\omega))$ and $p_2(Y, \omega) = \chi_Y(f_2(\omega))$. Conversely, if two observables are smearings of a third observable it is not entirely obvious, but nevertheless true in certain circumstances, that they are also functions of an observable (Theorem 11.1).

Biobservables arise in the context of sequential measurements and, more naturally, in detection schemes with two independent registration devices. Their margins, the partial observables, arise with ignoring one of the outcomes. If the observables E_1 and E_2 are functions of a third observable E , then

$$B_{\varrho}(X, Y) = E_{\varrho}(f_1^{-1}(X) \cap f_2^{-1}(Y)),$$

where $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$, $\varrho \in \mathcal{S}(\mathcal{H})$, defines a biobservable B of E_1 and E_2 . If observables E_1 and E_2 have a joint observable E , then they are also functions of E ,

namely, $E_i = E \circ \pi_i^{-1}$, where $\pi_i((\omega_1, \omega_2)) = \omega_i$, $i = 1, 2$. This leaves the question whether biobservables can be extended to joint observables.

By Theorem 4.2 such an extension is obtained whenever the value spaces $(\Omega_i, \mathcal{A}_i)$ have the property (D). This is the case, in particular, when the value spaces are closed or open subspaces of $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ (cf. Proposition 4.9). It then holds that $\mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2) = \mathcal{B}(\Omega_1 \times \Omega_2)$ (see Proposition 4.10).

Theorem 11.1 *Assume that the value spaces of the observables $E_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $E_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ have the property (D). The following conditions are equivalent:*

- (i) E_1 and E_2 have a biobservable;
- (ii) E_1 and E_2 have a joint observable;
- (iii) E_1 and E_2 are functions of a third observable;
- (iv) E_1 and E_2 are smearings of a third observable.

Proof Theorem 4.2 gives the implication (i) $\Rightarrow$ (ii). It remains to show that (iv) $\Rightarrow$ (i). Let $(\Omega, \mathcal{A})$ be a measurable space, $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ an observable, and p_1 and p_2 Markov kernels such that $\int_{\Omega} p_i(X_i, \omega) dE(\omega) = E_i(X_i)$ for all $X_i \in \mathcal{A}_i$ where $i = 1, 2$. Define a positive operator bimeasure $B : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ by

$$B(X_1, X_2) = \int_{\Omega} p_1(X_1, \omega) p_2(X_2, \omega) dE(\omega).$$

Since $E_1(X_1) = B(X_1, \Omega_2)$ and $E_2(X_2) = B(\Omega_1, X_2)$ for all $X_i \in \mathcal{A}_i$, it follows that B is a biobservable for E_1 and E_2 . $\square$

Definition 11.2 Two (or more) observables are said to be *jointly measurable* or *compatible* if they have one and thus all the properties of Theorem 11.1.

We stress that this definition presupposes the assumption that the value spaces of observables have the property (D).

For compatible observables E_1 and E_2 , the distributions $E_{1,\varrho}$, $E_{2,\varrho}$ can be inferred from single measurement outcome distributions like M_{ϱ} , E_{ϱ} , B_{ϱ} , G_{ϱ} . In addition, the probabilities $E_{1,\varrho}(X)$, $X \in \mathcal{A}_1$, and $E_{2,\varrho}(Y)$, $Y \in \mathcal{A}_2$, are bounded from below by the pair probabilities

$$M_{\varrho}(f_1^{-1}(X) \cap f_2^{-1}(Y)), \int_{\Omega} p_1(X, \omega) p_2(Y, \omega) dE_{\varrho}(\omega), B_{\varrho}(X, Y), G_{\varrho}(X \times Y).$$

By Lemma 4.6 and Proposition 4.8 these lower bounds are optimal if one of the observables E_1 or E_2 is sharp. Indeed, in this case we have

$$\begin{aligned} M(f_1^{-1}(X) \cap f_2^{-1}(Y)) &= \int_{\Omega} p_1(X, \omega) p_2(Y, \omega) dE(\omega) = B(X, Y) = G(X \times Y) \\ &= E_1(X) E_2(Y) = E_2(Y) E_1(X) = E_1(X) \wedge_{E(\mathcal{H})} E_2(Y) \end{aligned}$$

for each $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$, so that, for each state ϱ one has

$$\begin{aligned} \mathbf{M}_\varrho(f_1^{-1}(X) \cap f_2^{-1}(Y)) &= \int_{\Omega} p_1(X, \omega) p_2(Y, \omega) d\mathbf{E}_\varrho(\omega) \\ &= \mathbf{B}_\varrho(X, Y) = \mathbf{G}_\varrho(X \times Y) \\ &= \text{tr}[\varrho \mathbf{E}_1(X) \mathbf{E}_2(Y)] = \text{tr}[\varrho \mathbf{E}_1(X) \wedge_{E(\mathcal{H})} \mathbf{E}_2(Y)]. \end{aligned}$$

We take these results as an operational justification for Definition 11.2. Note, however, that for arbitrary observables $\mathbf{E}_1$, $\mathbf{E}_2$ the above pair probabilities need not be unique nor optimal.

The above results have an immediate but important application for sequential measurements. We formulate it in the following proposition, recalling from Sect. 10.3 that no topological assumptions are needed here.

Proposition 11.1 *Assume observables $\mathbf{E}_1$ and $\mathbf{E}_2$ are measured sequentially, with the $\mathbf{E}_1$ -measurement $\mathcal{M}_1$ performed first. The resulting bioobservable $\mathbf{B}_{12} : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ extends to a unique joint observable $\mathbf{E}_{12} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$, with the margins $\mathbf{E}_1$ and $\tilde{\mathbf{E}}_2 : Y \mapsto \mathcal{I}_1(\Omega_1)^*(\mathbf{E}_2(Y))$. If $\mathbf{E}_1$ is sharp, then $\mathbf{E}_1$ commutes with $\tilde{\mathbf{E}}_2$, that is, all the effects $\mathbf{E}_1(X)$ and $\mathbf{E}_2(Y)$ commute with each other, and their product equals $\mathbf{B}_{12}(X, Y) = \mathbf{E}_{12}(X \times Y)$:*

$$\mathbf{B}_{12}(X, Y) = \mathbf{E}_{12}(X \times Y) = \mathbf{E}_1(X) \tilde{\mathbf{E}}_2(Y) = \tilde{\mathbf{E}}_2(Y) \mathbf{E}_1(X).$$

Example 11.1 We illustrate Theorem 11.1 in the case of finite observables. Assume $\mathbf{E} = \{E_1, E_2, \dots, E_M\}$ and $\mathbf{F} = \{F_1, F_2, \dots, F_N\}$ satisfy condition (ii) of Theorem 11.1, that is, there exist an observable $\mathbf{G} = \{G_{ij} : i = 1, \dots, M, j = 1, \dots, N\}$ such that $\mathbf{E}$ and $\mathbf{F}$ are margins of $\mathbf{G}$; this can also be expressed directly in terms of smearings:

$$E_i = \sum_{j=1}^N G_{ij} = \sum_{k\ell} \delta_{ik} G_{k\ell}, \quad F_j = \sum_{i=1}^M G_{ij} = \sum_{k\ell} \delta_{j\ell} G_{k\ell}. \quad (11.1)$$

Here $\lambda_{ik} = \delta_{ik}$ and $\mu_{j\ell} = \delta_{j\ell}$ are stochastic matrices, which constitute discrete versions of Markov kernels. Hence (iii) and (iv) of Theorem 11.1 follow.

Assume then that (iv) holds, that is, $\mathbf{E}$, $\mathbf{F}$ are smearings of $\mathbf{H} = \{H_1, \dots, H_K\}$, with stochastic matrices (λ_{ik}) , (μ_{jk}) . Now notice that

$$\begin{aligned} E_i &= \sum_k \lambda_{ik} H_k = \sum_j \sum_k \lambda_{ik} \mu_{jk} H_k, \\ F_j &= \sum_k \mu_{jk} H_k = \sum_i \sum_k \lambda_{ik} \mu_{jk} H_k, \end{aligned} \quad (11.2)$$

so we obtain a joint observable $\mathbf{G}$ for $\mathbf{E}$ and $\mathbf{F}$ via $G_{ij} = \sum_k \lambda_{ik} \mu_{jk} H_k$. Then (ii) and (iii) follow. Notice that, in the discrete case, items (i) and (ii) are trivially equivalent. $\triangleleft$

Remark 11.1 Suppose that observable $\mathbf{E}_1$ is rank-1 and $\mathbf{E}_2$ is arbitrary. It can be proved that $\mathbf{E}_1$ and $\mathbf{E}_2$ have a joint observable $\mathbf{G}$ if and only if $\mathbf{E}_2$ is a smearing of $\mathbf{E}_1$ [1]. In this case, $\mathbf{G}$ is also a smearing of $\mathbf{E}_1$. $\triangleleft$

11.2 Alternative Definitions

In this section we briefly point out some more general ideas on joint measurability which, however, turn out to be of more limited use.

Let $\text{ran}(\mathbf{E}) = \{\mathbf{E}(X) \mid X \in \mathcal{A}\}$ denote the range of an observable. Observables $\mathbf{E}_1$ and $\mathbf{E}_2$ are *coexistent* if there is an observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ such that

$$\text{ran}(\mathbf{E}_1) \cup \text{ran}(\mathbf{E}_2) \subset \text{ran}(\mathbf{E}).$$

The notion of coexistence of observables probably describes the most general way of performing a measurement that results in statistics which contain the statistics of the two given observables. However, this definition does not provide the means for obtaining the association between the sets $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$ on the one hand and the sets Z_X , $Z_Y \in \mathcal{A}$ for which $\mathbf{E}(Z_X) = \mathbf{E}_1(X)$ and $\mathbf{E}(Z_Y) = \mathbf{E}_2(Y)$. The mere requirement of the existence of Z_X , Z_Y is thus not sufficient to make the definition of coexistence operationally useful. Moreover, this notion seems to be open to characterisations only after further specifications (see Sects. 11.3 and 11.4).

The following example demonstrates that the notion of coexistence of observables is, indeed, more general—and thus weaker—than any of the equivalent properties of Theorem 11.1.

Example 11.2 ([2]) Let $\{|1\rangle, |2\rangle, |3\rangle\}$ be an orthonormal basis in $\mathcal{H} = \mathbb{C}^3$ and let $\psi = \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle + |3\rangle)$. The effects $A_i = \frac{1}{2}(I - |i\rangle\langle i|)$, $i = 1, 2, 3$, and $B_1 = \frac{1}{2}|\psi\rangle\langle\psi|$, $B_2 = I - B_1$, constitute the three- and two-outcome observables $\mathbf{A} : i \mapsto A_i$ and $\mathbf{B} : j \mapsto B_j$, which are coexistent. Indeed, the range of the five-outcome observable $\mathbf{E}$, defined through the effects

$$\left\{ \frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|2\rangle\langle 2|, \frac{1}{2}|3\rangle\langle 3|, B_1, \frac{1}{2}I - B_1 \right\},$$

contains the ranges of $\mathbf{A}$ and $\mathbf{B}$. Assume now that $\mathbf{A}$ and $\mathbf{B}$ have also a joint observable $\mathbf{F} : (i, j) \mapsto F_{ij}$ so that $A_i = \sum_{j=1}^2 F_{ij}$ and $B_j = \sum_{i=1}^3 F_{ij}$. Since the range of B_1 is one dimensional, one has $F_{i1} = \lambda_i B_1$, $i = 1, 2, 3$, for some $\lambda_i \geq 0$, and therefore $A_i = \lambda_i B_1 + F_{i2}$. But then $0 = \langle i | A_i | i \rangle = \frac{\lambda_i}{2} |\langle i | \psi \rangle|^2 + \langle i | F_{i2} | i \rangle$ which implies that $\lambda_i = 0$ and hence $F_{i1} = 0$ for all i . Therefore, $B_1 = 0$, which is a contradiction. $\triangleleft$

Proposition 11.2 *Two-valued observables are coexistent if and only if they are compatible.*

Proof Let $\{\omega, \omega'\}$ and $\{\xi, \xi'\}$ be the outcome sets of the observables E_1 and E_2 , with $\text{ran}(E_i) = \{0, A_i, I - A_i, I\}$, $i = 1, 2$, and let E be an observable such that $E(X) = A_1$ and $E(Y) = A_2$ for some $X, Y \in \mathcal{A}$. Consider the partition $\mathcal{R} = \{X \cap Y, X^c \cap Y, X \cap Y^c, X^c \cap Y^c\}$ of the value space Ω of E into disjoint $\mathcal{A}$ -sets, and let

$$1 \mapsto E(X \cap Y), \quad 2 \mapsto E(X^c \cap Y), \quad 3 \mapsto E(X \cap Y^c), \quad 4 \mapsto E(X^c \cap Y^c)$$

constitute a corresponding coarse-grained observable $E^{\mathcal{R}}$ of E . The maps $f_1 : 1, 3 \mapsto \omega; 2, 4 \mapsto \omega'$, and $f_2 : 1, 2 \mapsto \xi; 3, 4 \mapsto \xi'$, allow one to write

$$\begin{aligned} A_1 = E(X) &= E((X \cap Y) \cup (X \cap Y^c)) = E(X \cap Y) + E(X \cap Y^c) = E^{\mathcal{R}}(f_1^{-1}\{\omega\}) \\ A_2 = E(Y) &= E(X \cap Y) + E(X^c \cap Y) = E^{\mathcal{R}}(f_2^{-1}\{\xi\}), \end{aligned}$$

showing that the coexistent pair E_1, E_2 is jointly measurable. $\square$

As an immediate corollary to Lemma 4.6 and Proposition 4.8 we also have the following.

Corollary 11.1 *Coexistent sharp observables are mutually commutative. If their value spaces have the property (D), then they are also jointly measurable and their joint observable is of the product form.*

In view of the probabilistic nature of quantum mechanics, a single measurement result has almost no meaning at all. It is the whole measurement outcome distribution which is to be considered as a result of a measurement. From that point of view one might suggest that a joint measurement of observables E_1, E_2 is any measurement $\mathcal{M}$, and thus observable M , whose distribution M_ϱ allows one to infer both $E_{1,\varrho}$ and $E_{2,\varrho}$ for some state or for a sufficiently large set of states or even for the full set of states ϱ . It is then possible that even for pairs of incompatible observables such measurements can be constructed. We shall confirm this in the case of position and momentum observables in Sect. 15.4 and for all pairs of observables in Chap. 18. However, in those cases one cannot, in general, obtain any information for the relevant pair probabilities and therefore we do not consider such joint determinations of the measurement outcome statistics as joint measurements of the observables in question.

11.3 Regular Observables

It has become evident in Sect. 9.2 that the order structure of the set of effects $E(\mathcal{H})$ is highly complicated. For instance, for an observable $E : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ the effect $E(X \cap Y)$ is a lower bound of the effects $E(X)$ and $E(Y)$ for any $X, Y \in \mathcal{A}$, but for

these effects, the greatest lower bound $\mathbf{E}(X) \wedge_{\mathbf{E}(\mathcal{H})} \mathbf{E}(Y)$ need not exist; and even if it did exist, it need not coincide with $\mathbf{E}(X \cap Y)$. It follows, in particular, that the range of an observable is generally not a Boolean algebra. Since the Boolean structure plays an important role in the functional calculus of observables, it is useful to investigate conditions under which this structure is present.¹

Proposition 11.3 *The range, $\text{ran}(\mathbf{E})$, of an observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is a Boolean subalgebra of the set $\mathbf{E}(\mathcal{H})$ of effects if and only if $\mathbf{E}$ is projection valued.*

Proof For any $X \in \mathcal{A}$, the product $\mathbf{E}(X)\mathbf{E}(X^c)$ is a positive lower bound of $\mathbf{E}(X)$ and $\mathbf{E}(X^c)$. If $\text{ran}(\mathbf{E})$ is Boolean, then $\mathbf{E}(X) \wedge \mathbf{E}(X^c) = 0$, and thus $\mathbf{E}(X)\mathbf{E}(X^c) = 0$, that is, $\mathbf{E}(X)^2 = \mathbf{E}(X)$. Conversely, if $\mathbf{E}$ is projection valued, the claim follows from the multiplicativity of the spectral measure and the fact that for any two projections P and R their greatest lower bound and smallest upper bound in $\mathbf{E}(\mathcal{H})$ are the same as in $\mathcal{P}(\mathcal{H})$, that is, $P \wedge R$ and $P \vee R$, respectively. $\square$

Restricting the order and the complement of $\mathbf{E}(\mathcal{H})$ to the range $\text{ran}(\mathbf{E})$ of $\mathbf{E}$, it becomes possible that the system $(\text{ran}(\mathbf{E}), \leq, ^c)$ is a Boolean σ -algebra without $\mathbf{E}$ being projection valued. To investigate this option, we introduce the notion of a regular observable.

A nontrivial effect is regular if its spectrum extends both below as well as above the value $\frac{1}{2}$. This property is important in the context of a realist interpretation of quantum mechanics as such effects and their complements can then represent approximate elements of reality, namely, in states where the probability of their occurrence, given by their expectations, is greater than $\frac{1}{2}$. The relevance of regularity here derives from that fact that regular observables are characterised by their Boolean range.

Definition 11.3 An observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is *regular* if for any $X \in \mathcal{A}$ with $0 \neq \mathbf{E}(X) \neq I$ one has

$$\mathbf{E}(X) \not\leq \frac{1}{2}I, \quad \frac{1}{2}I \not\leq \mathbf{E}(X).$$

An observable $\mathbf{E}$ is regular if and only if neither $\mathbf{E}(X) \leq \mathbf{E}(X)^c$ nor $\mathbf{E}(X)^c \leq \mathbf{E}(X)$ for any nontrivial $\mathbf{E}(X)$.

Lemma 11.1 *Let $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ be a regular observable. If $(A_i)_{i \in \mathbb{N}} \subset \text{ran}(\mathbf{E})$ is a summable sequence, that is, $A_1 + \dots + A_n \leq I$ for each $n \in \mathbb{N}$, then $\sup_{n \in \mathbb{N}} \{A_1 + \dots + A_n\} \in \text{ran}(\mathbf{E})$. Moreover, $\text{ran}(\mathbf{E})$ is a Boolean algebra with respect to the order and complement inherited from the set of effects $\mathbf{E}(\mathcal{H})$, and $\mathbf{E}$ is a Boolean σ -homomorphism $\mathcal{A} \rightarrow \text{ran}(\mathbf{E})$.*

Proof Let $A_1, A_2 \in \text{ran}(\mathbf{E})$ be such that $A_1 + A_2 \leq I$, and assume that $A_1 = \mathbf{E}(X)$, $A_2 = \mathbf{E}(Y)$. Then $X = (X \cap Y) \cup (X \setminus (X \cap Y))$ and $Y = (X \cap Y) \cup (Y \setminus (X \cap Y))$. Hence $\mathbf{E}(X \cap Y) \leq \mathbf{E}(X) = A_1$, $\mathbf{E}(X \cap Y) \leq \mathbf{E}(Y) = A_2$. Since $A_2 \leq I - A_1$, it follows that $\mathbf{E}(X \cap Y) \leq \mathbf{E}(X \cap Y)^c$, so that by the regularity assumption, $\mathbf{E}(X \cap Y) = 0$. Therefore $A_1 = \mathbf{E}(X_1)$ and $A_2 = \mathbf{E}(Y_1)$, where $X_1 = X \setminus (X \cap Y)$, $Y_1 = Y \setminus (X \cap Y)$

¹The results of this section were obtained in collaboration with S. Pulmannová [3, 4].

are disjoint sets. So we get $A_1 + A_2 = E(X_1) + E(Y_1) = E(X_1 \cup Y_1) \in \text{ran}(E)$. We note also that $X \cap Y_1 = \emptyset$ and $A_1 + A_2 = E(X \cup Y_1)$. This observation will be used presently.

Assume next that $(A_i)_{i \in \mathbb{N}}$ is a summable sequence in $\text{ran}(E)$. Using the above argument, we find disjoint sets X_1, X_2 such that $A_1 = E(X_1), A_2 = E(X_2)$. Now we proceed by induction. Assume that we have already found disjoint sets $X_1, \dots, X_{n-1}$ such that $A_i = E(X_i), i = 1, \dots, n-1$. Then $A_1 + \dots + A_{n-1} = E(X_1 \cup X_2 \cup \dots \cup X_{n-1})$. By the summability assumption, $(A_1 + \dots + A_{n-1}) \leq I - A_n$. Therefore, there is a set $X_n \in \mathcal{A}$ such that $(X_1 \cup \dots \cup X_{n-1}) \cap X_n = \emptyset$, and $A_n = E(X_n)$. Thus we find a sequence $X_i, i \in \mathbb{N}$, of disjoint sets such that $A_i = E(X_i), i \in \mathbb{N}$. From the σ -additivity of E we obtain $E(\bigcup_i X_i) = \sum_i E(X_i) = \sum_i A_i$, which shows that $\text{ran}(E)$ is closed under sums of summable sequences.

Let $E(X), E(Y) \in \text{ran}(E)$. We will prove that

$$E(X \cap Y) = E(X) \wedge_{\text{ran}(E)} E(Y),$$

that is, $E : \mathcal{A} \rightarrow \text{ran}(E)$ is a $\wedge$ -morphism. Evidently, $E(X \cap Y) \leq E(X), E(Y)$. Assume that for some $Z \in \mathcal{A}$, $E(Z) \leq E(X), E(Y)$ and write $Z = (Z \cap X \cap Y) \cup (Z \cap (X \cap Y)^c)$. Then we note:

$$\begin{aligned} E(Z \cap (X \cap Y)^c) &= E(Z \cap (X^c \cup Y^c)) \\ &= E((Z \cap X^c \cap Y) \cup (Z \cap X^c \cap Y^c) \cup (Z \cap X \cap Y^c)) \\ &= E(Z \cap X^c \cap Y) + E(Z \cap X^c \cap Y^c) + E(Z \cap X \cap Y^c) \\ &\leq E(Z) \leq E(X), E(Y). \end{aligned}$$

But we also have $E(Z \cap X^c \cap Y) \leq E(X^c), E(Z \cap X \cap Y^c) \leq E(Y^c), E(Z \cap X^c \cap Y^c) \leq E(X^c), E(Y^c)$, so that the effects $E(Z \cap X^c \cap Y), E(Z \cap X^c \cap Y^c)$, and $E(Z \cap X \cap Y^c)$ are irregular and thus equal 0. Therefore also $E(Z \cap (X \cap Y)^c) = 0$, and so $E(Z) = E(Z \cap X \cap Y) \leq E(X \cap Y)$. This concludes the proof that $E(X \cap Y) = E(X) \wedge_{\text{ran}(E)} E(Y)$. By de Morgan's law one obtains the dual result: for any $X, Y \in \mathcal{A}$, $E(X \cup Y) = E(X) \vee_{\text{ran}(E)} E(Y)$. Moreover, if the sets X and Y are disjoint, then $E(X) \vee_{\text{ran}(E)} E(Y) = E(X) + E(Y)$. Also, if $(X_i) \subset \mathcal{A}$ is a disjoint sequence, then

$$E\left(\bigcup_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} E(X_i) = \bigvee_{\text{ran}(E)} \{E(X_i) \mid i \in \mathbb{N}\}.$$

To prove that $\text{ran}(E)$ is a Boolean algebra, it remains to show the distributivity. But this follows immediately from the fact that E is a $\wedge$ -morphism and a $\vee$ -morphism from a Boolean set. $\square$

Corollary 11.2 *The range $\text{ran}(E)$ of an observable E is a Boolean algebra (with the ordering inherited from $E(\mathcal{H})$) if and only if E is regular.*

Proof We have to prove the ‘only if’ part. Hence, assume that $\text{ran}(\mathbf{E})$ is Boolean, and let $\mathbf{E}(X)$ be an irregular element. Then $\mathbf{E}(X) \leq \mathbf{E}(X)^c$, which in a Boolean algebra implies that $\mathbf{E}(X) = 0$. $\square$

Theorem 11.2 *Assume that the value spaces of observables $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{E}_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ have the property (D). If there is a regular observable $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ such that $\text{ran}(\mathbf{E}_1) \cup \text{ran}(\mathbf{E}_2) \subset \text{ran}(\mathbf{E})$, then $\mathbf{E}_1$ and $\mathbf{E}_2$ are compatible.*

Proof If $\mathbf{E}$ is regular, then from $\text{ran}(\mathbf{E}_1) \cup \text{ran}(\mathbf{E}_2) \subset \text{ran}(\mathbf{E})$ it follows that $\mathbf{E}_1$ and $\mathbf{E}_2$ are regular. Thus, by Lemma 11.1, the ranges $\text{ran}(\mathbf{E}_1)$, $\text{ran}(\mathbf{E}_2)$ and $\text{ran}(\mathbf{E})$ are Boolean. From this and the fact that $\text{ran}(\mathbf{E}_1) \cup \text{ran}(\mathbf{E}_2) \subset \text{ran}(\mathbf{E})$ it then follows that the map $(X, Y) \mapsto \mathbf{E}_1(X) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y)$ is a bioobservable of $\mathbf{E}_1$ and $\mathbf{E}_2$. Indeed, for a fixed $Y \in \mathcal{A}_2$, if $(X_i) \subset \mathcal{A}_1$ is a disjoint sequence, then

$$\begin{aligned} \mathbf{E}_1(\cup X_i) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y) &= \left(\sum \mathbf{E}_1(X_i) \right) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y) \\ &= \left(\sum \mathbf{E}(Z_{X_i}) \right) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y) \\ &= (\vee_{\text{ran}(\mathbf{E})} \mathbf{E}(Z_{X_i})) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y) \\ &= \vee_{\text{ran}(\mathbf{E})} (\mathbf{E}_1(X_i) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y)), \end{aligned}$$

where $(Z_{X_i}) \subset \mathcal{A}$ is a disjoint sequence such that $\mathbf{E}(Z_{X_i}) = \mathbf{E}_1(X_i)$ (which exists since $(\mathbf{E}_1(X_i)) \subset \text{ran}(\mathbf{E})$ is summable). Similarly, one shows that for a fixed $X \in \mathcal{A}_1$, if $(Y_i) \subset \mathcal{A}_2$ is a disjoint sequence, then

$$\mathbf{E}_1(X) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(\cup Y_i) = \vee_{\text{ran}(\mathbf{E})} (\mathbf{E}_1(X) \wedge_{\text{ran}(\mathbf{E})} \mathbf{E}_2(Y_i)).$$

Theorem 11.1 now entails that $\mathbf{E}_1$ and $\mathbf{E}_2$ are compatible. $\square$

11.4 Sharp Observables

Coexistent sharp observables commute with each other. If their value spaces have the property (D), then they are also jointly measurable and their joint observable is of the product form. We collect the results on sharp observables in the following theorem.

Theorem 11.3 *Assume that the measurable spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ have the property (D). For any two sharp observables $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{E}_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:*

- (i) $\mathbf{E}_1$ and $\mathbf{E}_2$ commute;
- (ii) $\mathbf{E}_1$ and $\mathbf{E}_2$ are coexistent;
- (iii) $\mathbf{E}_1$ and $\mathbf{E}_2$ are functions of a third observable;

- (iv) E_1 and E_2 are smearings of a third observable;
- (v) E_1 and E_2 have a unique bioobservable;
- (vi) E_1 and E_2 have a unique joint observable.

Remark 11.2 Let A and B be any two selfadjoint operators in $\mathcal{H}$. By definition, A and B commute if and only if all their spectral projections $A(X)$ and $B(Y)$, $X, Y \in \mathcal{B}(\mathbb{R})$, commute. By a well-known theorem of von Neumann [5], this is the case exactly when there is a selfadjoint operator C and real Borel functions f and g such that $A = f(C)$, $B = g(C)$, that is, $A = C \circ f^{-1}$ and $B = C \circ g^{-1}$. If A and B are bounded selfadjoint operators, then their commutativity can be equivalently defined by $AB = BA$. $\triangleleft$

The notion of compatibility of sharp observables E_1 and E_2 has a natural generalisation to partial compatibility or partial commutativity. We follow [6] to present one such formulation.

Definition 11.4 For two sharp observables E_1 and E_2 , the commutativity domain $\text{com}(E_1, E_2)$ is the set of vectors $\varphi \in \mathcal{H}$ for which

$$E_1(X)E_2(Y)\varphi = E_2(Y)E_1(X)\varphi \quad (11.3)$$

for all $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$. We say that E_1 and E_2 are commutative if $\text{com}(E_1, E_2) = \mathcal{H}$ and totally noncommutative if $\text{com}(E_1, E_2) = \{0\}$.

Lemma 11.2 The commutativity domain $\text{com}(E_1, E_2)$ of two sharp observables E_1 and E_2 is a closed subspace of $\mathcal{H}$ and it reduces E_1 and E_2 , that is, for any $X \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$,

$$E_1(X)(\text{com}(E_1, E_2)) \subset \text{com}(E_1, E_2), \quad E_2(Y)(\text{com}(E_1, E_2)) \subset \text{com}(E_1, E_2).$$

Proof The first claim follows since $\text{com}(E_1, E_2)$ can be expressed as the intersection of closed subspaces,

$$\text{com}(E_1, E_2) = \bigcap_{X,Y} \{\varphi \in \mathcal{H} \mid (E_1(X)E_2(Y) - E_2(Y)E_1(X))\varphi = 0\}.$$

Let $\varphi \in \text{com}(E_1, E_2)$. Then for any $X \in \mathcal{A}_1$, $E_1(X)\varphi \in \text{com}(E_1, E_2)$, since for all $Z \in \mathcal{A}_1$, $Y \in \mathcal{A}_2$,

$$\begin{aligned} E_2(Y)E_1(Z)E_1(X)\varphi &= E_2(Y)E_1(Z \cap X)\varphi = E_1(Z \cap X)E_2(Y)\varphi \\ &= E_1(Z)E_1(X)E_2(Y)\varphi = E_1(Z)E_2(Y)E_1(X)\varphi. \end{aligned}$$

Similarly one shows $E_2(Y)(\text{com}(E_1, E_2)) \subset \text{com}(E_1, E_2)$ for each $Y \in \mathcal{A}_2$. $\square$

Theorem 11.4 Consider two sharp observables E_1 and E_2 with the respective value spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ having the property (D). For any unit vector $\varphi \in \mathcal{H}$, the following conditions are equivalent:

- (i) $\varphi \in \text{com}(\mathbf{E}_1, \mathbf{E}_2)$;
- (ii) *there is a probability measure $\mu : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow [0, 1]$ such that*

$$\mu(X \times Y) = \langle \varphi | \mathbf{E}_1(X) \mathbf{E}_2(Y) \varphi \rangle = \langle \varphi | \mathbf{E}_1(X) \wedge \mathbf{E}_2(Y) \varphi \rangle$$

for all $X \in \mathcal{A}_1, Y \in \mathcal{A}_2$.

Proof The restrictions $\tilde{\mathbf{E}}_1$ and $\tilde{\mathbf{E}}_2$ of $\mathbf{E}_1$ and $\mathbf{E}_2$ on $\text{com}(\mathbf{E}_1, \mathbf{E}_2)$ are mutually commuting spectral measures, so that, by Theorem 11.3, the map $X \times Y \mapsto \tilde{\mathbf{E}}_1(X) \tilde{\mathbf{E}}_2(Y) = \tilde{\mathbf{E}}_1(X) \wedge \tilde{\mathbf{E}}_2(Y)$ extends to a joint projection valued observable $\tilde{\mathbf{F}} : \mathcal{A}_1 \otimes \mathcal{A}_2 \rightarrow \mathcal{L}(\text{com}(\mathbf{E}_1, \mathbf{E}_2))$. But then, for any $\varphi \in \text{com}(\mathbf{E}_1, \mathbf{E}_2)$, and $X \in \mathcal{A}_1, Y \in \mathcal{A}_2$, $\tilde{\mathbf{F}}_{\varphi, \varphi}(X \times Y) = \langle \varphi | \tilde{\mathbf{E}}_1(X) \tilde{\mathbf{E}}_2(Y) \varphi \rangle = \langle \varphi | \mathbf{E}_1(X) \mathbf{E}_2(Y) \varphi \rangle$. $\square$

11.5 Compatibility, Convexity, and Coarse-Graining

The set of observables on a given value space is convex. This leads to the following result:

Proposition 11.4 *Let $\mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{F}_1, \mathbf{F}_2$ be two pairs of jointly measurable observables such that $\mathbf{E}_i, \mathbf{F}_i$ share the same value space $(\Omega_i, \mathcal{A}_i)$. Let $\mathbf{G}, \mathbf{H}$ be joint observables of these pairs, respectively. Then for any $\lambda \in [0, 1]$, the observables $\lambda \mathbf{E}_1 + (1 - \lambda) \mathbf{F}_1$ and $\lambda \mathbf{E}_2 + (1 - \lambda) \mathbf{F}_2$ are jointly measurable, with a joint observable given by $\lambda \mathbf{G} + (1 - \lambda) \mathbf{H}$.*

The proof is straightforward and is left as an exercise.

The following result (due to [7] for the finite dimensional case) shows that the question of the compatibility of a pair of effects can always be decided by means of a single inequality.

Proposition 11.5 *Effects E, F in a separable Hilbert space are compatible if and only if the following inequality holds:*

$$\lambda_0 = \inf\{\lambda \in \mathbb{R} : E + F \leq \lambda I + G \text{ for some effect } G \leq E, F\} \leq 1. \quad (11.4)$$

Proof The fact that λ_0 is the minimum of the above set follows from a standard topological argument [8]. (In the finite dimensional case it results from a presentation of the problem of finding the infimum as a strictly feasible semidefinite programme.) If E, F are compatible, then there exists an effect G such that $G \leq E, F$ and $E + F \leq I + G$. This entails $\lambda_0 \leq 1$. Conversely, if $\lambda_0 \leq 1$, then for $\lambda = 1$ there exists an effect G satisfying $G \leq E, F$ and $E + F \leq \lambda I + G = I + G$. $\square$

Remark 11.3 The parameter λ_0 can be interpreted as follows. The quantity $\mu = \max\{0, 1 - \lambda_0^{-1}\}$ is the least amount of “noise” that has to be added to E, F (in the form of mixing E, F with the trivial effect 0 so that $\tilde{E} = (1 - \mu)E + \mu 0$, $\tilde{F} = (1 - \mu)F + \mu 0$ are compatible. Thus, a value $\lambda_0 > 1$ gives an indication of

how far away an incompatible pair of effects E, F are from a suitable “nearby” pair of compatible effects. $\triangleleft$

While the compatibility characterisation of Proposition 11.5 is operational (in the sense just described) and efficiently decidable via a semidefinite programme, it does nothing to elucidate the relationship—presumably of a trade-off—between the degree of noncommutativity of a pair of compatible effects E, F and their degrees of unsharpness without which they could not be compatible. This trade-off will be investigated for pairs of specific observables in later chapters, and it will be found, for example, in the case of qubit observables, that in addition to unsharpness and the degree of noncommutativity, other features are relevant for the compatibility question.

To conclude this chapter, we present a simple scheme of turning a set of effects or observables into a jointly measurable set. First we observe that in line with Remark 11.3, a set of effects $\{E_1, \dots, E_n\}$ can be turned into a compatible set simply by “scaling down” each effect with a suitable factor $\kappa \in (0, 1)$. A trivial way of ensuring the joint measurability of $\{\kappa E_1, \dots, \kappa E_n\}$ is by choosing κ small enough such that $\kappa E_1 + \kappa E_2 + \dots + \kappa E_n \leq I$. However, it is not necessary for compatibility to force the sum to be bounded by the identity operator; in fact, for two effects E, F and the choice $\kappa = 1 - \mu = \lambda_0^{-1}$, one will not have, in general, $\kappa E + \kappa F \leq I$ (see Exercise 13 of Chap. 14).

Proposition 11.6 *Let E and F be two observables on a finite outcome set, and let $0 \leq \mu \leq 1$. Then $\mu E + (1 - \mu)T$ and $(1 - \mu)F + \mu T'$ are compatible for any choice of trivial observables T and T' on the same outcome set.*

Proof First, let p and p' denote the probability distributions associated with T and T' . We define an observable G by the formula

$$G_\varrho(j, k) = \mu E_\varrho(j) p'(k) + (1 - \mu) p(j) F_\varrho(k). \quad (11.5)$$

For a fixed ϱ , the right hand side is clearly a probability distribution. Moreover, $\varrho \mapsto G_\varrho$ defines an affine mapping on $S(\mathcal{H})$; therefore G is an observable. The marginal observables are

$$\begin{aligned} \sum_k G_\varrho(j, k) &= \mu E_\varrho(j) + (1 - \mu) p(j), \\ \sum_j G_\varrho(j, k) &= (1 - \mu) F_\varrho(k) + \mu p'(k). \end{aligned} \quad \square$$

The physical idea behind this construction is the following. For each run of the measurement a coin is flipped and, depending on the result, one measures either E or F in the input state ϱ . In this way one obtains a measurement outcome for either E or F . In addition to this, one makes a random selection (according to either the distribution p' or p) of an outcome for the other observable. In this way one obtains an outcome for both observables $E^{(\mu)} = \mu E + (1 - \mu)T$ and $F^{(1-\mu)} = (1 - \mu)F + \mu T'$ simultaneously. The overall joint observable is the one given in formula (11.5).

Remark 11.4 The method of turning incompatible pairs of observables into compatible ones by adding random noise can be used to define a measure of the *degree of incompatibility*. We illustrate this for binary observables or simply for their generating effects E, F . The degree of incompatibility of E, F , denoted $\text{inc}(E, F)$ is $1 - \lambda^*$, where λ^* is the supremum of the numbers $\lambda \in [0, 1]$ such that

$$\begin{aligned} E^{(\lambda)} &= \frac{1+\lambda}{2}E + \frac{1-\lambda}{2}(I - E) = \lambda E + (1-\lambda)\frac{1}{2}I \\ F^{(\lambda)} &= \frac{1+\lambda}{2}F + \frac{1-\lambda}{2}(I - F) = \lambda F + (1-\lambda)\frac{1}{2}I \end{aligned} \quad (11.6)$$

are compatible. It can be shown that λ^* is in fact the maximum of this set of λ (Exercise). Then E, F are compatible if and only if $\text{inc}(E, F) = 0$.

Using the joint measurability condition for a pair of unbiased qubit effects reported here as Proposition 14.1, it has been shown by Banik et al. [9] that any pair of effects E, F in a finite dimensional Hilbert space is turned into a compatible pair $E^{(\lambda)}, F^{(\lambda)}$ for some $\lambda \geq 1/\sqrt{2}$. This limiting value is necessary: if, say E, F are spectral projections of the Pauli operators σ_1, σ_2 , then $E^{(\lambda)}, F^{(\lambda)}$ are compatible if and only if $\lambda = 1/\sqrt{2}$.

Refinements of the above definition can be used to compare the degree of incompatibility inherent in quantum mechanics with that of a set of more general probabilistic theories [9–11]. $\triangleleft$

Remark 11.5 It is important to note that a joint measurement of a set of appropriately rescaled effects $\{\kappa E_i\}$ does not constitute a joint measurement of the original set of effects $\{E_i\}$, despite the fact that the probabilities of the E_i can be computed once one has determined those of the κE_i . These sets of effects represent physically different collections of experimental events. The latter set can occur as possible outcomes within a common measurement, the former cannot if the E_i are incompatible. $\triangleleft$

11.6 Exercises

1. A collection of effects is called compatible if they occur in the range of a common observable. Show that effects E, F are compatible if and only if there are effects G, H , such that $G \leq E \leq H, G \leq F \leq H, G + H = E + F$, or, equivalently, if and only if there is an effect G such that $E + F - I \leq G \leq E, F$.
2. Prove Proposition 11.4.
3. Prove that two effects in a Hilbert space $\mathcal{H}$ are compatible if and only if they are compatible as effects in the closed subspace spanned by the union of their ranges.
4. Let $\mathcal{H} = \mathbb{C}^3$ and let $\{|1\rangle, |2\rangle, |3\rangle\}$ be its orthonormal basis. Define orthonormal unit vectors $\psi_1 = (|1\rangle + |2\rangle + |3\rangle)/\sqrt{3}, \psi_2 = (|1\rangle + \alpha|2\rangle + \alpha^2|3\rangle)/\sqrt{3}$ and $\psi_3 = (|1\rangle + \alpha^2|2\rangle + \alpha|3\rangle)/\sqrt{3}$ where $\alpha = \exp(2\pi i/3)$. Show that the set of effects $\{A_1, \dots, A_6\} = \{\frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|2\rangle\langle 2|, \frac{1}{2}|3\rangle\langle 3|, \frac{1}{2}|\psi_1\rangle\langle \psi_1|, \frac{1}{2}|\psi_2\rangle\langle \psi_2|,$

$\frac{1}{2}|\psi_3\rangle\langle\psi_3|$ constitutes a 6-outcome (rank-1) observable **A**. Define a (rank-2) observable $\{B_1, B_2, B_3\} = \{\frac{1}{2}|2\rangle\langle 2| + \frac{1}{2}|3\rangle\langle 3|, \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|3\rangle\langle 3|, \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|2\rangle\langle 2|\}$. Since the ranges of **A** and **B** belong to the range of **A** the observables **A** and **B** are coexistent. Show that **B** is not a smearing of **A**. Hence, **A** and **B** are not jointly measurable. This example is from [1].

5. Let E, F be two effects. Show that the set

$$\left\{ \lambda \mid \lambda E + (1 - \lambda) \frac{1}{2} I \text{ and } \lambda F + (1 - \lambda) \frac{1}{2} I \text{ are compatible} \right\}$$

is closed.

References

1. Pellonpää, J.-P.: On coexistence and joint measurability of rank-1 quantum observables. *J. Phys. A* **47**(5), 052002, 12 (2014)
2. Reeb, D., Reitzner, D., Wolf, M.M.: Coexistence does not imply joint measurability. *J. Phys. A* **46**, 462002 (2013)
3. Lahti, P., Pulmannová, S.: Coexistent observables and effects in quantum mechanics. *Rep. Math. Phys.* **39**(3), 339–351 (1997)
4. Lahti, P., Pulmannová, S.: Coexistence versus functional coexistence of quantum observables. *Rep. Math. Phys.* **47**(2), 199–212 (2001)
5. J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Die Grundlehren der mathematischen Wissenschaften, Band 38. Springer-Verlag, Berlin, (1968, 1996). (Reprint of the 1932 original). English translation: *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton (1955, 1996)
6. Ylinen, K.: On a theorem of Gudder on joint distributions of observables. In: *Symposium on the Foundations of Modern Physics* (Joensuu, 1985), pp. 691–694. World Scientific Publishing, Singapore (1985)
7. Wolf, M.M., Perez-Garcia, D., Fernandez, C.: Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Phys. Rev. Lett.* **103**, 230402 (2009)
8. Beneduci, R., Busch, P.: Incompatibility of effects and Bell inequality violations in general probabilistic models. Unpublished note (2012)
9. Banik, M., Gazi, M.R., Ghosh, S., Kar, G.: Degree of complementarity determines the nonlocality in quantum mechanics. *Phys. Rev. A* **87**(5), 052125 (2013)
10. Busch, P., Heinosaari, T., Schultz, J., Stevens, N.: Comparing the degrees of incompatibility inherent in probabilistic physical theories. *EPL (Europhys. Lett.)* **103**(1), 10002 (2013)
11. Stevens, N., Busch, P.: Steering, incompatibility, and Bell-inequality violations in a class of probabilistic theories. *Phys. Rev. A* **89**(2), 022123 (2014)

Chapter 12

Preparation Uncertainty

The probabilistic structure of quantum mechanics is a reflection of the fact that observations on quantum physical objects typically yield uncertain outcomes. Formally this uncertainty is encoded in the probability distribution of an observable in the state of the physical system. It is a fundamental feature of quantum mechanics that there are pairs of observables for which the degrees of uncertainty in the associated probability distributions cannot both be arbitrarily small in the same states. This feature constitutes one aspect of the *uncertainty principle* and is expressed in the form of trade-off inequalities called *preparation uncertainty relations*.

Another aspect of the uncertainty principle manifests itself as a limitation of the accuracy with which incompatible observables can be jointly measured. Such limitations are expressed in the form of trade-off relations for measurement errors and are called *measurement uncertainty relations* or *error-disturbance relations*. In subsequent chapters we will study a number of examples illustrating that measurement uncertainty is a necessary consequence of preparation uncertainty in quantum mechanics.

The statement of preparation and measurement uncertainty relations requires the definition of appropriate measures of uncertainty, measurement error and disturbance. The present chapter presents measures of uncertainty based on the probability distributions of observables and gives examples of preparation uncertainty relations that are found to be fundamental for associated measurement uncertainty relations.

We will first make precise what it means for the value of an observable to be definite or indeterminate in a given state, and then proceed to briefly review uncertainty relations based on uncertainty measures such as standard deviation and overall width.

12.1 Indeterminate Values of Observables

All measures of uncertainty considered here are adapted to observables with value spaces in $\mathbb{R}$. Moreover, we consider only observables E whose first moment operators $E[1] = \int_{\mathbb{R}} x dE(x)$ are selfadjoint. We recall that such an observable E is sharp exactly when its second moment operator is the square of its first moment operator, that is, the intrinsic noise operator $N(E) = E[2] - E[1]^2 = 0$ (Theorem 8.5). According to Corollary 9.1 the sharp observable A is characterised as the one with the least variance among all observables E for which $E[1] = A$.

The *values* of an observable E as the possible measurement outcomes are exactly the elements of the support $\text{supp}(E)$ of E . Thus a real number $x \in \mathbb{R}$ is a possible value of E if and only if for any $\epsilon > 0$ there is a state ϱ such that $E_{\varrho}((x - \epsilon, x + \epsilon)) \neq 0$. If E is sharp then $\text{supp}(E) = \sigma(A)$, the spectrum of $A = E[1]$; in this case the condition that x is a possible value may be expressed as the statement that for any $\epsilon > 0$ there exists a state ϱ with $E_{\varrho}((x - \epsilon, x + \epsilon)) = 1$.

We say that E is *definite* or has a (*definite*) *value* if one of its values is obtained with (probabilistic) certainty if measured. An observable E thus has a value $x \in \text{supp}(E)$ in a state ϱ exactly when E_{ϱ} is the point measure δ_x . This may occur if and only if the effect $E(\{x\})$ is nonzero and has eigenvalue 1. If E is sharp, then $E = A$ has a definite value in a vector state φ if and only if φ is an eigenstate of A , that is, $A\varphi = x\varphi$. If φ is not an eigenstate, then several different eigenvalues may occur with nonzero probability in a measurement of A . The values of the observable A are then said to be *indeterminate*.¹

Quantum uncertainty thus manifests itself in the randomness of measurement outcomes. The degree to which an observable is indeterminate can be quantified in terms of the width of its associated probability distribution in the given state.

12.2 Measures of Uncertainty

Standard Deviation

The most familiar measure of the width of a probability measure $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ is the standard deviation, given by the square root of the variance of μ ,

¹We will not dwell here on the distinction between the objective indeterminateness and subjective uncertainty of the values of an observable, which is explained in detail in [1]. Suffice it to say that in pure states the values are objectively indeterminate if their probabilities are not 1 or 0. A situation of subjective uncertainty (or ignorance) can be modelled by a state that is a mixture of eigenstates associated with at least two distinct eigenvalues, but even then one has to recognise that any mixed state can be decomposed in uncountably many ways into its pure components (see Theorem 9.2). A mixed state can arise as a reduced state of an entangled pure state, in which case application of the ignorance interpretation with respect to any particular decomposition is inconsistent with the indeterminacy represented by that entangled state.

$$\Delta(\mu)^2 = \text{Var}(\mu) = \int \left(x - \int x' d\mu(x') \right)^2 d\mu(x) = \mu[2] - \mu[1]^2,$$

provided that the integrals exist and are finite; otherwise we write $\Delta(\mu) = \infty$. One has $\Delta(\mu) = 0$ exactly when μ is the point measure at $\mu[1]$.

We will encounter many situations where a given probability distribution is randomised further (or smeared) by convolution with another probability measure. For probability measures of convolution form, $\mu * \nu$, Lemma 8.2 gives

$$\Delta(\mu), \Delta(\nu) \leq \Delta(\mu * \nu) = \sqrt{\Delta(\mu)^2 + \Delta(\nu)^2} \leq \Delta(\mu) + \Delta(\nu),$$

provided the relevant integrals defining all expressions are finite.

By the standard deviation of an observable $E : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ in a state ϱ we mean the standard deviation of the probability measure E_ϱ , which again is well-defined provided the first and second moments, $E_\varrho[1]$ and $E_\varrho[2]$, exist and are finite. If $\varrho = |\varphi\rangle\langle\varphi|$ is a pure state then these conditions simply mean that φ belongs in the domain of the operator $E[2]$; for a mixed state these conditions are more involved (see Sect. 9.3). In the following we do not always mention these obvious requirements. For a vector state φ we write $\Delta(E_\varphi)$ or $\Delta(E, \varphi)$. If E is sharp, we will also write $\Delta(A, \varrho)$ or $\Delta(A, \varphi)$, respectively, where $A = E[1]$.

We have $\Delta(E_\varrho) = 0$ exactly when the probability measure E_ϱ is a point measure at $E_\varrho[1]$, that is, E has the value $E_\varrho[1]$ in the state ϱ . We recall also that if E is sharp then for any $\epsilon > 0$ there is a state ϱ such that $\Delta(E_\varrho) < \epsilon$, that is, $\inf\{\Delta(E_\varrho) \mid \varrho \in \mathcal{S}(\mathcal{H})\} = 0$, and the lower bound is reached if and only if $A = E[1]$ has eigenvalues. More specifically, for any point x in the spectrum of A , there are sequences of vector states φ_n such that $\langle \varphi_n | A \varphi_n \rangle = x$ and $\Delta(A, \varphi_n) \rightarrow 0$ as $n \rightarrow \infty$.

Using the intrinsic noise operator $N(E) = E[2] - E[1]^2$, the variance $\text{Var}(E_\varrho)$ can be written as $\text{Var}(E_\varrho) = \text{Var}(E[1], \varrho) + \text{tr}[\varrho N(E)]$, see Eq. (9.15). We emphasise once more that this equation presents a splitting of the variance of E_ϱ into two terms that are not accessible through the measurement of E only.

For the special class of observables E that are convolutions of a spectral measure A and a probability measure μ (with finite variance), one has $\text{Var}(E_\varrho) = \text{Var}(\mu) + \text{Var}(A_\varrho)$. In this case the intrinsic noise contribution to the variance of the distribution E_ϱ is constant (state independent) and equal to $\text{Var}(\mu)$.

α -spread

The standard deviation of a probability measure is a special case of the so-called α -spreads. Let d be the Euclidean metric on $\mathbb{R}$, $d(x, y) = |x - y|$. For $1 \leq \alpha < \infty$, the *deviation of order α* , or *α -deviation*, of a probability measure μ from a point $y \in \mathbb{R}$ —or equivalently, from the point measure δ_y at y —is defined as²

²The case $\alpha = \infty$ can also be included in what follows, but as it would require separate considerations at various points, we omit this possibility and refer the reader to [2] for more detail.

$$\Delta_\alpha(\mu, \delta_y) = \left(\int d(x, y)^\alpha d\mu(x) \right)^{\frac{1}{\alpha}}. \quad (12.1)$$

The α -spread, or minimal deviation of order α , of μ is then defined as

$$\Delta_\alpha(\mu) = \inf_{y \in \mathbb{R}} \Delta_\alpha(\mu, \delta_y) = \inf_y \left(\int |x - y|^\alpha d\mu(x) \right)^{\frac{1}{\alpha}}. \quad (12.2)$$

For $\alpha = 2$ we may use both notations $\Delta_2(\mu)$ and $\Delta(\mu)$. When (12.1) is interpreted as distance, (12.2) represents the smallest distance of μ to the set of point measures. The point y to which a given measure is “closest” depends on α . For the absolute deviation ($\alpha = 1$) this is the median, for the standard deviation ($\alpha = 2$) it is the mean value. Like the standard deviation, the α -spread scales with the underlying metric: for the measure $\mu^{(\lambda)}$, defined via $\mu^{(\lambda)}(X) = \mu(\lambda X)$, $X \in \mathcal{B}(\mathbb{R})$, and fixed $\lambda > 0$, one has

$$\Delta_\alpha(\mu^{(\lambda)}) = \frac{1}{\lambda} \Delta_\alpha(\mu). \quad (12.3)$$

Lemma 12.1 *For any α and for any two probability measures μ, ν*

$$\Delta_\alpha(\mu), \Delta_\alpha(\nu) \leq \Delta_\alpha(\mu * \nu) \leq \Delta_\alpha(\mu) + \Delta_\alpha(\nu). \quad (12.4)$$

Proof For $\alpha = 2$ the result was already noted. For a general α we follow [2] to indicate a proof. Note that $\Delta_\alpha(\mu, \delta_y)$ is the standard α -norm $\|\cdot\|_{\mu, \alpha}$ in $L^\alpha(\mathbb{R}, \mu)$ of the function $x \mapsto x - y = f_y(x)$, that is

$$\Delta_\alpha(\mu, \delta_y) = \|f_y\|_{\mu, \alpha}. \quad (12.5)$$

For the first inequality in (12.4) we use translation invariance and concavity of Δ_α by considering $\nu * \mu$ as a convex combination of translates of μ with weight ν . For the second inequality in (12.4), using Lemma 8.2 (with its notation) and the Minkowski inequality we get

$$\begin{aligned} \Delta_\alpha(\mu * \nu, \delta_{x'+y'}) &= \|f_{x'+y'}\|_{\nu * \mu, \alpha} = \|f_{x'+y'} \circ \phi\|_{\nu \times \mu, \alpha} \\ &= \left(\int |x + y - (x' + y')|^\alpha d(\nu \times \mu)(x, y) \right)^{1/\alpha} \\ &\leq \left(\int |x - x'|^\alpha d(\nu \times \mu)(x, y) \right)^{1/\alpha} + \left(\int |y - y'|^\alpha d(\nu \times \mu)(x, y) \right)^{1/\alpha} \\ &= \|f_{x'}\|_{\nu, \alpha} + \|f_{y'}\|_{\mu, \alpha} = \Delta_\alpha(\mu, \delta_{x'}) + \Delta_\alpha(\nu, \delta_{y'}). \end{aligned}$$

The desired inequality (12.4) now follows by minimising over x' and y' . $\square$

For an observable $F = \mu * E$ one thus has for any state ϱ ,

$$\Delta_\alpha(E_\varrho) \leq \Delta_\alpha(\mu * E_\varrho) \leq \Delta_\alpha(\mu) + \Delta_\alpha(E_\varrho).$$

Overall Width

Let $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be a probability measure. For any $\epsilon \in (0, 1)$ we consider the set $\mathbf{D}$ of all intervals $X \subset \mathbb{R}$ for which $\mu(X) \geq 1 - \epsilon$. Since $[-n, n] \in \mathbf{D}$ for $n \in \mathbb{N}$ large enough, the infimum of the lengths $|X|$ of the intervals $X \in \mathbf{D}$ is defined as a real number. This number,

$$W_\epsilon(\mu) = \inf_{X \in \mathbf{D}} \{|X| \mid \mu(X) \geq 1 - \epsilon\}, \quad (12.6)$$

is the *overall width* of the probability measure μ . As with the α -deviation, the overall width scales with the underlying metric of the value space.

Clearly the infimum does not change if we confine our attention to compact (closed and bounded) intervals. We show that this infimum is actually a minimum.

Lemma 12.2 *There is a compact interval $X \subset \mathbb{R}$ such that $\mu(X) \geq 1 - \epsilon$ and $W_\epsilon(\mu) = |X|$.*

Proof There are sequences of real numbers $a_n \leq b_n$ such that $\mu([a_n, b_n]) \geq 1 - \epsilon$ and $b_n - a_n \rightarrow W_\epsilon(\mu)$. We may assume that $b_n - a_n \leq W_\epsilon(\mu) + 1$. The sequences (a_n) and (b_n) are bounded, for otherwise we would get a sequence of intervals $[a_n, b_n]$ eventually disjoint from any fixed $[-p, p]$, contradicting the requirement $\mu([a_n, b_n]) \geq 1 - \epsilon$. Passing to suitable subsequences we may therefore assume that the limits $a = \lim_n a_n$ and $b = \lim_n b_n$ with $a \leq b$ exist. Clearly $b - a = W_\epsilon(\mu)$. We claim that $\mu([a, b]) \geq 1 - \epsilon$. Again passing to subsequences if necessary it is easily shown that we may assume that both sequences (a_n) , (b_n) are monotone. If (a_n) is increasing and (b_n) is decreasing, then $\mu([a, b]) = \lim_n \mu([a_n, b_n]) \geq 1 - \epsilon$. If (a_n) is decreasing and (b_n) is increasing, then $\mu([a, b]) \geq \mu([a_1, b_1]) \geq 1 - \epsilon$. The two remaining cases are essentially similar, and we only consider the case where both sequences are increasing. Then $\mu([a_n, b]) \geq \mu([a_n, b_n]) \geq 1 - \epsilon$, and so $\mu([a, b]) = \lim_n \mu([a_n, b]) \geq 1 - \epsilon$. $\square$

If $W_\epsilon(\mu) = 0$, then $\mu(\{x\}) = \lambda \geq 1 - \epsilon$ for some $x \in \mathbb{R}$, meaning that μ has a discrete part. For a continuous observable $\mathbf{E}$, like position or momentum, one thus has $W_\epsilon(\mathbf{E}_\varrho) \neq 0$ for any ϵ and ϱ .

For the overall width of a convolution of measures μ, ν it is straightforward to verify the following bound:

$$W_\epsilon(\mu * \nu) \geq \max\{W_\epsilon(\mu), W_\epsilon(\nu)\}. \quad (12.7)$$

There is a simple connection between the overall width and the α -deviation of a probability measure μ on $\mathbb{R}$, which arises as an expression of Chebyshev's inequality, given here in the form

$$\mu(\{x \in \mathbb{R} \mid |x - y| \geq \delta\}) \leq \frac{\int_{\mathbb{R}} |x - y|^\alpha d\mu(x)}{\delta^\alpha} \quad (\alpha \geq 1).$$

This translates readily into

$$\mu(\{x \in \mathbb{R} \mid |x - y| < \delta\}) \geq 1 - \frac{\Delta_\alpha(\mu, \delta_y)^\alpha}{\delta^\alpha} \equiv 1 - \varepsilon,$$

and (since $\delta = \Delta_\alpha(\mu, \delta_y)/(\varepsilon^{1/\alpha})$) this is equivalent to

$$W_\varepsilon(\mu) \leq 2 \frac{\Delta_\alpha(\mu, \delta_y)}{\varepsilon^{1/\alpha}}. \quad (12.8)$$

Consequently, then one also has

$$W_\varepsilon(\mu) \leq 2 \frac{\Delta_\alpha(\mu)}{\varepsilon^{1/\alpha}}. \quad (12.9)$$

By Lemma 12.1 this also gives an upper bound for $W_\varepsilon(\mu * \nu)$ in (12.7).

12.3 Examples of Preparation Uncertainty Relations

Preparation uncertainty relations express the mutual dependence of the widths of the measurement outcome distributions of two or more observables and they are usually given as inequalities limiting from below either the product or the sum of the widths of the distributions. Here we consider some of the most typical examples of such relations.

Uncertainty Relations for Products of Standard Deviations

Consider now any two real observables E and F , with $A = E[1]$ and $B = F[1]$. The standard uncertainty relation follows by application of the Cauchy–Schwarz inequality to $\Delta(A, \varphi)\Delta(B, \varphi)$:

$$\Delta(E_\varphi) \Delta(F_\varphi) \geq \Delta(A, \varphi) \Delta(B, \varphi) \geq \frac{1}{2} |\langle \varphi | (AB - BA) \varphi \rangle|. \quad (12.10)$$

It is well known that a more refined analysis yields a stronger form, which also involves a covariance term:

$$\begin{aligned} \Delta(E_\varphi)^2 \Delta(F_\varphi)^2 &\geq \Delta(A, \varphi)^2 \Delta(B, \varphi)^2 \\ &\geq \frac{1}{4} |\langle \varphi | (AB - BA) \varphi \rangle|^2 + \frac{1}{4} (\langle \varphi | (AB + BA) \varphi \rangle - 2 \langle \varphi | A \varphi \rangle \langle \varphi | B \varphi \rangle)^2. \end{aligned} \quad (12.11)$$

The above uncertainty relations can be generalised to arbitrary states ϱ (where due care has to be applied in the specification of the domain of states for which the expectation values are well defined) [3].

There is a sharpening of the inequality (12.10). Let E and F be two real observables such that their first moment operators $E[1] = A$ and $F[1] = B$ are selfadjoint, and let A, B be their spectral measures. Assume that E, F are jointly measurable and let G be an observable on $(\Omega, \mathcal{A})$ with functions $f_i : \Omega \rightarrow \mathbb{R}, i = 1, 2$, such that $E = G^{f_1}, F = G^{f_2}$, see Theorem 11.1. Let M be a (measurement type) dilation of G into a spectral measure acting on $\mathcal{H} \otimes \mathcal{K}$ such that

$$\begin{aligned} \langle \varphi \otimes \phi | M(f_1^{-1}(X)) \varphi \otimes \phi \rangle &= \langle \varphi | E(X) \varphi \rangle \\ \langle \varphi \otimes \phi | M(f_2^{-1}(Y)) \varphi \otimes \phi \rangle &= \langle \varphi | F(Y) \varphi \rangle \end{aligned}$$

for all $\varphi \in \mathcal{H}$. By the multiplicativity of M one then has

$$\int f_1(\omega) f_2(\omega) M_{\varphi \otimes \phi}(d\omega) = \langle L(f_1, M) \varphi \otimes \phi | L(f_2, M) \varphi \otimes \phi \rangle$$

with the mutually commuting selfadjoint operators $\hat{A} = L(f_1, M)$ and $\hat{B} = L(f_2, M)$ being such that

$$I \otimes P[\phi] \hat{A} I \otimes P[\phi] = A \otimes P[\phi] \quad \text{and} \quad I \otimes P[\phi] \hat{B} I \otimes P[\phi] = B \otimes P[\phi],$$

see Theorem 7.10. A direct computation gives:

$$\begin{aligned} \int f_1(\omega) f_2(\omega) M_{\varphi \otimes \phi}(d\omega) &= \langle A \varphi | B \varphi \rangle + \langle (\hat{A} - A \otimes I) \varphi \otimes \phi | (\hat{B} - B \otimes I) \varphi \otimes \phi \rangle \\ &= \langle B \varphi | A \varphi \rangle + \langle (\hat{B} - B \otimes I) \varphi \otimes \phi | (\hat{A} - A \otimes I) \varphi \otimes \phi \rangle. \end{aligned}$$

By the Cauchy–Schwarz inequality one finally obtains

$$|\langle A \varphi | B \varphi \rangle - \langle B \varphi | A \varphi \rangle| \leq 2 \|(\hat{A} - A \otimes I) \varphi \otimes \phi\| \|(\hat{B} - B \otimes I) \varphi \otimes \phi\|.$$

The squared norms are the expectations of the noise operators $N(E)$ and $N(F)$ in the vector state φ . Hence one has the following proposition (see, e.g. [4, Theorems 2 and 3]).

Proposition 12.1 *Let E, F be any two real observables, and assume their first moment operators A, B are selfadjoint. If E, F are jointly measurable, then*

$$\langle \varphi | N(E) \varphi \rangle \langle \varphi | N(F) \varphi \rangle \geq \frac{1}{4} |\langle A \varphi | B \varphi \rangle - \langle B \varphi | A \varphi \rangle|^2. \quad (12.12)$$

An immediate consequence is the following.

Corollary 12.1 *Let E, F be any two real observables, and assume their first moment operators A, B are selfadjoint. If E, F are jointly measurable, then*

$$\Delta(E_\varphi)\Delta(F_\varphi) \geq \frac{1}{2} |\langle A\varphi | B\varphi \rangle - \langle B\varphi | A\varphi \rangle|. \quad (12.13)$$

Proof This is a consequence of (9.16). $\square$

It is a fundamental feature of quantum mechanics that there are pairs of sharp observables A and B such that their uncertainty product $\Delta(A_\varrho)\Delta(B_\varrho)$ has a state-independent, strictly positive lower bound. The best known example is given by canonically conjugate pairs such as position Q and momentum P for which one has

$$\Delta(Q_\varrho)\Delta(P_\varrho) \geq \frac{\hbar}{2}. \quad (12.14)$$

This inequality is generalised in Proposition 15.1 into a trade-off inequality for all α -spreads.

The following proposition gives further structural insight into uncertainty relations with constant bound for the product of 2-spreads.

Proposition 12.2 *Let A and B be any two sharp observables for which there is a positive constant c such that*

$$\inf\{\Delta(A_\varrho)\Delta(B_\varrho) \mid \varrho \in \mathcal{S}(\mathcal{H})\} \geq c > 0.$$

Then

- (a) A and B are totally noncommutative, that is, $\text{com}(A, B) = \{0\}$;
- (b) $A = A[1]$ and $B = B[1]$ are unbounded;
- (c) No eigenvector of A (if any) is in the domain of B (and vice versa).

Proof Assume that the Hilbert space $\mathcal{K} = \text{com}(A, B)$ is not trivial (i.e. $\mathcal{K} \neq \{0\}$) and let P be the projection from $\mathcal{H}$ onto $\mathcal{K}$. Then the truncated spectral measures $A_P : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{K})$, $X \mapsto A_P(X) = PA(X)P$ and $B_P = PBP$ commute so that $\inf\{\Delta(A_P, \varrho)\Delta(B_P, \varrho) \mid \varrho \in \mathcal{S}(\mathcal{K})\} = 0$ (Exercise). Let $\varphi \in \mathcal{K}$. Then $A_P(X)\varphi = A(X)\varphi$ implies $\Delta(A_P, \varphi) = \Delta(A, \varphi)$. Similarly, $\Delta(B_P, \varphi) = \Delta(B, \varphi)$. Hence,

$$0 \leq \inf\{\Delta(A_\varrho)\Delta(B_\varrho) \mid \varrho \in \mathcal{S}(\mathcal{H})\} \leq \inf\{\Delta(A_P, \varrho)\Delta(B_P, \varrho) \mid \varrho \in \mathcal{S}(\mathcal{K})\} = 0,$$

yielding a contradiction and proving (a). The proofs of (b) and (c) are left as Exercises. $\square$

This proposition indicates that the expression of an uncertainty relation in terms of the product of the standard deviations is not always the optimal way of representing the idea that the two deviations cannot both be small in the same state. Indeed, for sharp observables A and B , with A or B bounded, the product $\Delta(A_\varrho)\Delta(B_\varrho)$ can be made arbitrarily small by choosing one of the quantities, for instance, $\Delta(A_\varrho)$ to be sufficiently small, yielding thus no information on the standard deviation of the other observable. In such a case a lower bound on the sum of the uncertainties could be more informative.

As a simple illustration, we recall from Exercise 4 of Chap. 9 that for the spin component of a spin- $\frac{1}{2}$ object in direction $\mathbf{a} \in \mathbb{R}^3$, $\|\mathbf{a}\| = 1$, one has $\text{Var}(S_{\mathbf{a}}, \varrho_{\mathbf{n}}) = (\frac{\hbar}{2})^2 [1 - (\mathbf{a} \cdot \mathbf{n})^2]$, which gives, for instance,

$$\text{Var}(S_x, \varrho_{\mathbf{n}}) + \text{Var}(S_y, \varrho_{\mathbf{n}}) + \text{Var}(S_z, \varrho_{\mathbf{n}}) \geq \frac{\hbar^2}{2},$$

for any state $\varrho = \varrho_{\mathbf{n}}$, with an equality whenever ϱ is a pure state. We consider examples in Chap. 14 where uncertainty relations for qubit observables are discussed more extensively.

Limitations of the standard deviations as an uncertainty measure become apparent, for instance, in the case where $\mathbf{E}$ is a periodic observable with a period $\ell > 0$, for which $\mathbf{E}([0, \ell)) = I$ (and points $x \in \mathbb{R}$ and $x + \ell$ can be identified). The value space of $\mathbf{E}$ is then homeomorphic with the circle $\mathbb{T}$, which admits many parameterisations (or charts). In this case, one can use the minimum variance

$$\inf_{x \in \mathbb{R}} \left(\inf_{y \in [z - \ell/2, z + \ell/2)} \int_{z - \ell/2}^{z + \ell/2} |x - y|^2 d\mu(x) \right)$$

as a meaningful measure of spread of a probability measure μ on the circle.

Generic Uncertainty Relations for Overall Widths

While the Cauchy–Schwarz inequality readily gives a (state-dependent) lower bound for the product of standard deviations of two observables, there seems to be no similarly simple way of establishing a trade-off relation for the overall widths of a generic pair of incompatible quantities.

There are interesting instances of such relations for specific pairs of observables, such as position and momentum; these are of the form

$$W_{\varepsilon_1}(\mathbf{Q}_{\varrho}) \cdot W_{\varepsilon_2}(\mathbf{P}_{\varrho}) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2, \quad (12.15)$$

valid for $\varepsilon_1, \varepsilon_2 \in (0, 1)$ sufficiently small such that $\varepsilon_1 + \varepsilon_2 \leq 1$. A precise statement with proof is given in Proposition 15.3.

A key ingredient in the proof of the above bound and similar uncertainty relations for overall widths is the following inequality.

Proposition 12.3 (Landau, Pollak [5]) *For two projections P , R , and for any state ϱ the following relation holds:*

$$\text{tr}[\varrho P] + \text{tr}[\varrho R] \leq 1 + \|PR\|. \quad (12.16)$$

According to inequality (12.16), the probabilities for the events or properties represented by P and R cannot both be too large if $\|PR\| < 1$; in particular, the inequality captures the fact the probabilities $\text{tr}[\varrho P]$ and $\text{tr}[\varrho R]$ can only be equal to 1 at the

same time, that is, in the same state ϱ if (and only if) $\|PR\| = 1$ (Exercise). This can happen even when P, R do not commute.

It is possible to prove generic trade-off relations for observables with finitely many outcomes in finite dimensions. For a metric space (Ω, d) , let $O_d(x, L)$ denote the (closed) ball in Ω with centre x and radius $L/2$, that is, $O_d(x, L) = \{y \in \Omega \mid d(y, x) \leq L/2\}$. Further, for a finite set Ω , let $|X|$ denote the counting measure (cardinality) of subsets X of Ω .

Proposition 12.4 (Miyadera [6]) *Let $\mathcal{H}$ be an N -dimensional Hilbert space. Consider a pair of sharp observables $\mathbf{A} = \{A_x \mid x \in \Omega_A\}$ and $\mathbf{B} = \{B_y \mid y \in \Omega_B\}$, where (Ω_A, d_A) and (Ω_B, d_B) are (finite) metric spaces. Then for $\varepsilon_1, \varepsilon_2 \in [0, 1]$ with $\varepsilon_1 + \varepsilon_2 \leq 1$ and for any state ϱ , the overall widths $W_{\varepsilon_1}(\mathbf{A}_\varrho)$ and $W_{\varepsilon_2}(\mathbf{B}_\varrho)$ satisfy*

$$\begin{aligned} \max_{x \in \Omega_A} |O_{d_A}(x, W_{\varepsilon_1}(\mathbf{A}_\varrho))| \cdot \max_{y \in \Omega_B} |O_{d_B}(y, W_{\varepsilon_2}(\mathbf{B}_\varrho))| \\ \geq \frac{(1 - \varepsilon_1 - \varepsilon_2)^2}{\max_{x \in \Omega_A, y \in \Omega_B} \text{tr}[A_x B_y]}. \end{aligned} \quad (12.17)$$

Proof The Landau–Pollak inequality (12.16) implies

$$\sum' \text{tr}[\varrho A_{x'}] + \sum' \text{tr}[\varrho B_{y'}] \leq 1 + \left\| \sum' A_{x'} \sum' B_{y'} \right\|,$$

with summations $\sum'$ over $x' \in O_{d_A}(x, W_{\varepsilon_1}(\mathbf{A}_\varrho))$ and $y' \in O_{d_B}(y, W_{\varepsilon_2}(\mathbf{B}_\varrho))$.

From $\|B\| \leq (\text{tr}[B^* B])^{1/2}$ for any operator B on $\mathcal{H}$, it follows that

$$\begin{aligned} \left\| \sum' A_{x'} \sum' B_{y'} \right\| &\leq \left(\sum' \sum' \text{tr}[B_{y'} A_{x'} B_{y'}] \right)^{1/2} \\ &\leq \max_{x \in \Omega_A} |O_{d_A}(x, W_{\varepsilon_1}(\mathbf{A}_\varrho))|^{1/2} \cdot \max_{y \in \Omega_B} |O_{d_B}(y, W_{\varepsilon_2}(\mathbf{B}_\varrho))|^{1/2} \\ &\quad \cdot \max_{x' \in \Omega_A, y' \in \Omega_B} \text{tr}[A_{x'} B_{y'}]^{1/2}. \end{aligned}$$

Finally, since there are $x \in \Omega_A, y \in \Omega_B$ with

$$1 - \varepsilon_1 \leq \sum' \text{tr}[\varrho A_{x'}], \quad 1 - \varepsilon_2 \leq \sum' \text{tr}[\varrho B_{y'}],$$

then inequality (12.17) follows. $\square$

The strength of inequality (12.17) is illustrated in Chap. 14 with an example for qubit observables.

12.4 Exercises

1. Prove the scaling relation (12.3).

2. Show that the overall width scales with the underlying metric of the value space.
3. Prove the inequality (12.7).
4. Prove the standard uncertainty relation (12.11).
5. Work out the details of the proof of Proposition 12.2 and show that it holds also for all α - and β -spreads.
6. Show that for any sharp observable $\inf\{\Delta(A, \varrho) \mid \varrho \in \mathcal{S}(\mathcal{H})\} = 0$.
7. Show that for any two sharp observables $\mathbf{A}$ and $\mathbf{B}$, if the sum $\Delta(\mathbf{A}_\varrho) + \Delta(\mathbf{B}_\varrho)$ has a state independent strictly positive lower bound then A and B cannot have common eigenstates.
8. Prove the Landau–Pollak inequality (12.16).
9. Let P, R be projections. Show that $\langle \varphi \mid P\varphi \rangle = 1 = \langle \varphi \mid R\varphi \rangle$ for some vector state φ if and only if $\|PR\| = 1$. Give an example of noncommuting projections P, R for which $\|PR\| = 1$. Characterise all pairs of projections for which $\|PR\| = 0$.

References

1. Busch, P., Lahti, P., Mittelstaedt, P.: The Quantum Theory of Measurement. Lecture Notes in Physics Monographs, vol. 2, 2nd revised edn. Springer, Heidelberg (1996)
2. Busch, P., Lahti, P., Werner, R.F.: Measurement uncertainty relations. J. Math. Phys. **55**, 042111 (2014)
3. Kraus, K., Schröter, J.: Expectation values of unbounded observables. Int. J. Theor. Phys. **7**(6), 431–442 (1973)
4. Ishikawa, S.: Uncertainty relations in simultaneous measurements for arbitrary observables. Rep. Math. Phys. **29**(3), 257–273 (1991)
5. Landau, H.J., Pollak, H.O.: Prolate spheroidal wave functions, Fourier analysis and uncertainty II. Bell Syst. Tech. J. **40**, 65–84 (1961)
6. Miyadera, T.: Uncertainty relations for joint localizability and joint measurability in finite-dimensional systems. J. Math. Phys. **52**(7), 072105 (2011)

Chapter 13

Measurement Uncertainty

In Chap. 11 we have seen that noncommutativity prohibits the joint measurability of two sharp observables but not necessarily that of two unsharp observables. We have also observed that adding noise to two incompatible observables can turn them into jointly measurable observables (see Proposition 11.6 and Remark 11.4). This leads to the idea of realising an *approximate joint measurement* of two incompatible sharp observables by means of a joint measurement of two compatible, possibly unsharp observables. In this chapter we develop tools to quantify the approximation errors and the degrees of unsharpness necessary to achieve compatible approximations. This will prepare the ground for formulating measurement uncertainty relations, which make precise Heisenberg's intuitive formulations of 1927 [1].

Heisenberg also considered heuristic, semi-classical *Gedanken* experiments to illustrate his idea that the limited accuracy of joint measurements of noncommuting quantities is the result of an unavoidable mutual disturbance that these measurements exert on each other. It is shown here that the degree of disturbance can be defined operationally as a measure of error in a sequential joint measurement.

We will also give examples of measurement uncertainty relations for specific types of observables such as covariant phase space observables. The problem of generic measurement uncertainty relations is raised and prototypical examples of such relations are formulated for finite-outcome observables in finite dimensional Hilbert spaces. Further cases and some important generalisations will be studied in subsequent chapters, including the question of tight uncertainty relations, which characterise the minimum-error boundary curve in the *uncertainty region* of, say, a given pair of observables.

13.1 Conceptualising Error and Disturbance

Measurement Error

Error measures are to be based on a comparison of the outcomes and statistics of the target observable (that one intends to measure) and the approximating observable (that is actually implemented by the measurement scheme).¹ This ensures that the error measures faithfully represent the results of an *error analysis*; thus, any error measure must

- (EA1) be determined from the statistics of the observables to be compared;
- (EA2) indicate faithfully the presence or absence of any differences between the observables, and suitably quantify the magnitude of the deviation.

In this way it becomes possible to compare theoretical values for approximation errors with the results of an operational error analysis. This is crucial when it comes to performing tests of measurement uncertainty relations. Such tests may be direct or indirect. In a *direct test*, the values of the errors (or disturbances) are estimated from error analysis procedures applied to the joint measurement scheme in question, in line with (EA1), (EA2). In an *indirect test*, the error quantities are determined from the statistics of measurements other than (and not even necessarily compatible with) the approximate joint measurement under investigation. It is only through direct tests that possible shortcomings or limitations of a given proposed error measure may be revealed.

The uncertainty measures introduced in Chap. 12 as indicators of the width of probability distributions will serve as models for the formulation of error measures; the associated preparation uncertainty relations turn out to play a fundamental role in proving measurement uncertainty relations.

All error, unsharpness and disturbance measures introduced here are adapted to observables with value spaces in $\mathbb{R}$.

The problem of quantifying measurement error and disturbance may be approached in two distinct ways. First, one may be interested in the question of how close a given measurement device comes to realising a good approximate measurement of some observable in a particular fixed state of the system. This question can be approached by defining *state-specific* error and disturbance measures. Such state-dependent measures would allow one to determine the imprecision that one has to accept in a measurement of some observable if it is required that the disturbance imparted on some other observable should be limited to a specified amount.

The second approach to quantifying measurement errors is one of interest to a device manufacturer, who would wish to specify a worst-case limit on error and disturbance of a device; this would assure the customers of bounds to error and disturbance that apply to *all* states they wish to measure. Such *device figures of merit* will thus be *state-independent* measures of error and disturbance.

¹This notion of measurement error based on the statistical differences between observables was proposed, for instance, by Ludwig [2, pp. 197–198].

It does not appear to be straightforward to define meaningful state-dependent error measures based on the comparison of distributions. Such measures do not yield nontrivial joint measurement error bounds or error-disturbance trade-off relations. Indeed, any (sharp or unsharp) observable A can be measured precisely, with perfect accuracy. For a given state ϱ one may choose a measurement of A such that ϱ^f is close to (or exactly) ϱ so that for any other observable B the distributions B_{ϱ^f} and B_{ϱ} are almost (or exactly) the same.

Similarly, one can mimick a perfect accuracy joint measurement of any pair of incompatible observables. For any state ϱ and any pair of observables A, B , choose a joint measurement of two trivial observables $C = A_{\varrho}I, D = B_{\varrho}I$, which, if performed on ϱ , is perfectly error free in that it reproduces the statistics A_{ϱ} and B_{ϱ} .

In a sense, this observation is not surprising: it means that if there is prior information about the state, measurements can be adapted so as to produce perfect statistics with negligible disturbance. It also means that the issue of an error-disturbance trade-off will be better addressed with error measures that represent device figures of merit.

There are (at least) two ways of obtaining state-independent error measures. The first is to define a state-dependent measure for all states and define *the worst case error* as the least upper bound of these numbers. Alternatively, one can focus on a representative subset of states, namely, the (near-)eigenstates, and define the mean or the worst-case error across these. Error measures obtained by the latter method will be called *calibration errors*. They are applicable in the case of realistic measuring devices with limited operating ranges, which work only for a restricted set of input states.

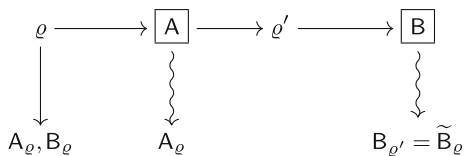
We will introduce several distinct error measures in the coming sections. In each case we will say that an observable is an *approximation* of a given sharp observable if the chosen error measure has a finite value. Each error measure comes with a specific way of splitting it up into two parts—*systematic error* or *bias*, and *random error*. We often adopt a widely accepted definition of *unbiased approximation* as one for which the first moments of the target observable and the approximating observable coincide—a situation addressed already in Corollary 9.1. We emphasise that while this has sometimes been taken to be the only criterion for a “good” approximation, it is by no means enough to characterise approximations in the sense of finite errors as considered here.

Disturbance Caused by a Measurement

The joint measurement of two noncommuting sharp observables such as position and momentum is obstructed by the mutual disturbance of the two measurement schemes involved if one attempts to apply them together or sequentially. The distance between the distorted observables and their original target observables is then a measure of the disturbance. In other words, such sequential schemes can be interpreted as approximate joint measurements of the two observables in question, and the approximation error of the second measured observable is a measure of its disturbance.

As noted in Sects. 10.3 and 11.1, Proposition 11.1, two measurements will in general disturb each other if executed in immediate succession. Thus it makes a difference whether or not a measurement of observable A is carried out before a

Fig. 13.1 Sequential measurement of observables A and then B as a realisation of a joint measurement of A and $\tilde{B}$



measurement of B . If A is sharp, the distorted effects $\tilde{B}(Y) = \mathcal{I}^*(\mathbb{R})(B(Y))$ must commute with $A(X)$ for all X, Y , whatever the second observable B is (here $\mathcal{I}$ denotes the instrument of the measurement of A). If A is a rank-1 observable then $\tilde{B}$ is a smearing of A .

However, if A is an unsharp observable and not rank-1, the distortion of B into $\tilde{B}$ may be less dramatic, and the measurement sequence that defines $\tilde{B}$ may actually constitute an approximate measurement of B . One may cast the irreducible disturbance of B by a measurement of A in an even more general and radical way: rather than measuring B after A , one may measure *any* observable $\tilde{B}$; the resulting sequential joint observable still defines a joint measurement of A and some observable $\tilde{B}$. One can then consider the task of finding the observable $\tilde{B}$ that is optimal in the sense that $\tilde{B}$ is the best approximation to B . This would amount to a way of characterising the minimal amount of unavoidable disturbance of B due to a measurement of A , forced by the incompatibility of A and B .

This consideration confirms the important realisation that the disturbance can be quantified in terms of the error of the distorted observable $\tilde{B}$ relative to B (see Fig. 13.1); in other words, the issue of disturbance is conceptually and operationally reduced to an instance of a (sequential) joint measurement, where the second-measured observable is bound to be approximate due to the disturbance. To conclude: error-disturbance trade-off relations are an instance of joint measurement error trade-off relations. In what follows we will therefore focus on the latter.

13.2 Comparing Distributions

We use the Wasserstein distance of order α to compare probability measures.

For any two probability measures μ, ν on $\mathbb{R}$ we let $\Gamma(\mu, \nu)$ denote the set of their joint measures, also called *couplings*, on $\mathbb{R} \times \mathbb{R}$. Given $\gamma \in \Gamma(\mu, \nu)$, one may define, for any $\alpha, 1 \leq \alpha < \infty$, the ‘cost in transporting’ μ to ν along γ as

$$\Delta_\alpha^\gamma(\mu, \nu) = \left(\int |x - y|^\alpha d\gamma(x, y) \right)^{\frac{1}{\alpha}}, \quad (13.1)$$

with the distance to the power α as the cost function. The infimum of the numbers $\Delta_\alpha^\gamma(\mu, \nu)$, $\gamma \in \Gamma(\mu, \nu)$, is the α -distance or *Wasserstein distance of order α* [3] of μ and ν :²

$$\Delta_\alpha(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \Delta_\alpha^\gamma(\mu, \nu). \quad (13.2)$$

In the case $\alpha = 2$, this expression represents the *root-mean-square (rms)* deviation between two random variables with distributions μ and ν , respectively. For $\nu = \delta_y$, the definition (13.2) coincides with (12.1), since the only coupling is then the product measure $\mu \times \delta_y$. In particular, $\Delta_\alpha(\mu, \delta_y) \geq \Delta_\alpha(\mu)$.

The existence of an optimal coupling minimising (13.1) is known, see, for instance, [3, Theorem 4.1], but it does not imply that $\Delta_\alpha(\mu, \nu)$ is finite. Though the distance can become infinite, the triangle inequality

$$\Delta_\alpha(\mu, \zeta) \leq \Delta_\alpha(\mu, \nu) + \Delta_\alpha(\nu, \zeta) \quad (13.3)$$

still holds [3, after Example 6.3]. The proof relies on Minkowski's inequality and the use of a “Gluing Lemma”, which builds a coupling from μ to ζ out of couplings from μ to ν and from ν to ζ .

We define the (*Wasserstein*) α -distance between any two real observables E and F :

$$\Delta_\alpha(E, F) = \sup_{\varrho \in S(\mathcal{H})} \Delta_\alpha(E_\varrho, F_\varrho). \quad (13.4)$$

The α -distances are global error measures, taking into account the largest possible deviations between corresponding probability measures of the observables being compared. As such they lend themselves to *direct* tests of associated measurement uncertainty relations.

The following two lemmas, from [4], will be crucial.

Lemma 13.1 *For probability measures μ, ν, η the following relations hold.*

$$\Delta_\alpha(\eta * \mu, \eta * \nu) \leq \Delta_\alpha(\mu, \nu); \quad (13.5)$$

$$\Delta_\alpha(\eta * \mu, \mu) \leq \Delta_\alpha(\eta, \delta_0). \quad (13.6)$$

Proof First we note that any coupling γ for μ, ν gives a coupling $\tilde{\gamma}(X \times Y) = \int \gamma(X - x, Y - x) d\eta(x)$ for $\eta * \mu, \eta * \nu$, for which we get $\int |x - y|^\alpha d\tilde{\gamma}(x, y) = \int |x - y|^\alpha d\gamma(x, y)$. Since the infimum may be attained at a coupling different from $\tilde{\gamma}$, Eq. (13.5) follows.

To prove (13.6), note that it is a special case of (13.5) since $\Delta_\alpha(\eta * \mu, \mu) = \Delta_\alpha(\mu * \eta, \mu * \delta_0)$. $\square$

²Again, an extension to $\alpha = \infty$ is possible [4] but will not be considered here.

Lemma 13.2 *Let E be a sharp observable on $\mathbb{R}$, η some probability measure on $\mathbb{R}$, and $F = \eta * E$, that is, $F_\varrho = \eta * E_\varrho$ for all states ϱ . Then*

$$\Delta_\alpha(F, E) = \Delta_\alpha(\eta, \delta_0). \quad (13.7)$$

Proof By (13.6) we have $\Delta_\alpha(F_\varrho, E_\varrho) \leq \Delta_\alpha(\eta, \delta_0)$. We claim that this upper bound is nearly attained whenever E_ϱ is sharply concentrated, say, $\Delta_\alpha(E_\varrho, \delta_q) \leq \varepsilon$; this is possible, because E is sharp. Indeed, we then have

$$\begin{aligned} \Delta_\alpha(\eta, \delta_0) &= \Delta_\alpha(\eta * \delta_q, \delta_q) \\ &\leq \Delta_\alpha(\eta * \delta_q, \eta * E_\varrho) + \Delta_\alpha(\eta * E_\varrho, E_\varrho) + \Delta_\alpha(E_\varrho, \delta_q) \\ &\leq 2\varepsilon + \Delta_\alpha(\eta * E_\varrho, E_\varrho) = 2\varepsilon + \Delta_\alpha(F_\varrho, E_\varrho). \quad \square \end{aligned}$$

This lemma combined with the preparation uncertainty relation (12.14) yields the following basic result.

Proposition 13.1 (Measurement uncertainty relations for covariant phase space measurements) *For any covariant phase space observable G^T , with the marginal observables $\mu * Q$ and $\nu * P$, one has*

$$\Delta_2(\mu * Q, Q) \Delta_2(\nu * P, P) \geq \frac{\hbar}{2}. \quad (13.8)$$

Proof Since $\mu = Q_{\Pi T \Pi}$ and $\nu = P_{\Pi T \Pi}$, the result is immediate by Lemma 13.2 and the inequality (12.14). $\square$

With Proposition 15.1 this result generalises to all powers α, β .

It is important to note that the inequality (13.8) constitutes a limitation for the accuracy of joint measurements within the restricted class of covariant phase space observables. It is a fundamental feature of quantum mechanics, proved only recently in full generality, that this limitation extends to all possible attempted joint measurements of position and momentum. This result will be stated and proven in Chap. 15 in terms of the α -distance between observables.

For any $\gamma \in \Gamma(\mu, \nu)$ and for any pair of functions $(f, g) \in L^1(\nu) \times L^1(\mu)$, with

$$f(y) - g(x) \leq |x - y|^\alpha, \quad (13.9)$$

one has

$$\int f d\nu - \int g d\mu \leq \Delta_\alpha(\mu, \nu)^\alpha. \quad (13.10)$$

The *Kantorovich Duality Theorem* asserts that the supremum of the left hand side over the functions satisfying (13.9) closes the gap:

$$\Delta_\alpha(\mu, \nu)^\alpha = \sup_{f, g} \left\{ \int f d\nu - \int g d\mu \right\}. \quad (13.11)$$

The statement of Kantorovich Duality in [3, Theorem 5.10] includes that the supremum (13.11) is attained when one restricts the set of functions to bounded continuous functions. Moreover, this set can be further reduced to positive continuous functions of compact support without changing the value of the supremum [4].

Calibration Error

In the calibration of a new measuring device one typically tests the apparatus by inputting states with (nearly) sharp values of the quantity to be measured, to determine the spread of outcome readings, which ought to be narrowly distributed around the “actual” value. This way of estimating the *calibration error* also accommodates the fact that every measurement device has only a finite operating range, which means that the supremum (13.4) over all states is not easily accessible to experimental implementation. It seems more reasonable to just calibrate the performance of a measurement of $\mathbf{F}$ as an approximate measurement of $\mathbf{E}$ by looking at the distributions $\mathbf{F}_\varrho$ for preparations for which $\mathbf{E}_\varrho$ is nearly a point measure, i.e., those for which $\mathbf{E}$ has a nearly sharp value. The idea of calibration error was formalised in [5] as a measure of *error bar width*, which was shown to obey a measurement uncertainty relation using the method developed in [6] and applied here.

For $\varepsilon > 0$, we define the ε -calibration error (resp. the *calibration error*) of $\mathbf{F}$ with respect to $\mathbf{E}$ as

$$\Delta_\alpha^\varepsilon(\mathbf{F}, \mathbf{E}) = \sup_{\varrho, x} \left\{ \Delta_\alpha(\mathbf{F}_\varrho, \delta_x) \mid \Delta_\alpha(\mathbf{E}_\varrho, \delta_x) \leq \varepsilon \right\}, \quad (13.12)$$

$$\Delta_\alpha^c(\mathbf{F}, \mathbf{E}) = \lim_{\varepsilon \rightarrow 0} \Delta_\alpha^\varepsilon(\mathbf{F}, \mathbf{E}), \quad (13.13)$$

respectively. Here the limit in (13.13) exists because (13.12) is a monotonely decreasing function. By the triangle inequality, we have $\Delta_\alpha(\mathbf{F}_\varrho, \delta_x) \leq \Delta_\alpha(\mathbf{F}_\varrho, \mathbf{E}_\varrho) + \Delta_\alpha(\mathbf{E}_\varrho, \delta_x)$ and, taking the supremum over ϱ and x as in (13.12)

$$\Delta_\alpha^\varepsilon(\mathbf{F}, \mathbf{E}) \leq \Delta_\alpha(\mathbf{F}, \mathbf{E}) + \varepsilon \quad \text{and} \quad \Delta_\alpha^c(\mathbf{F}, \mathbf{E}) \leq \Delta_\alpha(\mathbf{F}, \mathbf{E}). \quad (13.14)$$

When $\mathbf{F}$ just adds independent noise, there is also the complementary inequality [4], a direct analog of Lemma 13.2.

Lemma 13.3 *Let $\mathbf{E}$ be a sharp observable on $\mathbb{R}$, η some probability measure on $\mathbb{R}$, and $\mathbf{F} = \eta * \mathbf{E}$. Then*

$$\Delta_\alpha(\eta, \delta_0) - \varepsilon \leq \Delta_\alpha^\varepsilon(\mathbf{F}, \mathbf{E}) \leq \Delta_\alpha(\eta, \delta_0) + \varepsilon, \quad (13.15)$$

and so

$$\Delta_\alpha^c(\mathbf{F}, \mathbf{E}) = \Delta_\alpha(\eta, \delta_0). \quad (13.16)$$

Proof For a calibration state ϱ , i.e., $\Delta_\alpha(\mathbf{E}_\varrho, \delta_x) \leq \varepsilon$, one has the upper bound $\Delta_\alpha(\mathbf{F}_\varrho, \delta_x) = \Delta_\alpha(\eta * \mathbf{E}_\varrho, \delta_x) \leq \Delta_\alpha(\eta * \mathbf{E}_\varrho, \mathbf{E}_\varrho) + \Delta_\alpha(\mathbf{E}_\varrho, \delta_x) \leq \Delta_\alpha(\eta, \delta_0) + \varepsilon$.

For the complementary bound one can use $\Delta_\alpha(\eta, \delta_0) = \Delta_\alpha(\eta * \delta_x, \delta_x) \leq \Delta_\alpha(\eta * \delta_x, \eta * \mathbf{E}_\varrho) + \Delta_\alpha(\eta * \mathbf{E}_\varrho, \delta_x) \leq \varepsilon + \Delta_\alpha(\eta * \mathbf{E}_\varrho, \delta_x)$. Thus,

$$\begin{aligned}\Delta_\alpha(\eta, \delta_0) - \varepsilon &\leq \Delta_\alpha(\mathbf{F}_\varrho, \delta_x) \leq \Delta_\alpha(\eta, \delta_0) + \varepsilon, \\ \Delta_\alpha(\eta, \delta_0) - \varepsilon &\leq \Delta_\alpha^\varepsilon(\mathbf{F}, \mathbf{E}) \leq \Delta_\alpha(\eta, \delta_0) + \varepsilon,\end{aligned}$$

where the second row is obtained by taking the supremum of the term $\Delta_\alpha(\mathbf{F}_\varrho, \delta_x)$ in the first over all x and all calibrating states. $\square$

Hence, in the case of convolution observables $\mathbf{F} = \eta * \mathbf{E}$, with $\mathbf{E}$ sharp, we have

$$\Delta_\alpha(\mathbf{F}, \mathbf{E}) = \Delta_\alpha^c(\mathbf{F}, \mathbf{E}). \quad (13.17)$$

In general, however, the inequality (13.14) is strict (exercise). In view of Eq. (13.17), the error trade-off inequality (13.8) for covariant phase space observables can also be expressed in terms of calibration errors.

13.3 Error Bar Width

Gross Error Bar

Next we give a definition of measurement error that formalises the common practice of calibrating measuring instruments in terms of likely error intervals. In the process of calibration, one tests the device by applying it to a sufficiently large family of input states in which the observable one wishes to measure with this setup has fairly sharp values. The error is then characterised as an overall measure of the bias, or systematic error, and the overall width of the output distribution as the random error, across a range of input values [5, 7]. Error bars give the minimal average interval lengths that one has to allow to contain all output values with a given confidence level. We first consider (*gross*) *error bars*, which disregard the distinction between systematic and random error, and will show later how to separate out the bias contribution.

For simplicity, we give the following definitions only for approximations of a sharp observable $\mathbf{E}$, so that the assumption of localised input states can be described as $\mathbf{E}_\varrho(J_{x;\delta}) = 1$, for intervals $J_{x;\delta} = [x - \delta/2, x + \delta/2]$, $x \in \mathbb{R}$, $\delta > 0$. Let $\mathbf{E}$, $\mathbf{F}$ be observables on $\mathbb{R}$ and $\mathbf{E}$ be sharp. For each $\varepsilon \in (0, 1)$, $\delta > 0$, we define the *error* of $\mathbf{F}$ relative to $\mathbf{E}$

$$\begin{aligned}\mathcal{W}_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E}) = & \quad (13.18) \\ \inf\{w \geq 0 \mid \forall x \in \mathbb{R} \forall \varrho \in \mathcal{S}(\mathcal{H}) : \mathbf{E}_\varrho(J_{x;\delta}) = 1 \Rightarrow \mathbf{F}_\varrho(J_{x,w}) \geq 1 - \varepsilon\}.\end{aligned}$$

Lemma 13.4 The infimum in Eq. (13.18) is a minimum.

Proof For $x \in \mathbb{R}$, $\varrho \in \mathcal{S}(\mathcal{H})$, denote

$$\Delta_{x,\varrho} = \{w \geq 0 \mid \mathbf{E}_{\varrho}(J_{x,\delta}) = 1 \implies \mathbf{F}_{\varrho}(J_{x,w}) \geq 1 - \epsilon\}.$$

If $\mathbf{E}_{\varrho}(J_{x,\delta}) \neq 1$, then $\Delta_{x,\varrho} = [0, \infty)$. If $\mathbf{E}_{\varrho}(J_{x,\delta}) = 1$, then $\Delta_{x,\varrho}$ is more generally of the form $[\alpha_{x,\varrho}, \infty)$ for some $\alpha_{x,\varrho} \geq 0$. Indeed, if $w \in \Delta_{x,\varrho}$, then $y \in \Delta_{x,\varrho}$ for any $y \geq w$, and taking a decreasing sequence in $\Delta_{x,\varrho}$ converging to $\inf \Delta_{x,\varrho}$ we see that $\inf \Delta_{x,\varrho} \in \Delta_{x,\varrho}$. Since

$$\mathcal{W}_{\epsilon,\delta} = \bigcap_{x \in \mathbb{R}, \varrho \in \mathcal{S}(\mathcal{H})} [\alpha_{x,\varrho}, \infty)$$

and the latter set is closed and contains the interval $[t, \infty)$ for its every element t , we have $\mathcal{W}_{\epsilon,\delta} = [\alpha, \infty)$ where $\alpha = \inf \mathcal{W}_{\epsilon,\delta}$. $\square$

The error describes the range within which the input values can be inferred from the output distributions, with confidence level $1 - \epsilon$, given initial localisations within δ .

$\mathbf{F}$ is called an ϵ -approximation to $\mathbf{E}$ if $\mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}) < \infty$ for all $\delta > 0$. The error is an increasing function of δ , so that one can define the (gross) error bar width of $\mathbf{F}$ relative to $\mathbf{E}$:

$$\mathcal{W}_{\epsilon}(\mathbf{F}, \mathbf{E}) = \inf_{\delta} \mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}) = \lim_{\delta \rightarrow 0} \mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}). \quad (13.19)$$

In case $\mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}) = \infty$ for all $\delta > 0$, we write $\mathcal{W}_{\epsilon}(\mathbf{F}, \mathbf{E}) = \infty$.

$\mathbf{F}$ is called an *approximation (in the sense of finite error bar width)* to $\mathbf{E}$ if $\mathcal{W}_{\epsilon}(\mathbf{F}, \mathbf{E}) < \infty$ for all $\epsilon \in (0, \frac{1}{2})$. The restriction to $\epsilon < \frac{1}{2}$ reflects the idea that a “good” approximation should have confidence levels greater than $\frac{1}{2}$.

It is worth noting that if $\mathbf{F}$ is an approximation to $\mathbf{E}$, the map $\epsilon \mapsto \mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E})$ is a decreasing function of $\epsilon \in (0, \frac{1}{2})$ for every $\delta > 0$.

It is possible to characterise the case of an accurate measurement of the sharp observable $\mathbf{E}$.

Proposition 13.2 *Let $\mathbf{F}$ be an approximation to the sharp observable $\mathbf{E}$. Then the following are equivalent:*

- (a) $\mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}) \leq \delta$ for all $\epsilon \in (0, 1)$, $\delta > 0$;
- (b) $\mathbf{F} = \mathbf{E}$.

If either of these condition is fulfilled then $\mathcal{W}_{\epsilon}(\mathbf{F}, \mathbf{E}) = 0$ for all $\epsilon \in (0, 1)$.

Proof Assume (b) holds. Let $\epsilon \in (0, 1)$, $\delta > 0$. Choose $w = \delta$; then for any $x \in \mathbb{R}$ and any state ϱ with $\mathbf{E}_{\varrho}(J_{x;\delta}) = 1$, we also have $\mathbf{E}_{\varrho}(J_{x;\delta}) \geq 1 - \epsilon$. This shows that $\mathcal{W}_{\epsilon,\delta}(\mathbf{E}, \mathbf{E}) \leq \delta$.

Conversely, assume that (a) holds. Consider any $\epsilon \in (0, 1)$, $\delta > 0$. For $w = \mathcal{W}_{\epsilon,\delta}(\mathbf{F}, \mathbf{E}) \leq \delta$, we have, for all $x \in \mathbb{R}$ and all ϱ with $\mathbf{E}_{\varrho}(J_{x;\delta}) = 1$, that $\mathbf{F}_{\varrho}(J_{x;\delta}) \geq \mathbf{F}_{\varrho}(J_{x;w}) \geq 1 - \epsilon$. This entails for any vector state φ for which $\mathbf{E}(J_{x;\delta})\varphi = \varphi$ that

$\langle \varphi | F(J_{x;\delta}) \varphi \rangle \geq 1 - \varepsilon$. As this holds for any $\varepsilon \in (0, 1)$, we obtain $\langle \varphi | F(J_{x;\delta}) \varphi \rangle = 1$. This entails that $E(J_{x;\delta}) \leq F(J_{x;\delta})$. Since $x \in \mathbb{R}$ and $\delta > 0$ are arbitrary, this ordering holds for any closed interval $J = [a, b]$. Using the weak σ -additivity of E and F we get $E((-\infty, t]) = \text{w-lim}_{n \rightarrow \infty} E([t - n, t]) \leq \text{w-lim}_{n \rightarrow \infty} F([t - n, t]) = F((-\infty, t])$ and $E((t, \infty)) = \text{w-lim}_{n \rightarrow \infty} E([t + \frac{1}{n}, t + n]) \leq \text{w-lim}_{n \rightarrow \infty} F([t + \frac{1}{n}, t + n]) = F(t, \infty)$. Since $E(\mathbb{R}) = F(\mathbb{R}) = I$, we must have $E((-\infty, t]) = F((-\infty, t])$ for all $t \in \mathbb{R}$, and so $E = F$ by Corollary 4.1. $\square$

We leave it as an exercise to confirm that the error bar width scales with the metric on the underlying value space.

Proposition 13.3 *The smeared position (or momentum) observable $\mu * Q$ is an approximation (in the sense of finite error bar width) to Q for any probability measure μ . In particular,*

$$\mathcal{W}_\varepsilon(\mu * Q, Q) \geq W_\varepsilon(\mu), \quad (13.20)$$

where $W_\varepsilon(\mu)$ is the overall width of μ (Eq. (12.6)).

Proof Let $\varepsilon \in (0, 1)$, $\delta > 0$ be given. We have to show that there is a finite number $w > 0$ such that for all $q \in \mathbb{R}$ one has $(\mu * Q)_\varrho(J_{q;w}) \geq 1 - \varepsilon$ whenever $Q_\varrho(J_{q;\delta}) = 1$.

Thus, fix an arbitrary q , and let ϱ be any state such that $Q_\varrho(J_{q;\delta}) = 1$. Let q_0, w_0 be such that $\mu(J_{q_0;w_0}) \geq 1 - \varepsilon$. Next, we note that $J_{q;\delta} \subset J_{q;w} - x$ holds for all $x \in J_{q_0;w_0}$ exactly when $w \geq 2|q_0| + w_0 + \delta$. In that case, $Q_\varrho(J_{q;w} - x) = 1$ for all such x . Then:

$$\begin{aligned} \mu * Q_\varrho(J_{q;w}) &= \int \mu(dx) Q_\varrho(J_{q;w} - x) \geq \int_{J_{q_0;w_0}} \mu(dx) Q_\varrho(J_{q;w} - x) \\ &= \mu(J_{q_0;w_0}) \geq 1 - \varepsilon. \end{aligned}$$

The inequality $\mathcal{W}_\varepsilon(\mu * Q, Q) \geq W_\varepsilon(\mu)$ follows from the above constraint on w by taking the limit $\delta \rightarrow 0$ and noting that the infimum of the w_0 is $W_\varepsilon(\mu)$. This gives $\mathcal{W}_\varepsilon(\mu * Q, Q) \geq 2|q_0| + W_\varepsilon(\mu)$. $\square$

If the distribution μ is *unbiased* in the sense that q_0 can be chosen to be zero, so that $\mu(J_{0;w_0}) \geq 1 - \varepsilon$ in the above proof, then it follows that $\mathcal{W}_\varepsilon(\mu * Q, Q) = W_\varepsilon(\mu)$. This will be shown below in terms of the bias-free error bar width.

It is not hard to construct approximations of Q (with finite error bars) that do not share the translation covariance of Q .

Example 13.1 Let f be a continuous function on $\mathbb{R}$ which is one-to-one and such that $f(q) - q$ is not constant but $|f(q) - q| \leq \alpha$ for all $q \in \mathbb{R}$ and some fixed $\alpha > 0$. An example is $f(q) = q + \alpha \cos(q)$. Let Q_μ be a smeared position observable. Then $(\mu * Q) \circ f^{-1}$ is a non-covariant approximation to Q in the sense of finite error bars. (The proof is left as an exercise.) $\triangleleft$

Example 13.2 For any bounded Borel function f on $\mathbb{R}$, the observable $\mathbf{Q} \circ f^{-1}$ has infinite error bars with respect to $\mathbf{Q}$. Indeed, let J be a bounded interval which contains the range of f . Then for any finite $w > 0$, one can find q such that $J_{q;w} \cap J = \emptyset$. Then $f^{-1}(J_{q;w}) = \emptyset$ and it follows that $(\mathbf{Q} \circ f^{-1})_{\varrho}(J_{q;w}) = \mathbf{Q}_{\varrho}(f^{-1}(J_{q;w})) = 0$ for all ϱ . Hence $\mathcal{W}_{\varepsilon,\delta}(\mathbf{Q} \circ f^{-1}, \mathbf{Q}) = \infty$ for all $\varepsilon \in (0, 1)$ and all $\delta > 0$. $\triangleleft$

Proposition 13.4 Let $\mu * \mathbf{Q}, \nu * \mathbf{P}$ be the margins of a covariant phase space observable. Then for $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 < 1$ the following error bar trade-off relation holds:

$$\mathcal{W}_{\varepsilon_1}(\mu * \mathbf{Q}, \mathbf{Q}) \mathcal{W}_{\varepsilon_2}(\nu * \mathbf{P}, \mathbf{Q}) \geq 2\pi\hbar(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (13.21)$$

Proof This is a consequence of inequalities (13.20) and (12.15). $\square$

Bias-Free Error

We show next how the gross error can be decomposed into a (positive) bias term (systematic error) and a bias-free (random) error. Let $\varepsilon \in (0, 1)$ and $\delta > 0$ be given. Let $\mathbf{F}, \mathbf{E}$ be observables on $\mathbb{R}$ and $\mathbf{E}$ be sharp. Note that the condition $\mathbf{E}_{\varrho}(J_{x;\delta}) = 1$ (for some $x \in \mathbb{R}$) can be expressed as $W_0(\mathbf{E}_{\varrho}) \leq \delta$. We define the *bias-free error* $\mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E})$ as follows:

$$\mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E}) = \sup \{ \mathcal{W}_{\varepsilon}(\mathbf{F}_{\varrho}) \mid \varrho \in \mathcal{S}(\mathcal{H}), W_0(\mathbf{E}_{\varrho}) \leq \delta \}. \quad (13.22)$$

This is a measure of the overall minimal error, determined by the overall widths of all output distributions, given input distributions supported in intervals $J_{q;\delta}$. If this quantity is finite for some δ_0 , it is an increasing function for all $\delta \leq \delta_0$. In that case we can define the *bias-free error bar width*,

$$\mathcal{W}_{\varepsilon}^0(\mathbf{F}, \mathbf{E}) = \lim_{\delta \rightarrow 0} \mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E}). \quad (13.23)$$

The following is obvious:

$$\mathcal{W}_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E}) \geq \mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E}). \quad (13.24)$$

If these quantities are finite, one then has in the limit $\delta \rightarrow 0$:

$$\mathcal{W}_{\varepsilon}(\mathbf{F}, \mathbf{E}) \geq \mathcal{W}_{\varepsilon}^0(\mathbf{F}, \mathbf{E}). \quad (13.25)$$

The difference between $\mathcal{W}_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E})$ and $\mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E})$ disappears when the output distributions are concentrated at the locations of the input distributions, that is, around the intervals $J_{x;\delta}$. This is to say that the difference is a measure of the overall magnitude of the *bias* $\beta_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E})$ inherent in $\mathbf{F}$ relative to $\mathbf{E}$:

$$\beta_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E}) = \mathcal{W}_{\varepsilon,\delta}(\mathbf{F}, \mathbf{E}) - \mathcal{W}_{\varepsilon,\delta}^0(\mathbf{F}, \mathbf{E}) \geq 0. \quad (13.26)$$

Rephrasing this as

$$\mathcal{W}_{\varepsilon, \delta}(\mathbf{F}, \mathbf{E}) = \mathcal{W}_{\varepsilon, \delta}^0(\mathbf{F}, \mathbf{E}) + \beta_{\varepsilon, \delta}(\mathbf{F}, \mathbf{E}), \quad (13.27)$$

we see that the gross error is decomposed into the bias-free error and the magnitude of the bias. Note that one can take the limit of $\delta \rightarrow 0$:

$$\beta_{\varepsilon}(\mathbf{F}, \mathbf{E}) = \mathcal{W}_{\varepsilon}(\mathbf{F}, \mathbf{E}) - \mathcal{W}_{\varepsilon}^0(\mathbf{F}, \mathbf{E}). \quad (13.28)$$

As an immediate consequence of these definitions, we can say that $\mathbf{F}$ is an ε -approximation to $\mathbf{E}$ if and only if the bias and random errors are finite for all $\delta > 0$.

Proposition 13.5 *Let $\mu * \mathbf{Q}$ be a smeared position (or momentum) observable. Then*

$$\mathcal{W}_{\varepsilon}^0(\mu * \mathbf{Q}, \mathbf{Q}) = W_{\varepsilon}(\mu). \quad (13.29)$$

Proof We note first that $W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho}) \geq W_{\varepsilon}(\mu)$ (see (12.7)). Using this, we show that whenever $W_0(\mathbf{Q}_{\varrho}) \leq \delta$, then $W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho}) \leq W_{\varepsilon}(\mu) + \delta$.

We are given that $\mathbf{Q}_{\varrho}(J_{q_0; \delta}) = 1$ for some $q_0 \in \mathbb{R}$. Assume $w \geq W_{\varepsilon}(\mu)$, that is, $\mu(J_{q_1; w}) \geq 1 - \varepsilon$ for some q_1 . We have to show that $w + \delta \geq W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho})$, that is, $\mu * \mathbf{Q}_{\varrho}(J_{q_2; w+\delta}) \geq 1 - \varepsilon$ for some $q_2 \in \mathbb{R}$.

Let $q_2 = q_0 - q_1$. Then it follows that $q + J_{q_2; w+\delta} \supset J_{q_0; \delta}$ for all $q \in J_{q_1; w}$. Then

$$\begin{aligned} \mu * \mathbf{Q}_{\varrho}(J_{q_2; w+\delta}) &= \int \mu(dq) \mathbf{Q}_{\varrho}(q + J_{q_2; w+\delta}) \\ &\geq \int_{J_{q_1; w}} \mu(dq) = \mu(J_{q_1; w}) \geq 1 - \varepsilon. \end{aligned}$$

This shows that $w \geq W_{\varepsilon}(\mu)$ implies $w + \delta \geq W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho})$ whenever $W_0(\mathbf{Q}_{\varrho}) \leq \delta$. Thus, under this assumption we let w approach $W_{\varepsilon}(\mu)$ to obtain $W_{\varepsilon}(\mu) + \delta \geq W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho})$.

To summarise, we have shown: $W_{\varepsilon}(\mu) \leq W_{\varepsilon}(\mu * \mathbf{Q}_{\varrho}) \leq W_{\varepsilon}(\mu) + \delta$, where the latter inequality holds if $W_0(\mathbf{Q}_{\varrho}) \leq \delta$. This entails $W_{\varepsilon}(\mu) \leq \mathcal{W}_{\varepsilon, \delta}^0(\mu * \mathbf{Q}, \mathbf{Q}) \leq W_{\varepsilon}(\mu) + \delta$. Now we can take the limit $\delta \rightarrow 0$ to obtain the result. $\square$

Corollary 13.1 *Let $\mathbf{G}^T$ be a covariant phase space observable with margins $\mu * \mathbf{Q}$ and $\nu * \mathbf{P}$. Then for $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 < 1$ the following trade-off relation holds:*

$$\mathcal{W}_{\varepsilon}^0(\mu * \mathbf{Q}, \mathbf{Q}) \mathcal{W}_{\varepsilon}^0(\nu * \mathbf{P}, \mathbf{P}) \geq 2\pi\hbar(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (13.30)$$

Proof Since $\mu = \mathbf{Q}_{\Pi T \Pi}$ and $\nu = \mathbf{P}_{\Pi T \Pi}$, this is a consequence of Eq. (13.29) and inequality (12.15) $\square$

13.4 Value Comparison Error

Instead of merely comparing the *distributions* of some sharp target observable $\mathbf{A}$ and its approximating observable $\mathbf{C}$, one may consider comparing their *values*. But to do so, one needs to be able to measure both observables on the same system, that is, they must be jointly measurable.

The mean error of the approximate measurement, a measurement of $\mathbf{C}$, can then be defined as the root-mean-square (rms) deviation of its outcomes c_k from the “true value”, a_k , obtained from the accurate reference measurement of $\mathbf{A}$, that is, symbolically, $\langle (c_k - a_k)^2 \rangle^{1/2}$. This error measure will be called (*rms*) *value deviation*.

For $\mathbf{C}$ to be compatible with the sharp observable $\mathbf{A}$, the two observables must commute, and their joint observable is of the product form $\mathbf{M}(X \times Y) = \mathbf{A}(X)\mathbf{C}(Y)$, by Proposition 4.8. Denoting the resulting state-dependent value deviation measure by $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$, one therefore has

$$\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2 = \iint (x - y)^2 \text{tr}[\varrho \mathbf{A}(dx) \mathbf{C}(dy)]. \quad (13.31)$$

Since $\mathbf{M}_\varrho$ is a coupling of $\mathbf{A}_\varrho$ and $\mathbf{C}_\varrho$, the value deviation $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$ is, in general, greater than the Wasserstein-2 deviation of $\mathbf{A}_\varrho$ and $\mathbf{C}_\varrho$. For later reference we expand this expression in terms of moment operators of $\mathbf{A}$ and $\mathbf{C}$, using the intrinsic noise operator $N(\mathbf{C}) = \mathbf{C}[2] - \mathbf{C}[1]^2$:

$$\begin{aligned} \Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2 &= \text{tr}[\varrho(\mathbf{C}[2] - \mathbf{C}[1]^2)] + \text{tr}[\varrho(\mathbf{C}[1] - \mathbf{A}[1])^2] \\ &= \text{tr}[\varrho N(\mathbf{C})] + \text{tr}[\varrho(\mathbf{C}[1] - \mathbf{A}[1])^2]. \end{aligned} \quad (13.32)$$

We stress that the value-comparison method of defining approximation errors is of limited utility in quantum mechanics as its applicability is restricted to approximating observables that are compatible with the target observable. Otherwise, there are no simultaneous values of $\mathbf{A}$, $\mathbf{C}$ and no distribution thereof, from which to compute a deviation. This can also be seen from the alternative expression above, showing $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2$ as the expectation of the operator $N(\mathbf{C}) + (\mathbf{C}[1] - \mathbf{A}[1])^2$.

If $\mathbf{A}$ and $\mathbf{C}$ are incompatible, there is no canonical way to determine the expectation of the squared difference operator $(\mathbf{C}[1] - \mathbf{A}[1])^2$ from the $\mathbf{C}$ and $\mathbf{A}$ measurements. To obtain this value, one should construct a measurement of the difference observable $\mathbf{E}^{\mathbf{C}[1]-\mathbf{A}[1]}$ to determine the second moment of its outcome statistics. Hence, while the right hand side of (13.32) is well defined also for incompatible observables $\mathbf{A}$ and $\mathbf{C}$, there is no reason to expect that this quantity represents a good definition of the approximation error for $\mathbf{A}$ by a measurement of $\mathbf{C}$ in the state ϱ .

We conclude that the value deviation cannot serve to formulate measurement uncertainty relations that are *universal* in the sense of being valid for arbitrary approximations. In fact, we will meet examples of joint approximations of two incompatible observables $\mathbf{A}$, $\mathbf{B}$ by means of compatible observables $\mathbf{C}$, $\mathbf{D}$ where the optimal

approximations $\mathbf{C}$, $\mathbf{D}$ are among the observables that do *not* commute with $\mathbf{A}$, $\mathbf{B}$, respectively.

If $\varrho = \varrho_a$ is an eigenstate of $\mathbf{A}$ with eigenvalue a , the value deviation becomes

$$\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho_a) = \left(\int (x - a)^2 \text{tr}[\varrho_a \mathbf{C}(dx)] \right)^{1/2}. \quad (13.33)$$

This correctly reproduces the rms deviation of the measured values of $\mathbf{C}$ from the “true” value. In the general case, however, the initial uncertainty of the $\mathbf{A}$ values entails that the value deviation encompasses not solely the approximation error of $\mathbf{C}$ relative to $\mathbf{A}$, but reflects also the preparation uncertainty of $\mathbf{A}$. This is consistent with the fact that $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho) \geq \Delta_2(\mathbf{C}_\varrho, \mathbf{A}_\varrho)$.

Example 13.3 Let $\mathbf{C}$ be a trivial observable of the form $\mathbf{C}(X) = \mu(X)I$, with a probability measure μ of finite variance. If ϱ is a state for which the sharp observable $\mathbf{A}$ has a nonzero finite variance, then

$$\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2 = \Delta(\mathbf{A}_\varrho)^2 + \Delta(\mu)^2 + (\mathbf{A}_\varrho[1] - \mu[1]I)^2. \quad (13.34)$$

This expression comprises the squared systematic error, given by the last term, and the random error (or imprecision), given by the middle term; in addition, there is a contribution arising from the preparation uncertainty of $\mathbf{A}$ in the state ϱ , which in the present simple example can be subtracted to obtain a measure of the error alone. $\triangleleft$

The interpretation of $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$ as an approximation error for the observable $\mathbf{C}$ relative to $\mathbf{A}$ is therefore not as clear-cut as it is in the case where $\mathbf{A}$ has a sharp value. In general, $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$ is a characteristic of the joint observable $\mathbf{M}$ of $\mathbf{A}$, $\mathbf{C}$ (where it exists), in that it is sensitive not only to the separate distributions of these observables but also to *correlations* between them as they are measured jointly.

Example 13.4 Consider an accurate measurement of $\mathbf{A}$, followed by a measurement of $\mathbf{C} = \mathbf{E}^{-\mathbf{A}}$. In that case one has

$$\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2 = \int 4x^2 \text{tr}[\varrho \mathbf{A}(dx)] = 4\mathbf{A}_\varrho[2] = 4\Delta(\mathbf{A}_\varrho)^2 + 4\mathbf{A}_\varrho[1]^2.$$

The last term correctly represents the systematic error since the mean of the distribution of $\mathbf{C}$ is always the negative of the mean of the $\mathbf{A}$ distribution. The first term on the right hand side reflects the maximal anticorrelation between the two observables. $\triangleleft$

Remark 13.1 For a (real) discrete sharp observable $\mathbf{A} = \{A_1, A_2, \dots\}$ with outcomes $\{a_1, a_2, \dots\}$ and a (real) discrete approximating observable $\mathbf{C} = \{C_1, C_2, \dots\}$ with outcomes $\{c_1, c_2, \dots\}$ the value comparison error can be implemented by a simple measurement scheme if $\mathbf{A}$ and $\mathbf{C}$ commute. If the measurement of $\mathbf{C}$ is preceded by a Lüders measurement of $\mathbf{A}$, the sequential joint probability for an outcome pair (a_k, c_ℓ) is given by $p(a_k, c_\ell) = \text{tr}[\varrho A_k C_\ell]$. This is just the joint probability for $\mathbf{A}$

and $\mathbf{C}$, and the value deviation for this measurement sequence is exactly the quantity $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$:

$$\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)^2 = \sum_{k,\ell} (a_k - c_\ell)^2 \text{tr}[\varrho A_k C_\ell]. \quad (13.35)$$

This scheme constitutes a *direct test* [8] of the value comparison error, and it is clear that it is restricted to the case of approximators which commute with the target observable. $\triangleleft$

We finally note a connection of the value comparison error with a concept of measurement noise. The noise of linear amplifiers is defined as the “root mean square” difference between operators A_{in} , A_{out} representing the amplifier input and output signals, respectively (see, e.g., [9–12]). Its use in the modelling of quantum measurement error can be traced to the work of Arthurs and Kelly [13], which was elaborated further by Arthurs and Goodman [14].

Thus, in a measurement scheme $\mathcal{M}$, one may associate the sharp *target* observable $\mathbf{A} = \mathbf{E}^A$ to be approximated with the input, with the operator $A_{\text{in}} = A \otimes I$, and the sharp pointer observable $\mathbf{Z} = \mathbf{E}^Z$ with the output via the operator $A_{\text{out}} = U^*(I \otimes Z)U$, where U is the unitary measurement coupling of $\mathcal{M}$. Let σ denote the initial probe state. Then for any object state ϱ one could consider the following as a state-dependent error measure, modeled after the classic rms deviation concept³:

$$\varepsilon_{\text{mn}}(\mathbf{Z}, \mathbf{A}; \varrho) = (\text{tr}[\varrho \otimes \sigma (A_{\text{out}} - A_{\text{in}})^2])^{1/2}. \quad (13.36)$$

We will refer to this quantity as the *measurement noise*.

The above definition of measurement noise can readily be recast in terms of the measured observable $\mathbf{C}$ and the observable $\mathbf{A}$ to be approximated (e.g., [15]):

$$\varepsilon_{\text{mn}}(\mathbf{Z}, \mathbf{A}; \varrho)^2 \equiv \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)^2 = \text{tr}[\varrho(\mathbf{C}[2] - \mathbf{C}[1]^2)] + \text{tr}[\varrho(\mathbf{C}[1] - \mathbf{A})^2]. \quad (13.37)$$

For $\varrho = P[\varphi]$, we will also use the notation $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varphi)$.

Comparing Eqs. (13.37) and (13.32), we see that the measurement noise $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)$ and the value deviation $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$ coincide on the domain of definition of the latter, that is, for commuting pairs of observables $\mathbf{A}, \mathbf{C}$ (with sharp $\mathbf{A}$). The expressions (13.36), (13.37) are formally well-defined independently of this commutativity assumption, so that one could be tempted to regard the definition of $\varepsilon_{\text{mn}}(\mathbf{Z}, \mathbf{A}; \varrho)$ as an extension of that of $\Delta_{\text{vd}}(\mathbf{C}, \mathbf{A}; \varrho)$. However, as shown in [16], the quantity $\varepsilon_{\text{mn}}(\mathbf{Z}, \mathbf{A}; \varrho) = \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)$ is of limited applicability; in fact, it loses its operational meaning as an error measure if applied to approximating observables $\mathbf{C}$ that do not commute with the target observable $\mathbf{A}$.

³It may happen that the measurement noise operator $A_{\text{out}} - A_{\text{in}}$ is unbounded; in this case appropriate care has to be taken with the specification of the states $\varrho \otimes \sigma$ for which the expectation values are well defined.

While the measurement noise is not reliable as an error measure on individual states, it does yield an indicator of the difference between observables if its values for sufficiently many states are taken into account. In fact, from the definition it is clear that $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varphi) = 0$ for all φ exactly when $\mathbf{C}$ is sharp and $\mathbf{C}[1] = \mathbf{A}$, that is, $\mathbf{C} = \mathbf{E}^{\mathbf{A}}$. This suggests the definition of the *global measurement noise* of an observable $\mathbf{C}$ relative to $\mathbf{A}$ as [17, 18]

$$\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}) = \sup_{\varrho} \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho), \quad (13.38)$$

where the supremum is taken over all states ϱ for which $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)$ is well defined and finite. We will say that $\mathbf{C}$ is a *finite noise approximation* to $\mathbf{A}$ if $\mathbf{C}$ has finite global measurement noise relative to $\mathbf{A}$. This will be the case whenever the *bias operator* $\mathbf{C}[1] - \mathbf{A}$ and the intrinsic noise operator $\mathbf{C}[2] - \mathbf{C}[1]^2$ appearing in (13.37) are bounded on a common dense domain.

13.5 Connections

We show that approximations in the sense of finite α -deviations or finite global measurement noise also have finite error bars.

Proposition 13.6 *Any observable $\mathbf{C}$ which is an approximation of a sharp observable $\mathbf{A}$ in the sense that $\Delta_{\alpha}(\mathbf{C}, \mathbf{A}) < \infty$ (for some $\alpha \in [1, \infty)$) is an approximation to $\mathbf{A}$ also in the sense of finite error bars. In that case the α -distance gives an upper bound for the error bar width: for any $\varepsilon \in (0, 1)$,*

$$\mathcal{W}_{\varepsilon}(\mathbf{C}, \mathbf{A}) \leq \frac{2}{\varepsilon^{1/\alpha}} \Delta_{\alpha}(\mathbf{C}, \mathbf{A}). \quad (13.39)$$

Proof The proof is a straightforward adaptation of the proof for the case $\alpha = 1$ given in [5, Proposition 5].

Using the definition of $\Delta_{\alpha}(\mathbf{C}_{\varrho}, \mathbf{A}_{\varrho})$ and Eq. (13.11), we are given that

$$\text{tr}[\varrho \mathbf{C}[f]] - \text{tr}[\varrho \mathbf{A}[g]] \leq \Delta_{\alpha}(\mathbf{C}, \mathbf{A})^{\alpha} = c^{\alpha}, \quad (13.40)$$

which holds for all $\varrho \in S$ and all functions f, g satisfying the constraint

$$f(y) - g(x) \leq |x - y|^{\alpha}, \quad x, y \in \mathbb{R}. \quad (13.41)$$

Let $\varepsilon \in (0, 1)$ and $\delta > 0$ be given. Put $w = \delta + 2n$, with $n \in \mathbb{N}$, $n^{\alpha} \geq c^{\alpha}/\varepsilon$. Consider an interval $J_{q;\delta}$ and a state ϱ with $\mathbf{A}_{\varrho}(J_{q;\delta}) = 1$. Define the functions $g_n = f_n \equiv h_n$ via

$$h_n(x) = \begin{cases} n^\alpha & \text{if } |x - q| \leq \delta/2; \\ (n + \delta/2 - |x - q|)^\alpha & \text{if } \delta/2 < |x - q| \leq \delta/2 + n; \\ 0 & \text{if } \delta/2 + n < |x - q|. \end{cases}$$

It is not hard to verify that f_n, g_n satisfy (13.41). Condition (13.40) for $f_n = g_n = h_n$ entails for $r_n = h_n/n^\alpha$ that $\text{tr}[\varrho \mathbf{C}[r_n]] - \text{tr}[\varrho \mathbf{A}[r_n]] \leq c^\alpha/n^\alpha$. We then have $\chi_{J_{q;\delta}} \leq r_n \leq \chi_{J_{q;w}}$.

Now $\mathbf{A}_\varrho(J_{q;\delta}) = 1$ implies $\text{tr}[\varrho \mathbf{A}[r_n]] = 1$, and so, using the assumption $n^\alpha \geq c^\alpha/\varepsilon$, we obtain

$$\text{tr}[\varrho \mathbf{C}(J_{q;w})] \geq \text{tr}[\varrho \mathbf{C}[r_n]] \geq \text{tr}[\varrho \mathbf{A}[r_n]] - c^\alpha/n^\alpha \geq 1 - \varepsilon.$$

To prove (13.39), we note that on putting $w = \delta + 2c/(\varepsilon^{1/\alpha})$, one still obtains $\text{tr}[\varrho \mathbf{C}(J_{q;w})] \geq 1 - \varepsilon$. This yields $\mathcal{W}_{\varepsilon,\delta}(\mathbf{C}, \mathbf{A}) \leq \delta + 2\Delta_\alpha(\mathbf{C}, \mathbf{A})/\varepsilon^{1/\alpha}$, and on letting δ approach 0, then (13.39) follows. $\square$

Proposition 13.7 *Any observable $\mathbf{C}$ that satisfies the condition of finite global measurement noise relative to a sharp observable $\mathbf{A}$, $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}) < \infty$, is an approximation to $\mathbf{A}$ in the sense of finite error bars. Then the measurement noise gives an upper bound for the error bar width: for any $\varepsilon \in (0, 1)$,*

$$\mathcal{W}_\varepsilon(\mathbf{C}, \mathbf{A}) \leq 2\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}) \left(1 + \sqrt{\frac{2}{\varepsilon}}\right). \quad (13.42)$$

The proof is given in the Appendix. We remark that it is an open question whether finite global measurement noise implies finite α -distance or vice versa.

13.6 Unsharpness

Two noncommuting sharp observables $\mathbf{A}, \mathbf{B}$ are incompatible. It is conceivable that they could be approximated by two commuting sharp observables $\mathbf{C}, \mathbf{D}$, which are jointly measurable. However, it is to be expected that better joint approximations could be achieved if commutativity is traded for unsharpness. Here we review various ways to make this idea precise. We consider two concepts: the intrinsic noise that was already introduced, and the resolution width.

Intrinsic Noise Operator and Inherent Noise

In Proposition 12.1 we have encountered a necessary trade-off in the degrees of unsharpness, measured in terms of the intrinsic noise, required for two noncommuting observables to be jointly measurable. However, the inequality does not represent the connection between (non-)commutativity and unsharpness in a faithful way. For instance, a pair of trivial observables is commutative and jointly measurable, but also maximally unsharp.

Following Polterovich [19, 20], we show that it is possible to isolate the *inherent noise* contribution contained in an observable due to its degree of noncommutativity. This leads to a strengthening of the noise trade-off inequality in such a way that the inherent noise is bounded from below by the noncommutativity and it vanishes exactly when the observable is commutative.

As noted in Sect. 8.3, every commutative semispectral measure can be expressed in the form of a smearing of a spectral measure with a Markov kernel, Eq. (8.4). This result can be interpreted by saying that among all observables, any unsharpness contained in commutative ones results from the smearing applied to a sharp observable. This provides a key to extracting a measure of the unsharpness associated with the noncommutativity of an unsharp observable. To start with, we introduce the concepts of overall noise and overall degree of noncommutativity of an observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$.

Let $K(\Omega)$ denote the set of measurable functions $h : \Omega \rightarrow [-1, 1]$. For $h \in K(\Omega)$, consider the induced (real) observable $E^h = E \circ h^{-1}$.

The *overall noise* is defined as

$$\mathcal{N}(E) = \sup_{h \in K(\Omega)} \|N(E^h)\|,$$

where $N(E^h) = E^h[2] - E^h[1]^2 = L(h^2, E) - L(h, E)^2$ is the noise operator of E^h . Clearly, $0 \leq \mathcal{N}(E^h) \leq 1$. We recall (Theorem 8.5) that $N(E^h) = 0$ exactly when E^h is sharp. Thus, if $\mathcal{N}(E) = 0$, then each E^h , $h \in K(\Omega)$ is sharp, and thus each $E(X)$, $X \in \mathcal{A}$, is a projection, that is, E is sharp. Conversely, if E is sharp, then $\mathcal{N}(E) = 0$.

The *overall degree of noncommutativity* of E is defined in terms of the commutators of the bounded selfadjoint operators $E^h[1] = L(h, E)$ as follows:

$$\nu(E) = \sup_{h, k \in K(\Omega)} \|E^h[1]E^k[1] - E^k[1]E^h[1]\|.$$

Now $0 \leq \nu(E) \leq 2$. Since $h \mapsto L(h, E)$ is bounded linear map and $\chi_X \in K(\Omega)$, $X \in \mathcal{A}$, we have by Proposition 4.12 that $\nu(E) = 0$ if and only if E is commutative.

Since $0 \leq N(E^h) \leq I$ for all $h \in K(\Omega)$, Proposition 12.1 implies that the overall noise is bounded from below by half of the overall degree of noncommutativity,

$$\mathcal{N}(E) \geq \frac{1}{2} \nu(E).$$

Let E on $(\Omega, \mathcal{A})$ be a smearing of another observable F on $(\Omega', \mathcal{A}')$ by a Markov kernel p , that is,

$$E(X) = \int_{\Omega'} p(X, \omega') dF(\omega'), \quad X \in \mathcal{A}.$$

We recall that if $\mathbf{F}$ is an extremal element of $\mathcal{O}(\Omega', \mathcal{A}', \mathcal{H})$ then p is unique. Especially, this holds for spectral measures. Given p , one can introduce the associated *smearing operator* as

$$\Pi : K(\Omega) \rightarrow K(\Omega'), \quad (\Pi h)(\omega') = \int_{\Omega} h(\omega) p(d\omega, \omega'), \quad \omega' \in \Omega'.$$

Finally, the *inherent noise* of $\mathbf{E}$ is defined as the infimum of overall noise among all smearings $(\mathbf{F}, p)$ of $\mathbf{E}$:

$$\mathcal{N}_i(\mathbf{E}) = \inf_{(\mathbf{F}, \Pi)} \sup_{h \in K(\Omega)} \|N(\mathbf{F}^{\Pi h})\|. \quad (13.43)$$

For any $h \in K(\Omega)$, $L(h, \mathbf{E}) = L(\Pi h, \mathbf{F})$ whereas $L(h^2, \mathbf{E}) \geq L((\Pi h)^2, \mathbf{F})$. This implies that

$$\mathcal{N}_i(\mathbf{E}) \leq \mathcal{N}(\mathbf{E}).$$

It follows that if $\mathbf{E}$ is sharp, then $\mathcal{N}_i(\mathbf{E}) = 0$.

We leave it as an exercise to show that both $\mathcal{N}_i$ and ν are monotonic under smearing, that is, if $\mathbf{E}$ is a smearing of $\mathbf{F}$ by p , then $\mathcal{N}_i(\mathbf{E}) \leq \mathcal{N}_i(\mathbf{F})$ and $\nu(\mathbf{E}) \leq \nu(\mathbf{F})$. Similarly to the overall noise, the inherent noise is also bounded from below by the noncommutativity.

Theorem 13.1 *For any observable $\mathbf{E}$,*

$$\mathcal{N}_i(\mathbf{E}) \geq \frac{1}{2}\nu(\mathbf{E}). \quad (13.44)$$

Proof This is again a consequence of the intrinsic noise relation (12.12). We leave the details as an exercise. $\square$

Since every commutative observable is a smearing of some sharp observable this entails that $\mathcal{N}_i(\mathbf{E}) = 0$ whenever $\nu(\mathbf{E}) = 0$. The above theorem gives the converse implication: vanishing inherent noise is characteristic of commutative observables.

The relationship between inherent unsharpness and noncommutativity readily extends to an unsharpness relation for joint measurements.

For any two observables $\mathbf{E}_1, \mathbf{E}_2$ with value spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$, their degree of noncommutativity is defined as

$$\nu(\mathbf{E}_1, \mathbf{E}_2) = \sup_{(h, k) \in K(\Omega_1) \times K(\Omega_2)} \|\mathbf{E}_1^h[1]\mathbf{E}_2^k[1] - \mathbf{E}_2^k[1]\mathbf{E}_1^h[1]\|.$$

One has $\nu(\mathbf{E}_1, \mathbf{E}_2) = 0$ exactly when $\mathbf{E}_1$ and $\mathbf{E}_2$ are mutually commutative.

Consider two observables $\mathbf{E}_1, \mathbf{E}_2$ whose value spaces have the property (D). By Theorem 11.1, $\mathbf{E}_1$ and $\mathbf{E}_2$ are jointly measurable exactly when there is an observable $\mathbf{E} \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ and Markov kernels $p_i : \mathcal{A}_i \times \Omega \rightarrow [0, 1]$ such that $\mathbf{E}_i$ is a smearing of $\mathbf{E}$ by p_i , $i = 1, 2$. Then, $\nu(\mathbf{E}_1) \leq \nu(\mathbf{E})$ and $\nu(\mathbf{E}_2) \leq \nu(\mathbf{E})$. Also, since

$L(h, E_i) = L(\Pi_i h, E)$, one now gets $\nu(E) \geq \nu(E_1, E_2)$. With these observations the following unsharpness relation for joint measurements has been established.

Proposition 13.8 *Let E_1, E_2 be jointly measurable observables, that is, smearings of an observable E with the Markov kernels p_1, p_2 . Then*

$$\mathcal{N}_i(E) \geq \frac{1}{2} \max\{\nu(E_1), \nu(E_2), \nu(E_1, E_2)\}. \quad (13.45)$$

Resolution Width

The intrinsic noise operator of an observable E , $N(E) = E[2] - E[1]^2$, taken over all states ϱ , is a faithful indicator of whether or not E is a sharp observable. But its expectation $\text{tr}[\varrho N(E)]$ in a state ϱ cannot be determined from the statistics of E alone. Furthermore, the overall and inherent noise $\mathcal{N}_i$ and the measure ν defined in the preceding subsection are designed to characterise unsharpness and noncommutativity in a way that is independent of any choice of scale. We now consider a measure of the intrinsic unsharpness of an observable whose value can be obtained from the statistics of the observable and which is sensitive to the scale of the underlying value space.

For an observable E with support $\text{supp}(E)$ given by $\mathbb{R}$ or a closed interval, the *resolution width* (at confidence level $1 - \varepsilon$) is defined as [21]:

$$\gamma_\varepsilon(E) = \inf\{w > 0 \mid \forall x \in \text{supp}(E) \exists \varrho \in \mathcal{S}(\mathcal{H}) : E_\varrho([x - \frac{w}{2}, x + \frac{w}{2}]) \geq 1 - \varepsilon\}.$$

This is the smallest interval size which is such that wherever the interval is placed, there are states for which the value of E is localised within the interval with probability at least equal to $1 - \varepsilon$.

We note that the resolution width scales with the metric of the underlying value space.

For a sharp observable E the resolution width is $\gamma_\varepsilon(E) = 0$ for all $\varepsilon \in (0, 1)$. It is worth noting that vanishing resolution width does not require the observable to be sharp: in fact, the resolution width is zero for any observable whose nonzero effects have norm 1; an example is given by the canonical phase observable (see Theorem 8.2).

Proposition 13.9 *Let A be a sharp observable with support $\mathbb{R}$ and F an ε -approximation to A . The error bar width of F relative to A is never smaller than the intrinsic resolution width of F :*

$$\mathcal{W}_\varepsilon(F, A) \geq \gamma_\varepsilon(F). \quad (13.46)$$

In particular, the resolution width is necessarily finite.

Proof If $\mathcal{W}_\varepsilon(F, A) = \infty$, (13.46) is trivially satisfied. Assume that $\mathcal{W}_\varepsilon(F, A)$ is finite. There is a $\delta_0 > 0$ such that $\mathcal{W}_{\varepsilon, \delta_0}(F, A) < \infty$. Since $\mathcal{W}_{\varepsilon, \delta}(F, A)$ is an increasing function of δ , then $\mathcal{W}_{\varepsilon, \delta}(F, A) < \infty$ for $\delta \leq \delta_0$. Let $w \geq \mathcal{W}_{\varepsilon, \delta}(F, A)$ for some δ ,

$0 < \delta \leq \delta_0$. Thus for all $x \in \mathbb{R}$ and all ϱ with $\mathbf{A}_\varrho(J_{x;\delta}) = 1$ we have $\mathbf{F}_\varrho(J_{x;w}) \geq 1 - \varepsilon$. This entails (given that $\text{supp}(\mathbf{A}) = \mathbb{R}$) that for all $x \in \mathbb{R}$ there is some ϱ such that $\mathbf{F}_\varrho(J_{x;w}) \geq 1 - \varepsilon$. Hence $w \geq \gamma_\varepsilon(\mathbf{F})$, and so $\mathcal{W}_{\varepsilon,\delta}(\mathbf{F}, \mathbf{A}) \geq \gamma_\varepsilon(\mathbf{F})$ for all $\delta > 0$, from which (13.46) follows. $\square$

Proposition 13.10 *Let $\mathbf{A}$ be a sharp observable with support $\mathbb{R}$. For the resolution width of the convolution $\mu * \mathbf{A}$ of $\mathbf{A}$ with a probability measure μ one has*

$$\gamma_\varepsilon(\mu * \mathbf{A}) \geq W_\varepsilon(\mu). \quad (13.47)$$

Equality holds if μ is absolutely continuous with respect to the Lebesgue measure.

Proof Let $w \geq \gamma_\varepsilon(\mu * \mathbf{A})$. Then, for any $x \in \mathbb{R}$ there is a state ϱ such that

$$(\mu * \mathbf{A})_\varrho(J_{x;w}) = \int \mu(J_{x-y;w}) \mathbf{A}_\varrho(dy) \geq 1 - \varepsilon.$$

Fix any x and take a suitable ϱ such that this inequality holds; then for some y one must have $\mu(J_{x-y;w}) \geq 1 - \varepsilon$. Therefore, $w \geq W_\varepsilon(\mu)$, and the inequality follows.

The remaining statement is proven in [21]. $\square$

Proposition 13.11 *Let $\mu * \mathbf{Q}$, $\nu * \mathbf{P}$ be margins of a covariant phase space observable $\mathbf{G}^T$. Then for $\varepsilon_1, \varepsilon_2 \in [0, 1]$ with $\varepsilon_1 + \varepsilon_2 \leq 1$, their resolution widths obey the inequality*

$$\gamma_{\varepsilon_1}(\mu * \mathbf{Q}) \gamma_{\varepsilon_2}(\nu * \mathbf{P}) \geq 2\pi\hbar (1 - \varepsilon_1 - \varepsilon_2). \quad (13.48)$$

Proof This is a consequence of the inequalities (13.47) and (12.15). $\square$

13.7 Finite Outcome Observables

In contrast to preparation uncertainty, it seems less straightforward to formulate measurement uncertainty relations in a generic way. Such relations would have to hold for any approximate joint measurement of some observables $\mathbf{A}, \mathbf{B}$, and they would take the form of inequalities for the errors where the bounds are governed by the degree of incompatibility of $\mathbf{A}, \mathbf{B}$. Here we consider some prototype examples of such relations obtained for finite-outcome observables in finite-dimensional Hilbert spaces [22]. The case of 2-valued observables in $\mathbb{C}^2$ is addressed extensively in Chap. 14. A first study of the error (or measurement uncertainty) region for finite-outcome observables appeared at the time of finalising this book [23].

Let Ω be a finite space with N elements and let $M(\Omega)_1^+$ denote the set of probability measures on 2^Ω . $M(\Omega)_1^+$ is a convex subset of an N -dimensional vector space, namely, the convex hull of the N point measures δ_ω , $\omega \in \Omega$.

Error Measures Induced by the Total Variation and Supremum Norms

We let $\mathcal{O}(\Omega)$ denote the set of observables with the value space $(\Omega, 2^\Omega)$ and acting on a Hilbert space $\mathcal{H}$. Suppose d is a metric on $M(\Omega)_1^+$. This induces a metric D_d on $\mathcal{O}(\Omega)$ via

$$D_d(\mathbf{E}, \mathbf{F}) = \sup_{\varrho \in \mathcal{S}(\mathcal{H})} d(\mathbf{E}_\varrho, \mathbf{F}_\varrho).$$

Specifically we will use the metrics induced by the total variation norm (modulo a factor $\frac{1}{2}$ which is introduced for convenience) and the sup norm, respectively: for any $\mu, \nu \in M(\Omega)_1^+$,

$$d_1(\mu, \nu) = \frac{1}{2} \sum_{\omega \in \Omega} |\mu(\omega) - \nu(\omega)|, \quad (13.49)$$

$$d_\infty(\mu, \nu) = \max_{\omega \in \Omega} |\mu(\omega) - \nu(\omega)|. \quad (13.50)$$

The associated metrics on $\mathcal{O}(\Omega)$ will be denoted D_1 and D_∞ , respectively. They take the form

$$D_1(\mathbf{E}, \mathbf{F}) = \max_{X \subset \Omega} \|\mathbf{E}(X) - \mathbf{F}(X)\|, \quad (13.51)$$

$$D_\infty(\mathbf{E}, \mathbf{F}) = \max_{\omega \in \Omega} \|E_\omega - F_\omega\|. \quad (13.52)$$

Remark 13.2 Going momentarily beyond the case of finite outcome observables and considering $\Omega \subset \mathbb{R}$, Theorem 6.15 of [3] states that the Wasserstein distance is controlled by a weighted total variation. In particular, if the diameter of Ω is bounded by K , that is, $d(\omega, \omega') \leq K$ for all $\omega, \omega' \in \Omega$, then $\Delta_1(\mu, \nu) \leq K d_1(\mu, \nu)$. A similar result holds for $\alpha = 2$. We leave it as an exercise to check whether the metric d_∞ gives an upper bound for $\alpha = \infty$. $\triangleleft$

We now consider two (sharp or unsharp) observables $\mathbf{A} = \{A_x\}_{x \in \Omega_A}$ (with $n = |\Omega_A|$ outcomes) and $\mathbf{B} = \{B_y\}_{y \in \Omega_B}$ (with $m = |\Omega_B|$ outcomes) and an approximate joint measurement by means of a joint measurement of two compatible observables $\mathbf{C}$ (defined on Ω_A) and $\mathbf{D}$ (defined on Ω_B).

The quality of the approximations of $\mathbf{A}$, $\mathbf{B}$ by means of $\mathbf{C}$, $\mathbf{D}$ will be measured via the metrics defined above. The degree of noncommutativity of $\mathbf{A}$ and $\mathbf{B}$ entails a constraint on the distances $D(\mathbf{A}, \mathbf{C})$ and $D(\mathbf{B}, \mathbf{D})$. The resulting trade-off relation also involves quantities that represent measures of the intrinsic unsharpness of $\mathbf{A}$, $\mathbf{B}$; their definition depends on the choice of distance D . In the case of D_1 and D_∞ the unsharpness measures are defined as

$$V_1(\mathbf{A}) = \max_{X \subset \Omega_A} \|\mathbf{A}(X) - \mathbf{A}(X)^2\|,$$

$$V_\infty(\mathbf{A}) = \max_{\omega \in \Omega_A} \|A_\omega - A_\omega^2\|.$$

Further, we define the overall degree of commutativity of $\mathbf{A}$, $\mathbf{B}$ on Ω_A , Ω_B as follows:

$$\begin{aligned}\nu_1(\mathbf{A}, \mathbf{B}) &= \max_{X \subset \Omega_A, Y \subset \Omega_B} \|\mathbf{A}(X)\mathbf{B}(Y) - \mathbf{B}(Y)\mathbf{A}(X)\|, \\ \nu_\infty(\mathbf{A}, \mathbf{B}) &= \max_{\omega \in \Omega_A, \omega' \in \Omega_B} \|A_\omega B_{\omega'} - B_{\omega'} A_\omega\|,\end{aligned}$$

We quote (without proof) the following theorem [22].

Theorem 13.2 *Let $\mathbf{A}$, $\mathbf{B}$ be discrete observables with finite value spaces Ω_A , Ω_B defined in a finite dimensional Hilbert space, and let $\mathbf{C}$, $\mathbf{D}$ be jointly measurable observables on Ω_A , Ω_B , respectively. Let the triple (D, V, ν) be one of (D_1, V_1, ν_1) and $(D_\infty, V_\infty, \nu_\infty)$. Then the following approximation error trade-off relation holds:*

$$\begin{aligned}2D(\mathbf{A}, \mathbf{C})D(\mathbf{B}, \mathbf{D}) + D(\mathbf{A}, \mathbf{C}) + D(\mathbf{B}, \mathbf{D}) \\ + 2[2D(\mathbf{A}, \mathbf{C}) + V(\mathbf{A})]^{\frac{1}{2}}[2D(\mathbf{B}, \mathbf{D}) + V(\mathbf{B})]^{\frac{1}{2}} \geq \nu(\mathbf{A}, \mathbf{B}).\end{aligned}$$

As an immediate consequence one obtains an unsharpness trade-off relation.

Corollary 13.2 *Let $\mathbf{A}$, $\mathbf{B}$ be two jointly measurable observables on the finite value spaces Ω_A , Ω_B , respectively defined in a finite dimensional Hilbert space. Then*

$$V_i(\mathbf{A})^{\frac{1}{2}} V_i(\mathbf{B})^{\frac{1}{2}} \geq \frac{1}{2} \nu_i(\mathbf{A}, \mathbf{B}) \quad (13.53)$$

for the cases $i = 1$ and $i = \infty$.

The proof is straightforward; one chooses $\mathbf{C} = \mathbf{A}$ and $\mathbf{D} = \mathbf{B}$ in the above theorem. Thus, the joint measurability of two noncommuting observables requires that they are sufficiently unsharp.

A further consequence is obtained by choosing $\mathbf{B} = \mathbf{A}$.

Corollary 13.3 *For an observable $\mathbf{A}$ with finite value space Ω_A defined in a finite dimensional Hilbert space one has*

$$V_i(\mathbf{A}) \geq \frac{1}{2} \nu_i(\mathbf{A}, \mathbf{A}), \quad (13.54)$$

for the cases $i = 1$ and $i = \infty$.

The intrinsic unsharpness of an observable $\mathbf{A}$ is bounded below by its degree of noncommutativity. Note the similarity in spirit between these relations and those of Sect. 13.6.

Error Bar Width for Finite Outcome Spaces

The probabilistic distances D_1 , D_∞ (Eqs. (13.51), (13.52)) are not optimally suitable to quantify the quality of approximate measurements of continuous observables such as position. To illustrate this point, consider an attempt to approximate the observable $\mathbf{Q}$ by a slightly translated version $\mathbf{Q}_x = \mathbf{Q} \circ \tau_x^{-1}$, where $\tau_x(y) = y - x$. Consider a

vector state φ with $\mathbf{Q}$ -distribution $\mathbf{Q}_\varphi$ supported in an interval J of width $|x|/3$. Then the $\mathbf{Q}_x$ -distribution $(\mathbf{Q}_x)_\varphi$ is supported in $J - x$, which is disjoint to J . Hence the total variation norm of the difference of these distributions assumes the maximal value 2, and therefore $D_1(\mathbf{Q}, \mathbf{Q}_x) = 2$, which is also maximal, independently of the value of $x \neq 0$. This shows that the error in approximating $\mathbf{Q}$ by $\mathbf{Q}_x$ jumps discontinuously to the maximal value, for arbitrarily small shift distances x , while intuitively the approximation should be better, the smaller x .

By contrast, in the case of observables on a finite value set equipped with the discrete metric, one can find a direct connection between the probabilistic distance and distances based on this discrete metric. This will be illustrated here with the concept of error bar width. A connection between the Wasserstein 2-distance and D_1 will be exhibited in Chap. 14 in the case of 2-outcome qubit observables.

From its inception in Heisenberg's 1927 paper [1], the joint measurement error relation was seen as an expression of the fact that the possibilities of measurement cannot exceed the possibilities of preparation. Starting with his famous Como lecture [24], Niels Bohr consistently emphasised the consistency of the possibilities of observation (measurement) and definition (preparation).

We have seen illustrations of this principle in the case of covariant phase space measurements, where the error relations (13.8) and (13.21) in terms of Wasserstein 2-distances and error bar widths were consequences of the corresponding preparation relations (12.14) and (12.15) for Wasserstein 2-deviations and overall widths, respectively. This connection was shown to be generic by Miyadera in 2011 [25] for finite-outcome measurements in finite dimensional Hilbert spaces.

Let (Ω_A, d) be a finite metric space, and let $O_d(x, w)$ denote the closed ball with centre $x \in \Omega$ and radius $\frac{w}{2}$. Let $\mathbf{A}$ and $\mathbf{C}$ be, respectively, a spectral and semispectral measure in the Hilbert space $\mathcal{H}$ with the same outcome space (Ω_A, d) . For such finite observables, the *error bar width* is

$$\mathcal{W}_\varepsilon(\mathbf{C}, \mathbf{A}) = \inf \left\{ w > 0 \mid \forall x \in \Omega_A \forall \varrho \in \mathcal{S}(\mathcal{H}) : \mathbf{A}_\varrho(\{x\}) = 1 \Rightarrow \mathbf{C}_\varrho(O_d(x, w)) \geq 1 - \varepsilon \right\}. \quad (13.55)$$

The definition of the Wasserstein 1-distance can be applied without modification. Then the following result holds [25].

Theorem 13.3 *Let $\mathbf{A}, \mathbf{B}$ be sharp observables defined on a finite dimensional Hilbert space $\mathcal{H}$, with outcome spaces given by finite metric spaces (Ω_A, d_A) and (Ω_B, d_B) . Let $\mathbf{C}, \mathbf{D}$ be jointly measurable observables on Ω_A and Ω_B , respectively. Then for any $\varepsilon_1, \varepsilon_2 \in [0, 1]$ with $\varepsilon_1 + \varepsilon_2 \leq 1$ there exists a state $\varrho \in \mathcal{S}(\mathcal{H})$ such that*

$$\begin{aligned} \mathcal{W}_{\varepsilon_1}(\mathbf{C}, \mathbf{A}) &\geq W_{\varepsilon_1 + \varepsilon_2}(\mathbf{A}_\varrho), & \mathcal{W}_{\varepsilon_2}(\mathbf{D}, \mathbf{B}) &\geq W_{\varepsilon_1 + \varepsilon_2}(\mathbf{B}_\varrho); \\ \Delta_1(\mathbf{C}, \mathbf{A}) &\geq \frac{\varepsilon_1}{2} W_{\varepsilon_1 + \varepsilon_2}(\mathbf{A}_\varrho), & \Delta_1(\mathbf{D}, \mathbf{B}) &\geq \frac{\varepsilon_2}{2} W_{\varepsilon_1 + \varepsilon_2}(\mathbf{B}_\varrho). \end{aligned}$$

The proof is based on a remote state preparation method known from quantum information theory. Together with inequality (12.15), it is seen that the above inequalities

entail an error relation for joint measurements of two finite observables if the parameters $\varepsilon_1, \varepsilon_2$ are small enough (such that $\varepsilon_1 + \varepsilon_2 \leq \frac{1}{2}$). A stronger result was also proven in [25].

Theorem 13.4 *Let $\mathbf{A} = \{A_x\}_{x \in \Omega_A}$, $\mathbf{B} = \{B_y\}_{y \in \Omega_B}$ be sharp observables in a finite dimensional Hilbert space $\mathcal{H}$ defined on finite metric spaces (Ω_A, d_A) , (Ω_B, d_B) . Let $\mathbf{C}, \mathbf{D}$ be jointly measurable observables on Ω_A and Ω_B , respectively. Then for $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 \leq 1$,*

$$\begin{aligned} \max_{x \in \Omega_A} |O_{d_A}(x, \mathcal{W}_{\varepsilon_1}(\mathbf{C}, \mathbf{A}))| \cdot \max_{y \in \Omega_B} |O_{d_B}(y, \mathcal{W}_{\varepsilon_2}(\mathbf{D}, \mathbf{B}))| \\ \geq \frac{(1 - \varepsilon_1 - \varepsilon_2)^2}{\max_{x \in \Omega_A, y \in \Omega_B} \text{tr}[A_x B_y]}, \end{aligned}$$

where $|X|$ denotes the counting measure of the finite set $|X|$.

The following example shows that uncertainty relations similar to (12.15), (13.21) hold for position and momentum on finite spaces [25].

Example 13.5 The position and momentum of a particle on $\mathbb{Z}^N$ are represented by orthonormal basis systems of eigenvectors in $\mathcal{H} = \mathbb{C}^N$, $\{\varphi_n \mid n = 0, 1, \dots, N-1\}$, $\{\psi_k \mid k = 0, 1, \dots, N-1\}$, where

$$\psi_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-i \frac{2\pi nk}{N}} \varphi_n. \quad (13.56)$$

Let $\Omega_Q = \Omega_P = \mathbb{Z}^N \simeq \{0, 1, \dots, N-1\}$ and define sharp observables $\mathbf{Q} = \{Q_n = P[\varphi_n]\}_{n \in \Omega_Q}$ and $\mathbf{P} = \{P_k = P[\psi_k]\}_{k \in \Omega_P}$. A metric on these spaces is given by

$$d(m, n) = \min \left\{ \frac{2\pi}{N} (m - n + Nt) \mid t \in \mathbb{Z}, m - n + Nt \geq 0 \right\}. \quad (13.57)$$

Noting that $|O_d(m, w)| \leq \frac{wN}{2\pi} + 1$ for any $m \in \mathbb{Z}_N$ and $w \geq 0$, it follows that for any $\varrho \in \mathcal{S}(\mathcal{H})$ and $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 \leq 1$,

$$\left(W_{\varepsilon_1}(\mathbf{Q}_\varrho) + \frac{1}{N} \right) \cdot \left(W_{\varepsilon_2}(\mathbf{P}_\varrho) + \frac{1}{N} \right) \geq 2\pi\hbar(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (13.58)$$

Further, for any jointly measurable observables $\mathbf{Q}'$ on Ω_Q and $\mathbf{P}'$ on Ω_P , one has

$$\left(W_{\varepsilon_1}(\mathbf{Q}', \mathbf{Q}) + \frac{1}{N} \right) \cdot \left(W_{\varepsilon_2}(\mathbf{P}', \mathbf{P}) + \frac{1}{N} \right) \geq 2\pi\hbar(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (13.59)$$

◁

13.8 Appendix

Proof of Proposition 13.7

Since $\mathbf{A}$ is sharp we will write A for $\mathbf{A}[1]$ and use the facts that $A^2 = \mathbf{A}[1]^2 = \mathbf{A}[2]$ and $\Delta(\mathbf{A}, \varrho) = \Delta(A, \varrho)$.

We begin by rewriting the definition of the measurement noise for general states ϱ , denoted $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)$, and expressing the condition of bounded measurement noise: for all ϱ and $c = \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}) < \infty$,

$$\begin{aligned} \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)^2 &= \text{tr}[\varrho(\mathbf{C}[1] - A)^2] + \text{tr}[\varrho(\mathbf{C}[2] - \mathbf{C}[1]^2)] \\ &= \text{tr}[\varrho(\mathbf{C}[1] - A)^2] + \Delta(\mathbf{C}, \varrho)^2 - \Delta(\mathbf{C}[1], \varrho)^2 \leq c^2. \end{aligned}$$

(This follows readily from the corresponding condition stipulated for all vector states.) The first term can be estimated as follows: using the inequality

$$\begin{aligned} |\text{cov}_\varrho(\mathbf{C}[1], A)| &= \frac{1}{2} |\text{tr}[\varrho \mathbf{C}[1] A] + \text{tr}[\varrho A \mathbf{C}[1]] - 2\text{tr}[\varrho \mathbf{C}[1]]\text{tr}[\varrho A]| \\ &\leq \Delta(\mathbf{C}[1], \varrho)\Delta(A, \varrho), \end{aligned}$$

we see that

$$\begin{aligned} \text{tr}[\varrho(\mathbf{C}[1] - A)^2] &= \Delta(\mathbf{C}[1] - A, \varrho)^2 + (\text{tr}[\varrho(\mathbf{C}[1] - A)])^2 \\ &= \Delta(\mathbf{C}[1], \varrho)^2 + \Delta(A, \varrho)^2 - 2\text{cov}_\varrho(\mathbf{C}[1], A) + (\text{tr}[\varrho(\mathbf{C}[1] - A)])^2 \\ &\geq (\Delta(\mathbf{C}[1], \varrho) - \Delta(A, \varrho))^2 + (\text{tr}[\varrho(\mathbf{C}[1] - A)])^2. \end{aligned}$$

The boundedness of $\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)$ then gives:

$$\begin{aligned} (\Delta(\mathbf{A}[1], \varrho) - \Delta(A, \varrho))^2 + (\text{tr}[\varrho \mathbf{C}[1]] - \text{tr}[\varrho A])^2 \\ + (\Delta(\mathbf{C}, \varrho)^2 - \Delta(\mathbf{C}[1], \varrho)^2) \leq \varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}; \varrho)^2 \leq c^2. \end{aligned}$$

Each of the three bracketed terms is nonnegative and hence bounded above by c^2 . This implies:

$$\Delta(\mathbf{C}[1], \varrho)^2 - c^2 \leq \Delta(\mathbf{C}, \varrho)^2 \leq \Delta(\mathbf{C}[1], \varrho)^2 + c^2, \quad (\#1)$$

$$\Delta(A, \varrho) - c \leq \Delta(\mathbf{C}[1], \varrho) \leq \Delta(A, \varrho) + c, \quad (\#2)$$

$$\text{tr}[\varrho A] - c \leq \text{tr}[\varrho \mathbf{C}[1]] \leq \text{tr}[\varrho A] + c; \quad (\#3)$$

the first two inequalities taken together yield:

$$\Delta(\mathbf{C}, \varrho)^2 \leq (\Delta(A, \varrho) + c)^2 + c^2. \quad (\#4)$$

Now observe that the variance on the l.h.s. is the variance of the distribution $\mathbf{C}_\varrho$. We use the following variant of Chebyshev's inequality, valid for any $w > 0$:

$$\begin{aligned} \Delta(\mathbf{C}_\varrho)^2 &= \int (x - \mathbf{C}_\varrho[1])^2 \mathbf{C}_\varrho(dx) \\ &\geq \begin{cases} \int_{\mathbb{R} \setminus J_{q;w}} (x - \mathbf{C}_\varrho[1])^2 \mathbf{C}_\varrho(dx) \geq \left(\frac{w}{2} - |\mathbf{C}_\varrho[1] - q|\right)^2 (1 - \mathbf{C}_\varrho(J_{q;w})) \\ \text{if } \mathbf{C}_\varrho[1] \in J_{q;w}; \\ \int_{J_{q;w}} (x - \mathbf{C}_\varrho[1])^2 \mathbf{C}_\varrho(dx) \geq \left(\frac{w}{2} - |\mathbf{C}_\varrho[1] - q|\right)^2 \mathbf{C}_\varrho(J_{q;w}) \\ \text{if } \mathbf{C}_\varrho[1] \notin J_{q;w}. \end{cases} \end{aligned}$$

We will only be using cases of large w where $\mathbf{C}_\varrho[1] \in J_{q;w}$ so that we obtain:

$$\left(\frac{w}{2} - |\mathbf{C}_\varrho[1] - q|\right)^2 (1 - \mathbf{C}_\varrho(J_{q;w})) \leq \Delta(\mathbf{C}_\varrho)^2. \quad (\#5)$$

Combining (#4) and (#5) yields:

$$\left(\frac{w}{2} - |\mathbf{C}_\varrho[1] - q|\right)^2 (1 - \mathbf{C}_\varrho(J_{q;w})) \leq (\Delta(A, \varrho) + c)^2 + c^2. \quad (\#6)$$

We will only use this in the case of states ϱ for which $\mathbf{C}_{\varrho^A}(J_{q;\delta}) = 1$. In this case we have $\Delta(A, \varrho) \leq \delta$ and $|\text{tr}[\varrho A] - q| \leq \delta$, and using (#3) we also obtain:

$$|\mathbf{C}_\varrho[1] - q| \leq |\mathbf{C}_\varrho[1] - \text{tr}[\varrho A]| + |\text{tr}[\varrho A] - q| \leq c + \delta.$$

We will also use only large (finite) w so that we can assume

$$\frac{w}{2} - |\mathbf{C}_\varrho[1] - q| \geq \frac{w}{2} - (\delta + c) > 0.$$

Note that this entails, in particular, that $\mathbf{C}_\varrho[1] \in J_{q;w}$, so that the use of (#6) is justified. Under these conditions (#6) entails

$$\left(\frac{w}{2} - (\delta + c)\right)^2 (1 - \mathbf{C}_\varrho(J_{q;w})) \leq (\delta + c)^2 + c^2. \quad (\#7)$$

Now, for any ε one can choose w large enough such that

$$\left(\frac{w}{2} - (\delta + c)\right)^2 = \frac{(\delta + c)^2 + c^2}{\varepsilon}$$

Then (#7) implies that $1 - \mathbf{C}_\varrho(J_{q;w}) \leq \varepsilon$. Moreover, since $\mathcal{W}_{\varepsilon,\delta}(\mathbf{C}, \mathbf{A}) \leq w$, we also have

$$\mathcal{W}_{\varepsilon,\delta}(\mathbf{C}, \mathbf{A}) \leq \frac{2}{\sqrt{\varepsilon}} \sqrt{(\delta + c)^2 + c^2} + 2(\delta + c),$$

which in the limit $\delta \rightarrow 0$ yields

$$\mathcal{W}_\varepsilon(\mathbf{C}, \mathbf{A}) \leq \left(1 + \frac{\sqrt{2}}{\sqrt{\varepsilon}}\right) 2\varepsilon_{\text{mn}}(\mathbf{C}, \mathbf{A}). \quad \square$$

13.9 Exercises

1. Show that Eq. (13.2) reduces to Eq. (12.1) when ν is the Dirac measure at $y \in \mathbb{R}$.
2. Show that $\Delta_\alpha(\mu^{(\lambda)}, \nu^{(\lambda)}) = \frac{1}{\lambda} \Delta_\alpha(\mu, \nu)$ for the scaled probability measures $\mu^{(\lambda)}, \nu^{(\lambda)}$. Likewise, defining $\mathbf{E}^{(\lambda)}(X) = \mathbf{E}(\lambda X)$ and similarly for $\mathbf{F}^{(\lambda)}$, show that $\Delta_\alpha(\mathbf{E}^{(\lambda)}, \mathbf{F}^{(\lambda)}) = \frac{1}{\lambda} \Delta_\alpha(\mathbf{E}, \mathbf{F})$.
3. Give an example showing that the inequality (13.14) is, in general, strict.
4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded Borel function. Show that the observable $\mathbf{Q} \circ f^{-1}$ has unbounded global measurement noise with respect to $\mathbf{Q}$, that is, $\varepsilon_{\text{mn}}(\mathbf{Q} \circ f^{-1}, \mathbf{Q}) = \infty$.
5. Show that any trivial observable $\mathbf{E} = \mu I$ on $\mathbb{R}$ has infinite distance from sharp position $\mathbf{Q}$: $\Delta_\alpha(\mu I, \mathbf{Q}) = \infty$.
6. Let g be a bounded measurable function on $\mathbb{R}$. Then the distance of $\mathbf{Q} \circ g^{-1}$ from $\mathbf{Q}$ is infinite, $\Delta_\alpha(\mathbf{Q} \circ g^{-1}, \mathbf{Q}) = \infty$.
7. Show that for the shifted position observable $\mathbf{Q}_x$ one has $\Delta_\alpha(\mathbf{Q}, \mathbf{Q}_x) = |x|$.
8. Show that both $\mathcal{N}_i$ and ν are monotonic under smearing, that is, $\mathbf{E}$ is a smearing of $\mathbf{F}$ by p , then $\mathcal{N}_i(\mathbf{E}) \leq \mathcal{N}_i(\mathbf{F})$ and $\nu(\mathbf{E}) \leq \nu(\mathbf{F})$.
9. Show that the observable $(\mathbf{Q} * \mu) \circ f^{-1}$ constructed in Example 13.1 is a non-covariant approximation to $\mathbf{Q}$ in the sense of finite error bars.

References

1. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Physik* **43**, 172–198 (1927)
2. Ludwig, G.: Foundations of Quantum Mechanics. I. Texts and Monographs in Physics. Springer, New York (1983). Translated from the German by Hein, C.A
3. Villani, C.: Optimal Transport - Old and New. Grundlehren der Mathematischen Wissenschaften, vol. 338. Springer, Berlin (2009)
4. Busch, P., Lahti, P., Werner, R.F.: Measurement uncertainty relations. *J. Math. Phys.* **55**, 042111 (2014)
5. Busch, P., Pearson, D.: Universal joint-measurement uncertainty relation for error bars. *J. Math. Phys.* **48**(8), 082103 (2007)
6. Werner, R.: The uncertainty relation for joint measurement of position and momentum. *Quantum Inf. Comput.* **4**(6–7), 546–562 (2004)
7. Busch, P., Pearson, D.: Error and unsharpness in approximate joint measurements of position and momentum (2014). [arXiv:1405.6956](https://arxiv.org/abs/1405.6956)
8. Busch, P., Stevens, N.: Direct tests of measurement uncertainty relations: What it takes. *Phys. Rev. Lett.* **114**(7), 070402 (2015)

9. Haus, H.A., Mullen, J.A.: Quantum noise in linear amplifiers. *Phys. Rev.* **128**, 2407–2413 (1962)
10. Haus, H.A.: Quantum noise, quantum measurement, and squeezing. *J. Opt. B* **6**(8), S626 (2004)
11. Yamamoto, Y., Haus, H.A.: Preparation, measurement and information capacity of optical quantum states. *Rev. Mod. Phys.* **58**, 1001–1020 (1986)
12. Clerk, A.A., Devoret, M.H., Girvin, S.M., Marquardt, F., Schoelkopf, R.J.: Introduction to quantum noise, measurement, and amplification. *Rev. Mod. Phys.* **82**, 1155–1208 (2010)
13. Arthurs, E., Kelly, J.: On the simultaneous measurements of a pair of conjugate observables. *Bell Syst. Tech.* **44**, 725–729 (1965)
14. Arthurs, E., Goodman, M.: Quantum correlations: a generalized Heisenberg uncertainty relation. *Phys. Rev. Lett.* **60**(24), 2447 (1988)
15. Ozawa, M.: Uncertainty relations for noise and disturbance in generalized quantum measurements. *Ann. Phys.* **311**(2), 350–416 (2004)
16. Busch, P., Lahti, P., Werner, R.: Quantum root-mean-square error and measurement uncertainty relations. *Rev. Mod. Phys.* **86**, 1261–1281 (2014)
17. Appleby, D.M.: Concept of experimental accuracy and simultaneous measurements of position and momentum. *Int. J. Theor. Phys.* **37**(5), 1491–1509 (1998)
18. Appleby, D.M.: Error principle. *Int. J. Theor. Phys.* **37**(10), 2557–2572 (1998)
19. Polterovich, L.: Quantum unsharpness and symplectic rigidity. *Lett. Math. Phys.* 1–20 (2012)
20. Polterovich, L.: Symplectic geometry of quantum noise. *Commun. Math. Phys.* **327**(2), 481–519 (2014)
21. Carmeli, C., Heinonen, T., Toigo, A.: Intrinsic unsharpness and approximate repeatability of quantum measurements. *J. Phys. A* **40**(6), 1303–1323 (2007)
22. Miyadera, T., Imai, H.: Heisenberg’s uncertainty principle for simultaneous measurement of positive-operator-valued measures. *Phys. Rev. A* **78**, 052119 (2008)
23. Schwonnek, R., Reeb, D., Werner, R.F.: Measurement uncertainty for finite quantum observables (2016). [arXiv:1604.00382](https://arxiv.org/abs/1604.00382)
24. Bohr, N.: The quantum postulate and the recent development of atomic theory. *Nature* **121**, Suppl., 580–590 (1928). *Symposium on the Foundations of Modern Physics. 1987* (Joensuu, 1987), pp. 1–18. World Sci. Publ, Teaneck, NJ (1987)
25. Miyadera, T.: Uncertainty relations for joint localizability and joint measurability in finite-dimensional systems. *J. Math. Phys.* **52**(7):072105 (2011)

Part III

Realisations

Chapter 14

Qubits

In this chapter we investigate the joint measurability of the observables of a qubit system, that is, a quantum system whose relevant degrees of freedom are represented by a two-dimensional complex Hilbert space. We will give a full characterisation of the pairs of simple qubit observables that are jointly measurable. We also develop the theory of approximate joint measurements of incompatible pairs of sharp qubit observables. The resulting measurement uncertainty relations for mutually unbiased observables are found to be closely related to corresponding optimal preparation uncertainty relations.

14.1 Qubit States and Observables

In a two-dimensional Hilbert space, $\mathcal{H} \cong \mathbb{C}^2$, linear operators correspond to 2×2 matrices with complex entries. These form the 4-dimensional complex vector space $\mathcal{L}(\mathcal{H}) \cong M_2(\mathbb{C})$, spanned by the Pauli matrices,

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Thus, $\mathcal{L}(\mathcal{H})$ is isomorphic to $\mathbb{C}^4$ via the correspondence

$$M_2(\mathbb{C}) \ni A = \sum_{j=0}^3 a_j \sigma_j = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma} \longleftrightarrow (a_0, \mathbf{a}) \in \mathbb{C}^4.$$

The trace-free subspace of $M_2(\mathbb{C})$ is spanned by the three Pauli matrices $\sigma_1, \sigma_2, \sigma_3$. Selfadjoint operators are represented as Hermitian 2×2 matrices, which form

a 4-dimensional real vector space M_4 , which is again spanned by the Pauli basis, now with coefficients $(a_0, \mathbf{a}) \in \mathbb{R}^4$.

The product of two operators $A, B \in M_2(\mathbb{C})$ is given as follows:

$$\begin{aligned} AB &= (a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma})(b_0 I + \mathbf{b} \cdot \boldsymbol{\sigma}) \\ &= (a_0 b_0 + \mathbf{a} \cdot \mathbf{b})I + (b_0 \mathbf{a} + a_0 \mathbf{b} + i \mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}. \end{aligned} \quad (14.1)$$

This is equivalent to the following algebraic identity satisfied by the Pauli matrices:

$$\sigma_k \sigma_\ell = i \sum_m \varepsilon_{k\ell m} \sigma_m + \delta_{k\ell} I \quad (14.2)$$

where $\varepsilon_{k\ell m}$ is the permutation symbol. It is now straightforward to verify that a one-dimensional projection operator is of the form $\varrho_{\mathbf{n}} = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma})$ where $\mathbf{n}$ is a unit vector (Exercise 2 of Sect. 9.6). The spectral decomposition of a selfadjoint operator A (with $\mathbf{a} \neq 0$) is then

$$A = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma} = (a_0 + \|\mathbf{a}\|)\varrho_{\hat{\mathbf{a}}} + (a_0 - \|\mathbf{a}\|)\varrho_{-\hat{\mathbf{a}}}, \quad \hat{\mathbf{a}} = \frac{\mathbf{a}}{\|\mathbf{a}\|}.$$

From this we see that $A \geq 0$ exactly when its eigenvalues $a_0 \pm \|\mathbf{a}\|$ are non-negative, that is, $a_0 \geq \|\mathbf{a}\|$.

Next we define the bilinear form

$$\langle A, B \rangle = a_0 b_0 - \sum_{i=1}^3 a_i b_i = a_0 b_0 - \mathbf{a} \cdot \mathbf{b}.$$

The vector space M_4 , equipped with $\langle \cdot, \cdot \rangle$ as a pseudo inner product and the order $\leq$, is isomorphic to the 4-dimensional Minkowski space, also denoted M_4 . Accordingly, we will apply freely the terminology of Minkowski geometry. Thus, selfadjoint operators A can be considered as (4-)vectors when considered as elements of M_4 , and we refer to $\langle \cdot, \cdot \rangle$ as the (Minkowski) scalar product. We also use the same notation for vectors and for points in M_4 as an affine space.¹

We define $A \geq_o 0$ (equivalently $0 \leq_o A$) to mean that $A \geq 0$ and at least one eigenvalue of A is equal to zero. Then for $A, B \in M_4$, $A \geq_o B$ (or $B \leq_o A$) is defined to mean $A - B \geq_o 0$. We then have

$$\begin{aligned} A \geq 0 &\iff \langle A, A \rangle \geq 0 \text{ and } a_0 \geq 0; \\ A \geq_o 0 &\iff \langle A, A \rangle = 0 \text{ and } a_0 \geq 0. \end{aligned}$$

The latter equivalence characterises those positive A that lie on the surface of the solid forward cone of 0, the positive cone M_4^+ of M_4 .

¹We recall that an affine space over a vector space is a set such that any two points can be joined by a unique vector in a natural way. In special relativity, the elements of M_4 are called point events and the vectors connecting them describe the possibility of joining the events with straight signals.

The *forward* and *backward cones* of an element $A \in M_4$ are defined as the sets

$$M_4^\uparrow(A) = \{B \in M_4 : A \leq_o B\}, \quad M_4^\downarrow(A) = \{B \in M_4 : A \geq_o B\}.$$

One can carry further the Minkowski geometric terminology. We recall that a vector $A \in M_4$ is called *lightlike* if $\langle A, A \rangle = 0$. If $\langle A, A \rangle > 0$ or < 0 , the vector A is called *timelike* or *spacelike*, respectively. Then $A \leq_o B$ is equivalent to $B - A$ being lightlike and $b_0 - a_0 \geq 0$. We will also say that elements $A, B \in M_4$ are called *spacelike separated*, or *related*, if $\langle A - B, A - B \rangle < 0$.

In addition to the Minkowski pseudo inner product one can define the Euclidean inner product on M_4 :

$$(A, B) = a_0 b_0 + \mathbf{a} \cdot \mathbf{b}.$$

The connection between the Euclidean and Minkowski inner products on M_4 can be expressed with the help of a linear map τ on M_4 , defined via $\tau(I) = I$, $\tau(\sigma_k) = -\sigma_k$, $k = 1, 2, 3$, that is:

$$\tau : M_4 \rightarrow M_4, \quad A \mapsto \tau(A) = \tau(a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma}) = a_0 I - \mathbf{a} \cdot \boldsymbol{\sigma}. \quad (14.3)$$

Then

$$\langle A, B \rangle = (A, \tau(B)). \quad (14.4)$$

The map τ extends to an isometry on $M_2(\mathbb{C}^2)$ that is induced by an antiunitary map C via $\tau(A) = CA^*C^*$. It can be used to represent time reversal.

States

States are positive operators of trace one; the (convex and compact) set of states $S(\mathbb{C}^2)$ is thus parametrised as

$$S(\mathbb{C}^2) = \{\varrho = \varrho_n = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma}) \mid \|\mathbf{n}\| \leq 1\}. \quad (14.5)$$

This is the set of operators $\frac{1}{2}(n_0 I + \mathbf{n} \cdot \boldsymbol{\sigma})$ represented by the solid ball of radius 1 and centre $(1, \mathbf{0})$ in the hyperplane $n_0 = 1$. The set of pure states is given by the surface of this ball; thus, ϱ_n is pure if and only if $\mathbf{n}$ is a unit vector, that is $\mathbf{n} = \hat{\mathbf{n}}$.

Effects

Effects are positive operators and bounded by the unit operator; thus, an operator $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$ is an effect if its eigenvalues $\frac{1}{2}(e_0 \pm \|\mathbf{e}\|)$ are in $[0, 1]$, that is,

$$\|\mathbf{e}\| \leq \min\{e_0, 2 - e_0\}. \quad (14.6)$$

This entails, in particular, that $0 \leq e_0 \leq 2$ and $\|\mathbf{e}\| \leq 1$.

Recalling that $\text{co}(X)$ denotes the convex hull of a set X , we note that the set of effects can be described as

$$E(\mathbb{C}^2) = \text{co}(M_4^\uparrow(0)) \cap \text{co}(M_4^\downarrow(I)).$$

This set is convex and compact.

The probability of an event represented by an effect $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$ in the state $\varrho = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma})$ can be expressed in terms of the Euclidean scalar product on M_4 ; one obtains

$$\text{tr}[\varrho E] = \frac{1}{2}(e_0 + \mathbf{n} \cdot \mathbf{e}) = (\varrho, E). \quad (14.7)$$

Remark 14.1 The Minkowski scalar product $\langle E, F \rangle$ of two effects E, F admits a simple physical interpretation: it is equal to the probability $\langle \Phi | E \otimes F \Phi \rangle$ of the joint occurrence of the effects E and F if measured simultaneously on a two-particle system in the entangled (singlet) state

$$\Phi = \frac{1}{\sqrt{2}}(\psi_+ \otimes \psi_- - \psi_- \otimes \psi_+)$$

where $\psi_\pm$ are the normalised eigenvectors of (say) σ_3 . ◁

For later use we note a relation that we may call *anti-Cauchy* inequality for obvious reasons.

Lemma 14.1 *Let $E, F \in E(\mathbb{C}^2)$. Then*

$$\langle E, F \rangle \geq \sqrt{\langle E, E \rangle \langle F, F \rangle}, \quad (14.8)$$

and equality holds if and only if $E = 0$ or $F = \gamma E$ for some $\gamma \geq 0$.

We leave the proof as an exercise.

14.2 Preparation Uncertainty Relations for Qubits

We now consider uncertainty relations for pairs sharp qubit observables given by the selfadjoint operators $A = \mathbf{a} \cdot \boldsymbol{\sigma}$, $B = \mathbf{b} \cdot \boldsymbol{\sigma}$ (where $\mathbf{a}, \mathbf{b}$ are unit vectors). Their eigenvalues are ± 1 and the spectral measure of A (say) takes the form

$$\mathbf{A} = \{\mathbf{A}(\{+1\}), \mathbf{A}(\{-1\})\} = \left\{ \frac{1}{2}(I + \mathbf{a} \cdot \boldsymbol{\sigma}), \frac{1}{2}(I - \mathbf{a} \cdot \boldsymbol{\sigma}) \right\}.$$

From (14.1) we have $[A, B] = AB - BA = 2i(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}$ which shows that $\mathbf{A}$ and $\mathbf{B}$ are compatible if and only if $\mathbf{b} = \mathbf{a}$ or $\mathbf{b} = -\mathbf{a}$. Otherwise, they are incompatible, they cannot be measured jointly.

For any state $\varrho = \varrho_{\mathbf{n}}$ one has $\mathbf{A}_{\varrho}(\{\pm 1\}) = \mathbf{A}_{\varrho}(\pm) = \frac{1}{2}(1 \pm \mathbf{a} \cdot \mathbf{n})$ so that, for instance, $\Delta(A, \varrho_{\mathbf{n}})^2 = 1 - (\mathbf{a} \cdot \mathbf{n})^2$, which allows one to compute the uncertainty

product for any two sharp qubit observables in any state. Rather than doing that, we note the following additive triple and pair uncertainty relations for the spin quantities along the co-ordinate axis:

$$\Delta(\sigma_1, \varrho)^2 + \Delta(\sigma_2, \varrho)^2 + \Delta(\sigma_3, \varrho)^2 \geq 2, \quad (14.9)$$

$$\Delta(\sigma_i, \varrho)^2 + \Delta(\sigma_j, \varrho)^2 = 2 - \Delta(\sigma_k, \varrho)^2 \geq 1. \quad (14.10)$$

Here the first inequality becomes an equation exactly when ϱ is a pure state, and the second inequality is saturated for those pure states ϱ for which $\Delta(\sigma_k, \varrho) = 1$, that is, for which $\text{tr}[\varrho\sigma_k] = 0$.

It is instructive to identify the optimal lower bound for the sum of uncertainties of the qubit observables A, B . The following inequalities then hold:

$$\Delta(A, \varrho) + \Delta(B, \varrho) \geq \|\mathbf{n} \times \mathbf{a}\| + \|\mathbf{n} \times \mathbf{b}\| \geq \|\mathbf{a} \times \mathbf{b}\| = \frac{1}{2} \|[A, B]\|, \quad (14.11)$$

$$\Delta(A, \varrho)^2 + \Delta(B, \varrho)^2 \geq 1 - \|\mathbf{a} \cdot \mathbf{b}\| = 1 - \sqrt{1 - \frac{1}{4} \|[A, B]\|^2}. \quad (14.12)$$

In (14.11), the first inequality is an equality if ϱ is a pure state, whereas $\|\mathbf{n} \times \mathbf{a}\| + \|\mathbf{n} \times \mathbf{b}\|$ assumes its minimum value at $\mathbf{n} = \pm \mathbf{a}$ or $\mathbf{n} = \pm \mathbf{b}$. For the second inequality the lower bound is attained at $\mathbf{n} = (\mathbf{a} + \mathbf{b})/\|\mathbf{a} + \mathbf{b}\|$ if $\mathbf{a} \cdot \mathbf{b} \geq 0$ and at $\mathbf{n} = (\mathbf{a} - \mathbf{b})/\|\mathbf{a} - \mathbf{b}\|$ if $\mathbf{a} \cdot \mathbf{b} \leq 0$.

It is evident that these additive uncertainty relations are more stringent in some respects than the traditional relations for the product of variances. In particular, they have nontrivial tight, state-independent bounds, which are determined by the degree of noncommutativity of the quantities involved.

It should be noted, however, that the above inequalities only give estimates of the actual lower boundary curve of the *uncertainty region* of A, B , defined as the region of all points $(\Delta(A, \varrho), \Delta(B, \varrho))$ in the ΔA - ΔB -plane,

$$\{(\Delta(A, \varrho), \Delta(B, \varrho)) \mid \varrho \in \mathcal{S}(\mathbb{C}^2)\}. \quad (14.13)$$

The lower boundary curve is given by

$$\Delta_{\min} B(\Delta(A)) = \min\{\Delta(B, \varrho) \mid \Delta(A, \varrho) = \Delta A\}.$$

As an illustration, and for later reference, we determine the uncertainty region for a pair A, B with $\mathbf{a} \cdot \mathbf{b} = 0$. Since $\Delta(A, \varrho)^2 = 1 - (\mathbf{a} \cdot \mathbf{n})^2$ and $\Delta(B, \varrho)^2 = 1 - (\mathbf{b} \cdot \mathbf{n})^2$, we see that for given value $\Delta A \in [0, 1]$, the set of ϱ with $\Delta(A, \varrho) = \Delta A$ is given by the vectors $\mathbf{n}$ which satisfy $\mathbf{a} \cdot \mathbf{n} = \pm [1 - (\Delta A)^2]^{1/2}$; among these vectors, the smallest value of $\Delta(B, \varrho)$ is assumed for the unique unit vector $\mathbf{n}_0$ that lies on the circular segment ‘between’ $\mathbf{a}$ and $\mathbf{b}$, so that $(\mathbf{a} \cdot \mathbf{n}_0)^2 + (\mathbf{b} \cdot \mathbf{n}_0)^2 = 1$ and therefore $\Delta_{\min} B(\Delta(A))^2 = 1 - (\Delta A)^2$, which finally entails

$$\Delta(A, \varrho)^2 + \Delta(B, \varrho)^2 \geq 1. \quad (14.14)$$

The lower boundary of the uncertainty region of A, B in $[0, 1] \times [0, 1]$ is the circular segment given by $(\Delta A)^2 + (\Delta B)^2 = 1$. This is the same as the curve given by the limiting case of inequality (14.12), but this coincidence does not arise when $\mathbf{a} \cdot \mathbf{b} \neq 0$.

The upper boundary curve of the uncertainty region in the present case is given by $\Delta B = 1$ since the set of vectors $\mathbf{n}$ satisfying $\Delta(A, \varrho) = \Delta A$ always contains one that is orthogonal to $\mathbf{b}$. Again, if $\mathbf{a} \cdot \mathbf{b} \neq 0$, the upper boundary will not coincide with the straight line $\Delta B = 1$. The full uncertainty region for this general case is described in [1].

14.3 Compatibility of a Pair of Qubit Effects

For any effect E we now write $E^c = I - E$. Let $\mathbf{E} = \{E, E^c\}$ and $\mathbf{F} = \{F, F^c\}$ be two valued (sharp or unsharp) qubit observables. We recall that they are compatible exactly when the generating effects E and F are compatible, that is, there is an effect G , such that $G \leq E$, $G \leq F$, $E + F - G \leq I$, so that $E = G + (E - G)$ and $F = G + (F - G)$.

Due to the simple structure of the qubit effects $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$ and $F = \frac{1}{2}(f_0 I + \mathbf{f} \cdot \boldsymbol{\sigma})$, their compatibility can be expressed in the form of a single inequality relating the defining vectors $(e_0, \mathbf{e})$, $(f_0, \mathbf{f})$. This inequality entails, among other features, the necessary degrees of unsharpness for two noncommuting qubit effects to be compatible.

In preparation of the following theorem we introduce the notations

$$\text{fuzz}(E) = \sqrt{\langle E, E \rangle} + \sqrt{\langle E^c, E^c \rangle}, \quad (14.15)$$

$$\text{bias}(E) = \sqrt{\langle E, E \rangle} - \sqrt{\langle E^c, E^c \rangle}. \quad (14.16)$$

We note the following.

Lemma 14.2 *Let $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$ be an effect in $\mathbb{C}^2$. Then:*

$$0 \leq \text{fuzz}(E) \leq 1, \quad -1 \leq \text{bias}(E) \leq 1. \quad (14.17)$$

Furthermore,

$$\text{fuzz}(E) = 1 \quad \text{iff } E = \frac{1}{2}e_0 I \text{ (trivial effect);} \quad (14.18)$$

$$\text{fuzz}(E) = 0 \quad \text{iff } E = \frac{1}{2}(I + \hat{\mathbf{e}} \cdot \boldsymbol{\sigma}) \text{ (nontrivial projection);} \quad (14.19)$$

$$\text{bias}(E) = 1 \quad \text{iff } E = I; \quad (14.20)$$

$$\text{bias}(E) = -1 \quad \text{iff } E = 0; \quad (14.21)$$

$$\text{bias}(E) = 0 \quad \text{iff } E = \frac{1}{2}(I + \mathbf{e} \cdot \boldsymbol{\sigma}) \text{ (trace-1 effect).} \quad (14.22)$$

It has been shown in [2] that the maps **fuzz** and **bias** meet natural requirements for measures of the fuzziness (or unsharpness) and the biasedness (or bias) of an effect, respectively, some of which correspond to the properties (14.18)–(14.22). We see already here that effects with $\text{fuzz}(E) = 1$ are most uninformative, hence maximally unsharp as they are trivial effects, and effects with $\text{fuzz}(E) = 0$ are nontrivial projections and as such *sharp properties*. Likewise, an effect with $\text{bias}(E) = 0$ is *unbiased*, in the operational sense that its probability is equal to $\frac{1}{2}$ in any state $\varrho = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma})$ that does not distinguish between $+\hat{\mathbf{e}}$ and $-\hat{\mathbf{e}}$, that is, for which $\mathbf{n} \cdot \hat{\mathbf{e}} = 0$. Then it is clear that the effects $I, 0$ are those that are maximally biased.

Further we define

$$F = \text{fuzz}(E)^2 + \text{fuzz}(F)^2, \quad B = \text{bias}(E)^2 + \text{bias}(F)^2, \quad (14.23)$$

and note the following identities for an effect $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$:

$$e_0 - 1 = \text{fuzz}(E)\text{bias}(E), \quad (14.24)$$

$$\|\mathbf{e}\|^2 = (1 - \text{fuzz}(E)^2)(1 - \text{bias}(E)^2). \quad (14.25)$$

Theorem 14.1 *Qubit effects $E, F \in \mathbf{E}(\mathbb{C}^2)$ are compatible if and only if either of the following equivalent inequalities is satisfied:*

$$\begin{aligned} \langle E, E^c \rangle \langle F, F^c \rangle - \sqrt{\langle E, E \rangle \langle F, F \rangle \langle E^c, E^c \rangle \langle F^c, F^c \rangle} \\ \leq \langle E, F^c \rangle \langle E^c, F \rangle + \langle E, F \rangle \langle E^c, F^c \rangle; \end{aligned} \quad (14.26)$$

$$\begin{aligned} -\frac{1}{4} \langle E - F, E - F \rangle \|[E, F]\|^2 \\ \leq \left(\langle E, F \rangle \sqrt{\langle E^c, E^c \rangle \langle F^c, F^c \rangle} + \langle E^c, F^c \rangle \sqrt{\langle E, E \rangle \langle F, F \rangle} \right)^2; \end{aligned} \quad (14.27)$$

$$(1 - F)(1 - B) \leq \langle E - E^c, F - F^c \rangle^2. \quad (14.28)$$

The proof of this theorem in the formulation given here is found in [3]. Equivalent formulations were shown in [4–6].

We have given three equivalent inequalities in the above theorem as each of them offers a somewhat different perspective on the joint measurability of two qubit effects. The first and third are manifestly invariant under the exchange $E \longleftrightarrow E^c$ or $F \longleftrightarrow F^c$, while the second is invariant under the simultaneous application of these exchanges; but since E and F are compatible if and only if E and F^c are compatible, there is another equivalent condition, obtained from (14.27) by (say) $F \longleftrightarrow F^c$.

As a consequence of Lemma 14.1, the left hand side of Eq. (14.26) is non-negative, and it is equal to zero only when both E and F are trivial effects (that is, multiples of I). Also note that this term is composed of quantities that characterise the effects E, F separately, while the term on the right hand side describes a relational feature of E, F .

The left hand side of (14.27) (and its variant for the pair E, F^c) is zero or negative in trivial cases of compatibility, namely, when E, F commute or are comparable ($E \leq F$ or $F \leq E$). In these cases, the inequality is trivially satisfied.

Finally, condition (14.28) is a compact version of the joint measurability criterion found in [6], expressed here in Minkowski-geometric notation. It shows immediately that if the left hand side becomes negative or zero, then compatibility is given: this happens when $\mathbf{F} \geq 1$ and at the same time $\mathbf{B} \leq 1$. This is to say that two effects are compatible as soon as their combined fuzziness is large enough while their combined biasedness remains sufficiently small. For $\mathbf{B} > 1$, one also has $\mathbf{F} > 1$ (since $\mathbf{B} \leq \mathbf{F}$), and similarly, $\mathbf{F} < 1$ entails $\mathbf{B} < 1$; in these cases the left hand side of (14.28) is positive and constitutes a constraint on compatibility. In the unbiased case, $\mathbf{B} = 0$ (which happens if and only if $\text{bias}(E) = \text{bias}(F) = 0$), then sufficient unsharpness guarantees joint measurability. This case will be fully investigated below.

Inequality (14.28) can be recast in a form where both sides are manifestly non-negative. To see this, we note that the left hand side equals $(1 - \mathbf{F})(1 - \mathbf{B}) = 1 - \frac{1}{2}[\mathbf{F}(2 - \mathbf{B}) + \mathbf{B}(2 - \mathbf{F})]$ and the right hand side is $[(e_0 - 1)(f_0 - 1) - \mathbf{e} \cdot \mathbf{f}]^2$. Thus (14.28) is equivalent to

$$\frac{1}{2}[\mathbf{F}(2 - \mathbf{B}) + \mathbf{B}(2 - \mathbf{F})] \geq 1 - [(e_0 - 1)(f_0 - 1) - \mathbf{e} \cdot \mathbf{f}]^2. \quad (14.29)$$

The expressions on both sides are non-negative. For the left hand side this follows from the fact that $0 \leq \mathbf{B}, \mathbf{F} \leq 2$. In the case of the right hand side, one may note that the positivity of E, F and E^c, F^c is equivalent to

$$|e_0 - 1| + \|\mathbf{e}\| \leq 1, \quad |f_0 - 1| + \|\mathbf{f}\| \leq 1,$$

and this immediately gives

$$1 \geq |e_0 - 1| |f_0 - 1| + \|\mathbf{e}\| \|\mathbf{f}\| \geq |(e_0 - 1)(f_0 - 1) - \mathbf{e} \cdot \mathbf{f}|.$$

What emerges from the form (14.29) of the compatibility condition is that *biasedness* figures as a third factor in the interplay between *noncommutativity* and *unsharpness* deciding over the compatibility of E and F .

We rewrite the right hand side of (14.29) (putting $x = e_0 - 1, y = f_0 - 1$):

$$\begin{aligned} & \frac{1}{2}[\mathbf{F}(2 - \mathbf{B}) + \mathbf{B}(2 - \mathbf{F})] \\ & \geq \left[1 - (|xy| + \|\mathbf{e}\| \|\mathbf{f}\|)^2\right] + 2[|xy| \|\mathbf{e}\| \|\mathbf{f}\| + xy \mathbf{e} \cdot \mathbf{f}] + \|\mathbf{e} \times \mathbf{f}\|^2. \end{aligned} \quad (14.30)$$

This shows that the lower bound of inequality (14.29) is composed of three non-negative contributions, one of which is directly related to the commutator of E and F since $\|\mathbf{e} \times \mathbf{f}\|^2 = 4\| [E, F] \|^2$. However, since this inequality will hold in the case of commuting effects E, F , for which $\mathbf{e}, \mathbf{f}$ are (anti-)parallel, one always has $1 - (|xy| - \|\mathbf{e}\| \|\mathbf{f}\|)^2$ as a lower bound for the left hand side, from which it follows that

inequality (14.30) becomes generally valid once the commutator term is removed:

$$\begin{aligned}
 \frac{1}{2}[F(2 - B) + B(2 - F)] &\geq 1 - (|xy| - \|e\|\|f\|)^2 \\
 &= 1 - (|xy| + \|e\|\|f\|)^2 + 4|xy|\|e\|\|f\| \\
 &\geq \left[1 - (|xy| + \|e\|\|f\|)^2\right] + 2[|xy|\|e\|\|f\| + xye \cdot f] \\
 &= 1 - |xy|^2 - \|e\|^2\|f\|^2 + 2xye \cdot f = G(E, F).
 \end{aligned} \tag{14.31}$$

Therefore, it is the commutator term, $\|e \times f\|^2$, that imposes the constraint on compatibility in (14.30). Taking into account the identities (14.24), (14.25), we see that the compatibility inequality (14.29) describes a relation between fuzziness, bias, and the commutator of E , F , additionally involving the term $2xye \cdot f$:

$$\frac{1}{2}[F(2 - B) + B(2 - F)] - G(E, F) \geq 4\| [E, F] \|^2. \tag{14.32}$$

We reiterate that $G(E, F) \geq 0$ and the left hand side is always non-negative and vanishes when E , F commute.

Finally, it is possible to give a generic representation of the term $2xye \cdot f$ using the anticommutator of $E - E^c$ and $F - F^c$ via the identity

$$\text{tr}[\{E - E^c, F - F^c\}] = 4(xy + e \cdot f). \tag{14.33}$$

The condition (14.28) possesses a simple invariance property.

Corollary 14.1 *The following statements are equivalent:*

- (i) *effects $E = \frac{1}{2}(e_0I + e \cdot \sigma)$ and $F = \frac{1}{2}(f_0I + f \cdot \sigma)$ are compatible;*
- (ii) *effects $\tau(E) = \frac{1}{2}(e_0I - e \cdot \sigma)$ and $\tau(F) = \frac{1}{2}(f_0I - f \cdot \sigma)$ are compatible;*
- (iii) *effects $\tau(E^c) = \frac{1}{2}((2 - e_0)I + e \cdot \sigma)$ and $\tau(F^c) = \frac{1}{2}((2 - f_0)I + f \cdot \sigma)$ are compatible.*

Proof It is straightforward to verify that all terms in (14.28) are invariant under the simultaneous substitutions $e_0 \leftrightarrow 2 - e_0$, $f_0 \leftrightarrow 2 - f_0$. This establishes the equivalence of (i) and (iii). Next, (ii) and (iii) are equivalent since $\tau(E) = I - \tau(E^c)$ and $\tau(F) = I - \tau(F^c)$. $\square$

Remark 14.2 The statement of Proposition 11.5 gives a joint measurability condition for pairs of effects in terms of a noise parameter, valid for arbitrary finite-dimensional or separable Hilbert spaces. Beyond this, no general criterion is known that relates joint measurability to unsharpness and bias (and possibly further factors). However, as shown in [7], Theorem 14.1 can be lifted to higher dimensions for a special family of pair of effects in finite-dimensional Hilbert space, namely, those which belong to the von Neumann algebra generated by two projections. $\triangleleft$

Unbiased Effects

We consider a special case of interest where the two effects E, F (as well as their complements) have null-components equal to $\frac{1}{2}$. This case was treated in [8]. Effects of this form are called *unbiased*; they satisfy $\text{bias}(E) = 0 = \text{bias}(F)$.

Proposition 14.1 *Effects $E = \frac{1}{2}(I + \mathbf{e} \cdot \boldsymbol{\sigma})$ and $F = \frac{1}{2}(I + \mathbf{f} \cdot \boldsymbol{\sigma})$ are compatible if and only if one of the following equivalent inequalities holds:*

$$\|\mathbf{e}\|^2 + \|\mathbf{f}\|^2 \leq 1 + (\mathbf{e} \cdot \mathbf{f})^2 \quad (14.34a)$$

$$(1 - \|\mathbf{e}\|^2)(1 - \|\mathbf{f}\|^2) \geq \|\mathbf{e} \times \mathbf{f}\|^2; \quad (14.34b)$$

$$\|\mathbf{e} + \mathbf{f}\| + \|\mathbf{e} - \mathbf{f}\| \leq 2. \quad (14.34c)$$

Proof The proof is obtained most directly from an evaluation of inequality (14.28) for unbiased effects. We leave the details as an exercise. $\square$

Remark 14.3 It is instructive to sketch the original proof of [8], which uses directly the compatibility condition proved in Exercise 1 of Chap. 11. Thus, effects E, F are compatible if and only if there is an operator $G = \frac{1}{2}(\gamma I + g \cdot \boldsymbol{\sigma})$ such that

$$0 \leq E + F - I \leq G \leq E, F.$$

This translates into the following set of geometric inequalities:

$$\|g\| \leq \gamma, \quad \|\mathbf{e} + \mathbf{f} - g\| \leq \gamma, \quad \|\mathbf{e} - g\| \leq 1 - \gamma, \quad \|\mathbf{f} - g\| \leq 1 - \gamma. \quad (14.35)$$

The compatibility of E, F is therefore equivalent to the existence of a vector g and a parameter γ such that g lies in the intersection of four balls centred at $\mathbf{0}, \mathbf{e}, \mathbf{f}$ and $\mathbf{e} + \mathbf{f}$ with radii γ and $1 - \gamma$, respectively. The four points are the vertices of a parallelogram, and the diagonally opposite balls have the same radius. Thus, the intersection of the four balls is non-empty if and only if it contains the intersection of the diagonals, which requires that

$$\|\mathbf{e} + \mathbf{f}\| \leq 2\gamma \leq 2 - \|\mathbf{e} - \mathbf{f}\|.$$

Therefore, in that case one can choose $g = \frac{1}{2}(\mathbf{e} + \mathbf{f})$. We conclude that E, F are jointly measurable if and only if (14.34c) holds. $\triangleleft$

The compatibility condition in the form (14.34b) has a simple operational interpretation. Note that the terms $\|\mathbf{e}\|, \|\mathbf{f}\|$ indicate how close E, F are to being rank-1 (hence non-trivial) projections. For E, F of the given form, we find $\text{fuzz}(E)^2 = 1 - \|\mathbf{e}\|^2$ and $\text{fuzz}(F)^2 = 1 - \|\mathbf{f}\|^2$, which confirms that fuzz measures the fuzziness or unsharpness of E, F . Furthermore, observing that $\|\mathbf{e} \times \mathbf{f}\|^2 = 4\|[E, F]\|^2$, we obtain the following.

Theorem 14.2 *Effects $E = \frac{1}{2}(I + \mathbf{e} \cdot \boldsymbol{\sigma})$ and $F = \frac{1}{2}(I + \mathbf{f} \cdot \boldsymbol{\sigma})$ are compatible if and only if*

$$\text{fuzz}(E)^2 \text{fuzz}(F)^2 \geq 4\| [E, F] \|^2. \quad (14.36)$$

This makes precise how unsharp two noncommuting effects must be in order to be compatible.

The inequalities (14.34a)–(14.34c) turn out to be necessary for the compatibility of any pair of qubit effects.

Proposition 14.2 *Suppose that the effects $E = \frac{1}{2}(e_0 I + \mathbf{e} \cdot \boldsymbol{\sigma})$ and $F = \frac{1}{2}(f_0 I + \mathbf{f} \cdot \boldsymbol{\sigma})$ are compatible. Then also the effects $E^{(1)} = \frac{1}{2}(I + \mathbf{e} \cdot \boldsymbol{\sigma})$ and $F^{(1)} = \frac{1}{2}(I + \mathbf{f} \cdot \boldsymbol{\sigma})$ are compatible, and therefore inequality (14.34c) holds.*

Proof Note that $E^{(1)} = \frac{1}{2}E + \frac{1}{2}\tau(E^c)$ and $F^{(1)} = \frac{1}{2}F + \frac{1}{2}\tau(F^c)$. Therefore the observables $\mathbf{E}^{(1)}$ and $\mathbf{F}^{(1)}$ defined by these effects are convex combinations of the observable pairs $\mathbf{E}, \mathbf{E}^{(\tau)}$ generated by $E, \tau(E^c)$ and $\mathbf{F}, \mathbf{F}^{(\tau)}$ generated by $F, \tau(F^c)$, respectively, with the mixing factor $\lambda = \frac{1}{2}$. It follows from Corollary 14.1 that the compatibility of the pair E, F entails the compatibility of the pair $\tau(E^c), \tau(F^c)$. Then Proposition 11.4 is applicable and gives the compatibility of the observables $\mathbf{E}^{(1)}$ and $\mathbf{F}^{(1)}$, and thus the inequalities (14.34a)–(14.34c). $\square$

Corollary 14.2 *Unbiased qubit observables $\mathbf{E}_1, \mathbf{E}_2$ with generating effects $E_{k,\pm} = \frac{1}{2}(I \pm \mathbf{e}_k \cdot \boldsymbol{\sigma})$ are compatible if and only if*

$$G_{++} = \frac{1}{2}(E_{1,+}E_{2,+} + E_{2,+}E_{1,+}) = \frac{1}{4}((1 + \mathbf{e}_1 \cdot \mathbf{e}_2)I + (\mathbf{e}_1 + \mathbf{e}_2) \cdot \boldsymbol{\sigma}) \quad (14.37)$$

is a positive operator. In this case, a joint observable is given by the four effects

$$\begin{aligned} G_{ij} &= \frac{1}{2}(E_{1,i}E_{2,j} + E_{2,j}E_{1,i}) \\ &= \frac{1}{4}[(1 + i\mathbf{e}_1 \cdot \mathbf{e}_2)I + (i\mathbf{e}_1 + j\mathbf{e}_2) \cdot \boldsymbol{\sigma}], \quad i, j \in \{+, -\}. \end{aligned} \quad (14.38)$$

Proof The positivity of G_{++} is equivalent to $1 + \mathbf{e} \cdot \mathbf{f} \geq |\mathbf{e} + \mathbf{f}|$ and hence to inequality (14.34a). The same holds for the three remaining operators. It is also clear that the four operators add up to I and their partial sums give E and F . $\square$

Remark 14.4 If the vectors $\mathbf{e}_1, \mathbf{e}_2$ have lengths $e \leq 1/\sqrt{2}$, then inequality (14.34a) is satisfied for all possible directions of these vectors, so that the effects are compatible. If these vectors are mutually orthogonal then this length constraint becomes necessary for compatibility. $\triangleleft$

14.4 Excursion: Compatibility of Three Qubit Effects

The question of characterising jointly measurable sets of effects or observables in terms of operational conditions is wide open if one looks beyond the case of a pair of qubit effects. Given the lack of general results, in later chapters we will give some

examples of constructions of joint observables for more than two observables, thus establishing some sufficient conditions for their joint measurability. Here we briefly look at three unbiased qubit observables with the generating effects

$$\begin{aligned} E_{1,\pm} &= \frac{1}{2}(I \pm \mathbf{e}_1 \cdot \boldsymbol{\sigma}), \\ E_{2,\pm} &= \frac{1}{2}(I \pm \mathbf{e}_2 \cdot \boldsymbol{\sigma}), \\ E_{3,\pm} &= \frac{1}{2}(I \pm \mathbf{e}_3 \cdot \boldsymbol{\sigma}). \end{aligned} \quad (14.39)$$

We introduce operators

$$G_{ijk} = \frac{1}{8} (\alpha_{ijk} I + g_{ijk} \cdot \boldsymbol{\sigma}), \quad i, j, k \in \{+, -\}, \quad (14.40)$$

where

$$\left. \begin{aligned} \alpha_{ijk} &= 1 + i j \mathbf{e}_1 \cdot \mathbf{e}_2 + i k \mathbf{e}_1 \cdot \mathbf{e}_3 + j k \mathbf{e}_2 \cdot \mathbf{e}_3 \\ g_{ijk} &= i \mathbf{e}_1 + j \mathbf{e}_2 + k \mathbf{e}_3 \end{aligned} \right\} \quad i, j, k \in \{+, -\}. \quad (14.41)$$

It is readily verified that these operators satisfy the marginality relations $E_{1,\pm} = \sum_{j,k} G_{\pm j,k}$, etc. Thus, the three observables are compatible if all eight operators $G_{ijk} \geq 0$, that is, $\|g_{ijk}\| \leq \alpha_{ijk}$, which is equivalent to

$$0 \leq [1 - \|\mathbf{e}_1\|^2 - \|\mathbf{e}_2\|^2 - \|\mathbf{e}_3\|^2] + (\|g_{ijk}\| - 1)^2. \quad (14.42)$$

Inequality (14.42), in turn, can be expressed in terms of the unsharpness $\text{fuzz}(E_{\nu,\lambda}) = 1 - \|\mathbf{e}_\nu\|^2$ as:

$$2 - (\|g_{ijk}\| - 1)^2 \leq \text{fuzz}(E_{1,i}) + \text{fuzz}(E_{2,j}) + \text{fuzz}(E_{3,k}). \quad (14.43)$$

With the choice $\|\mathbf{e}_\nu\|^2 \leq \frac{1}{3}$ the right hand side is ≥ 2 and the inequality is satisfied for all possible directions of the vectors $\mathbf{e}_\nu$. We thus have the following.

Proposition 14.3 *Three unbiased dichotomic qubit observables E_1, E_2, E_3 with effects $E_{1,i}, E_{2,j}, E_{3,k}$ as given in (14.39) are compatible if*

$$\text{fuzz}(E_{1,i}) + \text{fuzz}(E_{2,j}) + \text{fuzz}(E_{3,k}) \geq 2. \quad (14.44)$$

This condition is satisfied if each effect has fuzziness no less than $2/3$, or equivalently, $\|\mathbf{e}_\nu\| \leq 1/\sqrt{3}$.

The set of operators G_{ijk} of (14.40) forms a joint observable for E_1, E_2, E_3 if and only if (14.43) holds. If $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ have equal length e and are orthogonal, the operators G_{ijk} cannot form a joint observable unless $e \leq 1/\sqrt{3}$.

We note that the construction (14.41) does not give a joint observable in the trivial case where $\mathbf{e}_1 = \mathbf{e}_2 = \mathbf{e}_3$ and these three vectors have length 1: in that case we have $\alpha_{++-} = 0$ but $g_{++-} = \mathbf{e}_1$, so that G_{++-} is not an effect.

Remark 14.5 The method of intersecting spheres reviewed in the previous section for pairs of unbiased qubit effects was extended in [9] to the case of a triple of qubit observables with effects $E_{\nu,\pm} = \frac{1}{2}[(1 \pm x_\nu)I \pm \mathbf{e}_\nu]$, $\nu = 1, 2, 3$, where positivity is equivalent to $|x_\nu| + \|\mathbf{e}_\nu\| \leq 1$, revealing the following as a necessary condition for their joint measurability: there exists a vector g such that

$$\begin{aligned} & \|-\mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3 - g\| + \|\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - g\| \\ & + \|\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 - g\| + \|-\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - g\| \leq 4. \end{aligned} \quad (14.45)$$

Geometrically, this is the condition that the point with position vector g must lie in the intersection of four solid spheres centred at points A, B, C, D with position vectors $-(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)$, $\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$, $-\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, and $+\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$ and radii that add to 4. Recalling that the *Fermat-Torricelli point* associated with a set of points in $\mathbb{R}^3$ is defined as the point for which the sum of distances from all points in the set is minimised, the above necessary joint measurability condition can be phrased by saying that the Fermat-Torricelli point of A, B, C, D must have its sum of distances no greater than 4.

In the case of observables $\{E_{\nu,\pm}\}$ with orthogonal vectors, this condition reduces to (14.44), which is therefore necessary and sufficient if the effects $E_{\nu,\pm}$ are unbiased. Finally, in [10] it is shown that an unbiased triple of qubit effects is jointly measurable if and only if the Fermat-Torricelli point of the associated points A, B, C, D satisfies (14.45). $\triangleleft$

Remark 14.6 It has been shown in [9] that measurement realisations for joint measurements of compatible pairs of triples of qubit effects can be obtained by an adaptation of the Arthurs-Kelly model [11], which was originally formulated for position and momentum measurements (see Sect. 19.1). $\triangleleft$

14.5 Approximate Joint Measurements of Qubit Observables

Let $\mathbf{A}$ and $\mathbf{B}$ be any two incompatible sharp qubit observables. To be specific, and without loss of generality, we assume that $\cos \theta = \mathbf{a} \cdot \mathbf{b} > 0$, where $\mathbf{a}, \mathbf{b}$ are the unit vectors associated with $\mathbf{A}, \mathbf{B}$, respectively. We shall identify their optimal approximate joint measurements among the jointly measurable pairs of simple (unsharp) qubit observables $(\mathbf{C}, \mathbf{D})$. We will see that optimal approximations can be found among pairs of unbiased observables $\mathbf{C} = \{C_\pm = \frac{1}{2}(I \pm \mathbf{c} \cdot \boldsymbol{\sigma})\}$, $\mathbf{D} = \{D_\pm = \frac{1}{2}(I \pm \mathbf{d} \cdot \boldsymbol{\sigma})\}$; recall that these are jointly measurable exactly when the effects C_+, D_+ are compatible, which is the case if and only if

$$\|\mathbf{c} + \mathbf{d}\| + \|\mathbf{c} - \mathbf{d}\| \leq 2.$$

We note that any dichotomic qubit observable can be obtained as a smearing of some sharp qubit observable. Indeed, let $\mathbf{E} = \{E_{\pm} = \frac{1}{2}((1 \pm x)I \pm \mathbf{e} \cdot \boldsymbol{\sigma})\}$, where the positivity of the effects $E_{\pm}$ is equivalent to $|x| + \|\mathbf{e}\| \leq 1$. Using the spectral measure $\mathbf{E}^{\hat{\mathbf{e}} \cdot \boldsymbol{\sigma}} = \{\frac{1}{2}(I \pm \hat{\mathbf{e}} \cdot \boldsymbol{\sigma})\}$, one observes that $\mathbf{E} = p \star \mathbf{E}^{\hat{\mathbf{e}} \cdot \boldsymbol{\sigma}}$, where p is the Markov kernel

$$\begin{aligned} p(+, +) &= (1 + x) + \|\mathbf{e}\|, & p(-, +) &= 1 - p(+, +), \\ p(+, -) &= (1 + x) - \|\mathbf{e}\|, & p(-, -) &= 1 - p(+, -). \end{aligned}$$

Error Measure for Qubit Approximations

We shall use the metrics D_1, D_{∞} as well as the Wasserstein-2 distance Δ_2 to quantify the degree of approximation of the measurements.

For simple observables $\mathbf{E} = \{E_+, E_-\}$, $\mathbf{F} = \{F_+, F_-\}$, the distances $D_1(\mathbf{E}, \mathbf{F})$ and $D_{\infty}(\mathbf{E}, \mathbf{F})$ are readily computed. Observing that $E_+ - F_+ = F_- - E_-$ one gets

$$D_1(\mathbf{E}, \mathbf{F}) = D_{\infty}(\mathbf{E}, \mathbf{F}) = \|E_+ - F_+\| = \|E_- - F_-\| = \frac{1}{2}|e_0 - f_0| + \frac{1}{2}\|\mathbf{e} - \mathbf{f}\|.$$

By the result quoted in Remark 13.2 for the finite valued probability measures, the Wasserstein-2 distance is known to be controlled by the distance defined by total variation norm. Here we demonstrate this connection showing that $\Delta_2(\mathbf{E}, \mathbf{F})^2 = 4D_{\infty}(\mathbf{E}, \mathbf{F})$ [12].

For comparison, we note that for $\mathbf{C} = p \star \mathbf{E}^{\hat{\mathbf{e}} \cdot \boldsymbol{\sigma}}$ one has $D_{\infty}(\mathbf{C}, \mathbf{E}^{\hat{\mathbf{e}} \cdot \boldsymbol{\sigma}}) = \frac{1}{2}|c_0 - 1| + \frac{1}{2}(1 - \|\mathbf{c}\|)$.

In contrast to the metrics D_1, D_{∞} , the distance Δ_2 depends explicitly on the value space and on the metrics chosen. While the observables to be approximated are ± 1 -valued, the approximators could a priori be allowed to have different values (although we choose them to be dichotomic). Suppose that target observable $\mathbf{E}$, say, is ± 1 -valued, but the approximating observable $\mathbf{F}$ has the values $a_{\pm}$ (with $a_+ > a_-$). In order to calculate their 2-distance, we need to minimise the quantity (cf. Eq. (13.2))

$$\Delta_2^{\gamma}(\mathbf{E}_{\varrho}, \mathbf{F}_{\varrho})^2 = \int (x - y)^2 d\gamma(x, y),$$

where γ is any coupling for $\mathbf{E}_{\varrho}, \mathbf{F}_{\varrho}$. The 2-distance should vanish when the probabilities of $\mathbf{E}$ and $\mathbf{F}$ coincide for their corresponding values, $\pm 1 \leftrightarrow a_{\pm}$.

Any coupling $\gamma \in \Gamma(\mathbf{E}_{\varrho}, \mathbf{F}_{\varrho})$ is given by four positive numbers,

$$\begin{aligned} (1, a_+) &\mapsto \gamma_{++} = \gamma, \\ (1, a_-) &\mapsto \gamma_{+-} = \mathbf{E}_{\varrho}(+1) - \gamma, \\ (-1, a_+) &\mapsto \gamma_{-+} = \mathbf{F}_{\varrho}(a_+) - \gamma, \\ (-1, a_-) &\mapsto \gamma_{--} = 1 - \mathbf{E}_{\varrho}(+1) - \mathbf{F}_{\varrho}(a_+) + \gamma. \end{aligned}$$

It is then straightforward to obtain

$$\begin{aligned}\Delta_2^\gamma(\mathbf{E}_\varrho, \mathbf{F}_\varrho)^2 &= (1 + a_-)^2 - 4\gamma(a_+ - a_-) - 4\mathbf{E}_\varrho(+1)a_- \\ &\quad + \mathbf{F}_\varrho(a_+)[(1 + a_+)^2 - (1 + a_-)^2].\end{aligned}$$

To minimise this quantity, γ must be chosen as large as allowed by the positivity constraints (given that $a_+ - a_- > 0$), hence $\gamma = \min\{\mathbf{E}_\varrho(+1), \mathbf{F}_\varrho(a_+)\}$. Now it is easy to see that the minimum, $\Delta_2(\mathbf{E}_\varrho, \mathbf{F}_\varrho)$, can only vanish for $\mathbf{E}_\varrho(+1) = \mathbf{F}_\varrho(a_+)$ if $a_+ = 1$ and $a_- = -1$.

Thus we assume that the approximating observables are also ± 1 -valued. We then obtain

$$\Delta_2(\mathbf{E}_\varrho, \mathbf{F}_\varrho)^2 = 4|\mathbf{E}_\varrho(+1) - \mathbf{F}_\varrho(+1)| = 2|e_0 - f_0 + \mathbf{n} \cdot (\mathbf{e} - \mathbf{f})|.$$

By maximising this over all states $\varrho = \varrho_{\mathbf{n}}$ one has the worst-case error estimate

$$\Delta_2(\mathbf{E}, \mathbf{F})^2 = 2(|e_0 - f_0| + \|\mathbf{e} - \mathbf{f}\|) = 4D_\infty(\mathbf{E}, \mathbf{F}). \quad (14.46)$$

Optimal Approximations and Measurement Uncertainty Relation

We consider now the problem of finding optimal approximate joint measurements of two incompatible sharp qubit observables $\mathbf{A}, \mathbf{B}$ with outcome sets $\{-1, +1\}$. As discussed above, we choose the approximators to be dichotomic observables with the same outcome spaces. Hence the joint observables will have four outcomes. We choose the measure D_∞ to quantify the error as the distance between $\mathbf{A}, \mathbf{B}$ and the corresponding margins $\mathbf{C}, \mathbf{D}$ of the joint observable in question.

We call a point $(D_1, D_2) \in [0, 1] \times [0, 1]$ *admissible* if $D_1 = D_\infty(\mathbf{C}, \mathbf{A})$ and $D_2 = D_\infty(\mathbf{D}, \mathbf{B})$ for some jointly measurable qubit observables $\mathbf{C}$ and $\mathbf{D}$. Not all points in the square $[0, 1] \times [0, 1]$ are admissible; for instance the point $(0, 0)$ is not an admissible point since this would mean that $\mathbf{C} = \mathbf{A}$ and $\mathbf{D} = \mathbf{B}$, which is impossible since $\mathbf{C}, \mathbf{D}$ are compatible but $\mathbf{A}, \mathbf{B}$ are not. We shall characterise the region of admissible points. The search for admissible points (D_1, D_2) is narrowed down by the following simple observation. Let $\mathbf{C} = \{c_0 I, (1 - c_0)I\}$ be a trivial observable. Then $D_\infty(\mathbf{C}, \mathbf{A}) = \max\{c_0, 1 - c_0\}$ and therefore

$$\{D_\infty(\mathbf{C}, \mathbf{A}) \mid c_0 \in [0, 1]\} = \left[\frac{1}{2}, 1\right].$$

Thus, approximations by trivial observables will never give distances below $\frac{1}{2}$. Furthermore, since such a $\mathbf{C}$ is jointly measurable with any observable $\mathbf{D}$, and since $D_\infty(\mathbf{D}, \mathbf{B})$ can assume any value in $[0, 1]$, it follows that all points in the set $[0, 1] \times [0, 1] \setminus [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ are trivially admissible. We will therefore concentrate on the characterisation of admissible points (D_1, D_2) in the region $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$.

For the remainder of this Chapter, the approximators for the sharp target observables $\mathbf{A}, \mathbf{B}$ will be understood to be

$$\begin{aligned}\mathbf{C} &= \left\{ \frac{1}{2}(c_0 I + \mathbf{c} \cdot \boldsymbol{\sigma}), \frac{1}{2}((2 - c_0)I - \mathbf{c} \cdot \boldsymbol{\sigma}) \right\}, \\ \mathbf{D} &= \left\{ \frac{1}{2}(d_0 I + \mathbf{d} \cdot \boldsymbol{\sigma}), \frac{1}{2}((2 - d_0)I - \mathbf{d} \cdot \boldsymbol{\sigma}) \right\},\end{aligned}$$

respectively. We shall also use the associated unbiased approximators

$$\mathbf{C}^{(1)} = \left\{ \frac{1}{2}(I \pm \mathbf{c} \cdot \boldsymbol{\sigma}) \right\}, \quad \mathbf{D}^{(1)} = \left\{ \frac{1}{2}(I \pm \mathbf{d} \cdot \boldsymbol{\sigma}) \right\}.$$

The next two results will be proved in the Appendix.

Lemma 14.3 *Any admissible point $(D_1, D_2) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ has a realisation of the type $D_1 = D_\infty(\mathbf{C}^{(1)}, \mathbf{A})$, $D_2 = D_\infty(\mathbf{D}^{(1)}, \mathbf{B})$, where $\mathbf{c}$ and $\mathbf{d}$ are in the plane spanned by $\mathbf{a}$ and $\mathbf{b}$.*

Lemma 14.4 *The set of admissible points is a closed convex set which is reflection symmetric with respect to the axis $D_1 = D_2$; that is, with every admissible point (D_1, D_2) the point (D_2, D_1) is also admissible. Thus, the segment of the boundary curve defined as the graph of the function*

$$D_1 \mapsto D_2^*(D_1) = \min\{D_2 : (D_1, D_2) \text{ is admissible}\} \quad (14.47)$$

is convex, symmetric and belongs to the set of admissible points.

Example 14.1 If $D_1 = D_\infty(\mathbf{C}^{(1)}, \mathbf{A}) = 0$ (i.e. $\mathbf{c} = \mathbf{a}$), then the joint measurability requirement implies that $\mathbf{c}$ and $\mathbf{d}$ are parallel and thus,

$$D_\infty(\mathbf{D}^{(1)}, \mathbf{B}) = \frac{1}{2}\|\mathbf{d} - \mathbf{b}\| \geq \frac{1}{2}\sqrt{1 - (\mathbf{a} \cdot \mathbf{b})^2} = \frac{1}{2}\sin \theta.$$

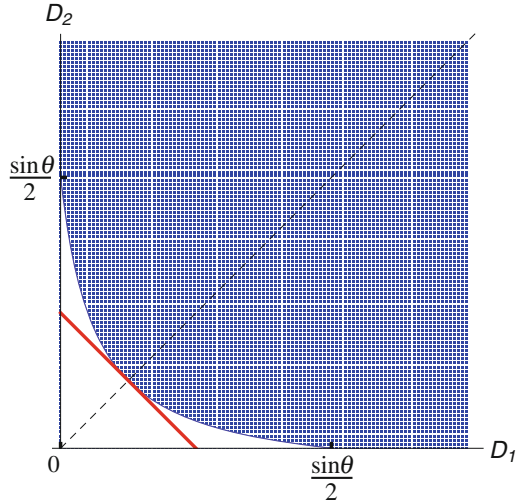
The lower bound is attained when $\mathbf{d} = \cos \theta \mathbf{a} = (\mathbf{a} \cdot \mathbf{b})\mathbf{a}$. We conclude that $(0, \frac{1}{2}\sin \theta)$ and $(\frac{1}{2}\sin \theta, 0)$ are points in the boundary of the admissible region. $\triangleleft$

We next determine the boundary point with $D_1 = D_2 = D_0$. Due to the convexity of the admissible region and its reflection symmetry with respect to the line $D_1 = D_2$, it follows immediately that the admissible region is bounded below tightly by the straight line $D_1 + D_2 = 2D_0$. This situation is sketched in Fig. 14.1. Determination of the value of D_0 yields the following result.

Proposition 14.4 *Any admissible point $(D_1, D_2) = (D_\infty(\mathbf{C}, \mathbf{A}), D_\infty(\mathbf{D}, \mathbf{B}))$ satisfies the error trade-off relation*

$$D_\infty(\mathbf{C}, \mathbf{A}) + D_\infty(\mathbf{D}, \mathbf{B}) \geq 2D_0, \quad (14.48)$$

Fig. 14.1 The admissible region (shaded area) and the line $D_1 + D_2 = 2D_0$ (thick line). The dashed line is the symmetry axis $D_1 = D_2$



where the lower bound is

$$2D_0 = \frac{1}{2\sqrt{2}} [\|\mathbf{a} + \mathbf{b}\| + \|\mathbf{a} - \mathbf{b}\| - 2] = \frac{1}{\sqrt{2}} \left(\sqrt{1 + 2\|[A_+, B_+]\|} - 1 \right).$$

The point (D_0, D_0) is admissible.

Before we prove this result, we comment on its significance. The second expression given for the lower bound $2D_0$ is a monotonic function of the commutator of the two projections defining the sharp observables being approximated. So, it is the degree of noncommutativity of these observables that limits the accuracy of their joint approximation. The first form of $2D_0$ is proportional to the expression $\|\mathbf{a} + \mathbf{b}\| + \|\mathbf{a} - \mathbf{b}\| - 2$, which appears in the qubit compatibility condition (14.34c). For projections (here $\mathbf{a}, \mathbf{b}$ are unit vectors), this expression is always positive unless $\mathbf{a}, \mathbf{b}$ are collinear, in which case its value is zero. Thus, the quantity $2D_0$ is once again a *measure of incompatibility* of the two sharp observables.

Proof The equality of the two expressions for the bound is due to the identities

$$\begin{aligned} \|\mathbf{a} + \mathbf{b}\| + \|\mathbf{a} - \mathbf{b}\| &= 2 \left(\cos \frac{\theta}{2} + \sin \frac{\theta}{2} \right) = 2\sqrt{1 + \sin \theta}, \\ \sin \theta &= \|\mathbf{a} \times \mathbf{b}\| = 2\|[A_+, B_+]\|. \end{aligned}$$

Consider the set of all jointly measurable unbiased observables $\mathbf{C}^{(1)}, \mathbf{D}^{(1)}$ such that $\mathbf{c}, \mathbf{d}$ have equal fixed distance from $\mathbf{a}, \mathbf{b}$, respectively: $\|\mathbf{c} - \mathbf{a}\| = \|\mathbf{d} - \mathbf{b}\| = 2D$ (so that $D_\infty(\mathbf{C}^{(1)}, \mathbf{A}) = D_\infty(\mathbf{D}^{(1)}, \mathbf{B}) = D$). If $(\mathbf{c}, \mathbf{d})$ is not symmetric under reflection with respect to the line through $\mathbf{a} + \mathbf{b}$, denote by $\bar{\mathbf{c}}$ and $\bar{\mathbf{d}}$ the mirror images of $\mathbf{d}$ and $\mathbf{c}$, respectively. Then, if $\mathbf{C}, \mathbf{D}$ are jointly measurable, so are $\bar{\mathbf{C}} = \{\frac{1}{2}(I \pm \bar{\mathbf{c}} \cdot \boldsymbol{\sigma})\}$ and

$\bar{\mathbf{D}} = \left\{ \frac{1}{2}(\mathbf{I} \pm \bar{\mathbf{d}} \cdot \boldsymbol{\sigma}) \right\}$ as the condition (14.34c) is invariant under reflections. Due to Proposition 11.4, the observables

$$\begin{aligned} \frac{1}{2}\mathbf{C}^{(1)} + \frac{1}{2}\bar{\mathbf{C}}^{(1)} &= \hat{\mathbf{C}}^{(1)} = \left\{ \frac{1}{2}(\mathbf{I} \pm \frac{1}{2}(\mathbf{c} + \bar{\mathbf{c}}) \cdot \boldsymbol{\sigma}) \right\}, \\ \frac{1}{2}\mathbf{D}^{(1)} + \frac{1}{2}\bar{\mathbf{D}}^{(1)} &= \hat{\mathbf{D}}^{(1)} = \left\{ \frac{1}{2}(\mathbf{I} \pm \frac{1}{2}(\mathbf{d} + \bar{\mathbf{d}}) \cdot \boldsymbol{\sigma}) \right\} \end{aligned}$$

are jointly measurable. It is clear from their definitions that the vectors $(\mathbf{c} + \bar{\mathbf{c}})$ and $(\mathbf{d} + \bar{\mathbf{d}})$ are mirror images of each other. As $\mathbf{c}, \mathbf{d}$ have equal distance $\frac{1}{2}D$ from $\mathbf{a}, \mathbf{b}$, respectively, this means that $\mathbf{c}$ and $\bar{\mathbf{c}}$ have equal distance $2D$ from $\mathbf{a}$. It follows that the distance from $\mathbf{a}$ to $(\mathbf{c} + \bar{\mathbf{c}})/2$ is less than $2D$ (or $= 2D$ if $\mathbf{c} = \bar{\mathbf{c}}$). We conclude that if $\mathbf{c}, \mathbf{d}$ are not mirror images of each other, there is a pair of jointly measurable observables (namely, $\hat{\mathbf{C}}^{(1)}, \hat{\mathbf{D}}^{(1)}$) with smaller (and equal) distances from $\mathbf{A}, \mathbf{B}$ and mirror symmetric vectors. This shows that the minimal equal distance approximations of $\mathbf{A}, \mathbf{B}$ by means of jointly measurable observables occur among the unbiased pairs $\mathbf{C}^{(1)}, \mathbf{D}^{(1)}$ with $\mathbf{c}, \mathbf{d}$ mirror symmetric with respect to $\mathbf{a} + \mathbf{b}$.

If coordinates are chosen such that $\mathbf{a} = (\sin \frac{\theta}{2}, \cos \frac{\theta}{2})$, $\mathbf{b} = (-\sin \frac{\theta}{2}, \cos \frac{\theta}{2})$, then let a symmetric pair $\mathbf{c}, \mathbf{d}$ be given by $\mathbf{c} = (u, v)$ and $\mathbf{d} = (-u, v)$, with $u, v > 0$. For such pairs, the joint measurability condition for $\mathbf{C}, \mathbf{D}$ assumes the form $u + v \leq 1$. It follows that the shortest (equal) distances $2D$ of $\mathbf{c}, \mathbf{d}$ from $\mathbf{a}, \mathbf{b}$ are assumed when $u + v = 1$ and $\mathbf{a} - \mathbf{c}$ is perpendicular to the line $u + v = 1$. But this distance $2D$ is equal to the distance of the lines $u + v = 1$ and $u + v = \cos \frac{\theta}{2} + \sin \frac{\theta}{2}$, hence

$$2D = \frac{1}{\sqrt{2}} (\cos \frac{\theta}{2} + \sin \frac{\theta}{2} - 1) = 2D_0.$$

This completes the proof. $\square$

The approximations $\mathbf{C}^{(1)}$ and $\mathbf{D}^{(1)}$ leading to the boundary point (D_0, D_0) are generally not among the smearings of $\mathbf{A}$ and $\mathbf{B}$. Indeed, let us denote by $2D_0^c$ the smallest distance achieved under the assumptions that $D_0^c = D_\infty(\mathbf{C}', \mathbf{A}) = \frac{1}{2}(1 - \|\mathbf{c}'\|) = D_\infty(\mathbf{D}', \mathbf{B}) = \frac{1}{2}(1 - \|\mathbf{d}'\|)$ where $\mathbf{C}', \mathbf{D}'$ are jointly measurable and (unbiased) smearings of $\mathbf{A}, \mathbf{B}$, respectively, with $\mathbf{c}' = c'\mathbf{a}$ and $\mathbf{d}' = d'\mathbf{b}$. If the vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal, then $D_0^c = D_0$. However, if $0 < \theta < \frac{\pi}{2}$, then

$$D_0^c = \frac{1}{2} \left(1 - \frac{\sqrt{1 - \sin \theta}}{\cos \theta} \right) > D_0. \quad (14.49)$$

The approximating vectors $\mathbf{c}, \mathbf{d}$ and $\mathbf{c}', \mathbf{d}'$ for $\mathbf{a}, \mathbf{b}$ are illustrated in Fig. 14.2. We conclude that to attain the best jointly measurable approximations of two sharp qubit observables, we are forced to seek approximating observables beyond their smeared versions.

The inequality (14.48) is tight only in the point $(D_1, D_2) = (D_0, D_0)$, for which the two approximation errors are equal. Using the function D_2^* defined in Eq. (14.47), the

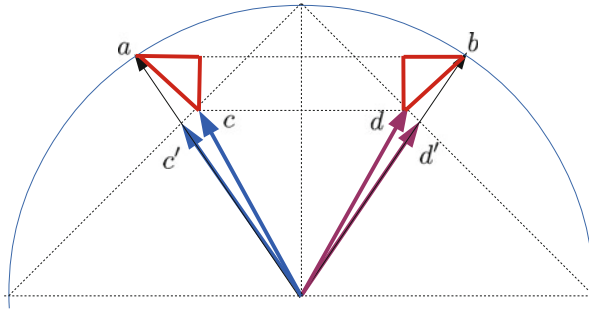


Fig. 14.2 The vectors $\mathbf{c}, \mathbf{d}$ correspond to the optimal compatible approximations $\mathbf{C}^{(1)}$ and $\mathbf{D}^{(1)}$, and the vectors $\mathbf{c}', \mathbf{d}'$ correspond to the closest (unbiased) compatible smearings of $\mathbf{A}, \mathbf{B}$. The latter are clearly suboptimal joint approximations

optimal, tight measurement uncertainty relation for two simple sharp qubit observables can be expressed as

$$D_{\infty}(\mathbf{D}, \mathbf{B}) \geq D_2^*(D_{\infty}(\mathbf{C}, \mathbf{A})). \quad (14.50)$$

The function D_2^* can be completely—albeit implicitly—characterised [13], and it has been shown that this optimal error bound is governed by an interplay between the incompatibility of $\mathbf{A}$ and $\mathbf{B}$ and the degrees of unsharpness of $\mathbf{C}, \mathbf{D}$ [14]. The explicit form of this trade-off and its derivation are rather involved and we sketch here only the solution for the case $\mathbf{a} \cdot \mathbf{b} = 0$.

We note first that the compatibility constraint $\|\mathbf{c} + \mathbf{d}\| + \|\mathbf{c} - \mathbf{d}\| \leq 2$ can be described geometrically as the condition that for given $\mathbf{c} \neq 0$, the vector $\mathbf{d}$ must lie within a closed ellipsoid with boundary given by the limiting case of equality in this inequality; the ellipsoid has its semi-major axis along $\mathbf{c}$ and lies within the unit ball centred at the origin, touching the surface at $\pm \hat{\mathbf{c}}$. Equivalently, one may consider $\mathbf{d} \neq 0$ fixed and find $\mathbf{c}$ constrained within a similar ellipsoid with axis along $\mathbf{d}$. It is clear that in any optimising constellation, the end points of $\mathbf{c}, \mathbf{d}$ must lie on the surfaces of their respective constraining ellipsoids; to be as close to $\mathbf{a}, \mathbf{b}$, respectively, they must lie within the plane spanned by the latter vectors. Moreover, the endpoint of the vector $\mathbf{c}$ ($\mathbf{d}$) will lie on the surface of a ball centred at $\mathbf{a}$ ($\mathbf{b}$), which is thus tangent to the constraining ellipsoid and has radius equal to the distance $\|\mathbf{c} - \mathbf{a}\|$ ($\|\mathbf{d} - \mathbf{b}\|$).

Given the orthogonality of $\mathbf{a}, \mathbf{b}$, it is not hard to see that these conditions are met when $\mathbf{c} = c\mathbf{a}, \mathbf{d} = d\mathbf{b}$ with $c, d > 0$, where the compatibility condition is met if and only if $c^2 + d^2 = 1$, in agreement with inequality (14.34a). In that case, the optimal errors are

$$D_{\infty}(\mathbf{C}, \mathbf{A}) = \frac{1}{2}(1 - c), \quad D_{\infty}(\mathbf{D}, \mathbf{B}) = \frac{1}{2}(1 - d),$$

and so

$$(2D_\infty(\mathbf{C}, \mathbf{A}) - 1)^2 + (2D_\infty(\mathbf{D}, \mathbf{B}) - 1)^2 = 1.$$

For general unbiased compatible approximators $\mathbf{C}, \mathbf{D}$, one thus obtains the qubit measurement uncertainty relation

$$(D_\infty(\mathbf{C}, \mathbf{A}) - \tfrac{1}{2})^2 + (D_\infty(\mathbf{D}, \mathbf{B}) - \tfrac{1}{2})^2 \leq \tfrac{1}{4}. \quad (14.51)$$

This shows that the admissible region is lower-bounded within $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$ by the segment of the circle with radius $\frac{1}{2}$ centred at $(\frac{1}{2}, \frac{1}{2})$, in line with Fig. 14.1.

Finally we note that there is a close connection between the measurement uncertainty relation just found for observables $\mathbf{A}, \mathbf{B}$ with orthogonal vectors $\mathbf{a}, \mathbf{b}$ and the preparation uncertainty relation (14.14) [12]. If we define the state $\varrho_0 = \frac{1}{2}(I + (\mathbf{c} + \mathbf{d}) \cdot \boldsymbol{\sigma})$ with orthogonal vectors $\mathbf{c} = c\mathbf{a}, \mathbf{d} = d\mathbf{b}$ satisfying $c^2 + d^2 \leq 1$, we find that

$$\Delta(\mathbf{A}, \varrho_0)^2 = 1 - (\mathbf{a} \cdot (\mathbf{c} + \mathbf{d}))^2 = 1 - c^2, \quad \Delta(\mathbf{B}, \varrho_0)^2 = 1 - (\mathbf{b} \cdot (\mathbf{c} + \mathbf{d}))^2 = 1 - d^2,$$

and so, in agreement with (14.14),

$$\Delta(\mathbf{A}, \varrho)^2 + \Delta(\mathbf{B}, \varrho)^2 = 2 - c^2 - d^2 \geq 1.$$

The connection with the measurement uncertainty relation (14.51) becomes apparent if we observe that both relations are equivalent to $c^2 + d^2 \leq 1$. This condition ensures both the conditions for ϱ_0 to be a state and for $\mathbf{c}, \mathbf{d}$ to define compatible qubit observables $\mathbf{C}, \mathbf{D}$. We can shed more light on this connection by recalling that a joint observable of the kind (14.38) can be given for $\mathbf{C}, \mathbf{D}$, which here assumes the form

$$G_{kl} = \tfrac{1}{4}(I + (k\mathbf{c} + l\mathbf{d}) \cdot \boldsymbol{\sigma});$$

the positivity of these operators is equivalent to $c^2 + d^2 \leq 1$, hence again to the positivity of ϱ_0 .

There is also a close analogy with a similar connection between preparation and measurement uncertainty in the case of phase space measurements [12]. As we have seen, a covariant phase space observable is generated by a positive trace-one operator T by application of the Weyl operators (and integration over phase space cells). In the present case we can similarly define a (projective) representation of the shift group on the discrete ‘phase space’ $\mathbb{Z}^2 \times \mathbb{Z}^2$ (already encountered in Example 13.5). Thus, let $\mathbf{a} = \mathbf{i}, \mathbf{b} = \mathbf{j}, \mathbf{c} = \mathbf{k}$ be a right-handed orthogonal triple of unit vectors; consider $X = \mathbf{a} \cdot \boldsymbol{\sigma} = \sigma_1$ as the ‘position’, $Y = \mathbf{b} \cdot \boldsymbol{\sigma} = \sigma_2$ as the ‘momentum’, and put $Z = \mathbf{c} \cdot \boldsymbol{\sigma} = \sigma_3$. A shift of the values of X alone is generated by the unitary and selfadjoint operator Y as $YXY = -X$, and similarly a shift of the values of Y alone is generated by X as $XYX = -Y$. Then $XY = iZ$, which generates simultaneous shifts, $ZXZ = -Z, ZYZ = -Y$. Finally $X^2 = Y^2 = Z^2 = I$.

We see that the effects of the above joint observable can then be generated from $\frac{1}{2}\varrho_0 = G_{++}$ as $G_{+-} = \frac{1}{2}X\varrho_0X$, $G_{-+} = \frac{1}{2}Y\varrho_0Y$, $G_{--} = \frac{1}{2}Z\varrho_0Z$. In this way one can say that the measurement error relation for the margins $\mathbf{C}$, $\mathbf{D}$ is reduced to the preparation uncertainty relation for $\mathbf{A}$, $\mathbf{B}$ in the state ϱ_0 .

Remark 14.7 The steps performed in this section towards obtaining the qubit measurement uncertainty relation (14.48) are illustrated in an interactive demonstration available at the Wolfram Demonstrations Project web page with the URL <http://demonstrations.wolfram.com/HeisenbergTypeUncertaintyRelationForQubits/>. ◁

Qubit Measurement Uncertainty in Terms of Error Bar Width

In Proposition 12.4 and Theorem 13.4 we have seen versions of preparation and measurement uncertainty relations for discrete observables formulated in terms of overall width and error bar width, respectively. To illustrate that these measures yield nontrivial relations even in the simplest discrete case of two-point value spaces, we specify these relations to the case of the two sharp qubit observables $\mathbf{Z} = \mathbf{E}^{\sigma_3}$ and $\mathbf{X} = \mathbf{E}^{\sigma_1}$. The value spaces are $\Omega_Z = \Omega_X = \{1, -1\}$, and we equip them with the discrete metric d . (We note that the choice of the values does not affect the results below.)

Let $\varepsilon_1, \varepsilon_2 \in [0, 1]$ with $\varepsilon_1 + \varepsilon_2 \leq 1$, and let ϱ be any state. The overall widths $W_{\varepsilon_1}(\mathbf{Z}_\varrho)$ and $W_{\varepsilon_2}(\mathbf{X}_\varrho)$ then satisfy the following inequality.

$$\max_{z \in \Omega_Z} |O_d(z, W_{\varepsilon_1}(\mathbf{Z}_\varrho))| \cdot \max_{x \in \Omega_X} |O_d(x, W_{\varepsilon_2}(\mathbf{X}_\varrho))| \geq 2(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (14.52)$$

For the choice $\varepsilon_1 = \varepsilon_2 = 0$, the value of the bound is 2. Suppose $\mathbf{Z}$ (say) is sharply localised, so that one of the outcomes has probability 1; then $\max_{z \in \Omega_Z} |O_d(z, W_{\varepsilon_1}(\mathbf{Z}_\varrho))| = 1$. It follows that $\max_{x \in \Omega_X} |O_d(x, W_{\varepsilon_2}(\mathbf{X}_\varrho))| = 2$; this means that $\mathbf{X}$ cannot be sharply localised in the eigenstates of $\mathbf{Z}$, and vice versa.

On the other hand, if one stipulates

$$\max_{z \in \Omega_Z} |O_d(z, W_{\varepsilon_1}(\mathbf{Z}_\varrho))| = \max_{x \in \Omega_X} |O_d(x, W_{\varepsilon_2}(\mathbf{X}_\varrho))| = 1,$$

then the above inequality entails that one must accept $\varepsilon_1 + \varepsilon_2 \geq 1 - 1/\sqrt{2}$. This simply reproduces the Landau–Pollak relation (12.16) (which can, of course, also be verified by direct calculation in the present case):

$$\max_{z \in \Omega_Z} Z_\varrho(\{z\}) + \max_{x \in \Omega_X} X_\varrho(\{x\}) \leq 1 + \frac{1}{\sqrt{2}}.$$

Let $\mathbf{M}$ be an observable on $\Omega_Z \times \Omega_X$, with margins $\mathbf{M}_1, \mathbf{M}_2$ on the value spaces Ω_Z and Ω_X , respectively. Then for $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 \leq 1$,

$$\max_{z \in \Omega_Z} |O_d(z, \mathcal{W}_{\varepsilon_1}(\mathbf{M}_1, \mathbf{Z}))| \cdot \max_{x \in \Omega_X} |O_d(x, \mathcal{W}_{\varepsilon_2}(\mathbf{M}_2, \mathbf{X}))| \geq 2(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (14.53)$$

If one assumes

$$\max_{z \in \Omega_Z} |O_d(z, \mathcal{W}_{\varepsilon_1}(\mathbf{M}_1, \mathbf{Z}))| = \max_{x \in \Omega_X} |O_d(x, \mathcal{W}_{\varepsilon_2}(\mathbf{M}_2, \mathbf{X}))| = 1, \quad (14.54)$$

it follows again that $\varepsilon_1 + \varepsilon_2 \geq 1 - 1/\sqrt{2}$. Let $\varepsilon_1, \varepsilon_2$ be such that

$$\begin{aligned} 1 - \varepsilon_1 &= \min_{z \in \Omega_Z} \text{tr}[\varrho_z \mathbf{M}_1(\{z\})], \quad \varrho_z = \frac{1}{2}(I + z\hat{\mathbf{z}} \cdot \boldsymbol{\sigma}), \\ 1 - \varepsilon_2 &= \min_{x \in \Omega_X} \text{tr}[\varrho_x \mathbf{M}_2(\{x\})], \quad \varrho_x = \frac{1}{2}(I + x\hat{\mathbf{x}} \cdot \boldsymbol{\sigma}). \end{aligned}$$

(Here ϱ_z, ϱ_x are the eigenstates of $\mathbf{Z}, \mathbf{X}$.) Then, since $\text{tr}[\varrho_z \mathbf{M}_1(\{z\})] \geq 1 - \varepsilon_1$ and $\text{tr}[\varrho_x \mathbf{M}_2(\{x\})] \geq 1 - \varepsilon_2$, Eq. (14.54) must hold, and therefore

$$\min_{z \in \Omega_Z} \text{tr}[\varrho_z \mathbf{M}_1(\{z\})] + \min_{x \in \Omega_X} \text{tr}[\varrho_x \mathbf{M}_2(\{x\})] \leq 1 + \frac{1}{\sqrt{2}}.$$

Finally, we can replace the eigenstate notation by the spectral projections of $\mathbf{Z}, \mathbf{X}$ to obtain

$$\min_{z \in \Omega_Z} \text{tr}[\mathbf{Z}(\{z\}) \mathbf{M}_1(\{z\})] + \min_{x \in \Omega_X} \text{tr}[\mathbf{X}(\{x\}) \mathbf{M}_2(\{x\})] \leq 1 + \frac{1}{\sqrt{2}}. \quad (14.55)$$

This is a bound on the proximity of $\mathbf{M}_1$ to $\mathbf{Z}$ and $\mathbf{M}_2$ to $\mathbf{X}$, measured in terms of the minimal “overlaps” of the associated effects, which are not allowed to become too large due to the competing properties of the noncommutativity of $\mathbf{Z}, \mathbf{X}$ and the compatibility of $\mathbf{M}_1, \mathbf{M}_2$. Note that in this example we see once more the close connection between preparation and measurement uncertainty, here expressed in terms of overall width and error bar width.

14.6 Appendix

Proof of Propositions 14.3 and 14.4

(a) If (D_1, D_2) is an admissible point, then also (D_2, D_1) is an admissible point.

Proof If $\mathbf{C}$ with $C_+ = \frac{1}{2}(c_0 I + \mathbf{c} \cdot \boldsymbol{\sigma})$ and $\mathbf{D}$ with $D_+ = \frac{1}{2}(d_0 I + \mathbf{d} \cdot \boldsymbol{\sigma})$ realise the distances D_1 and D_2 , respectively, then choose $c'_0, \mathbf{c}'$ and $d'_0, \mathbf{d}'$ as follows: $c'_0 = d_0$, $\mathbf{c}'$ has the length of $\mathbf{d}$ and its angle relative to $\mathbf{a}$ is equal to the angle between $\mathbf{d}$ and $\mathbf{b}$; similarly, $d'_0 = c_0$, $\mathbf{d}'$ has the length of $\mathbf{c}$ and its angle relative to $\mathbf{b}$ is the same as the angle between $\mathbf{c}$ and $\mathbf{a}$. This ensures that $(D'_1, D'_2) = (D_2, D_1)$. $\square$

(b) Assume that $(D_1, D_2) = (D_\infty(\mathbf{C}, \mathbf{A}), D_\infty(\mathbf{D}, \mathbf{B}))$ is an admissible point. As shown in Proposition 14.2, the joint measurability of $\mathbf{C}$ and $\mathbf{D}$ implies that of $\mathbf{C}^{(1)}$

and $\mathbf{D}^{(1)}$. Define $\mathbf{c}_0$ and $\mathbf{d}_0$ to be the projections of the vectors $\mathbf{c}$ and $\mathbf{d}$, respectively, onto the plane spanned by $\mathbf{a}$ and $\mathbf{b}$. Then

$$\|\mathbf{c} + \mathbf{d}\| \geq \|\mathbf{c}_0 + \mathbf{d}_0\|, \quad \|\mathbf{c} - \mathbf{d}\| \geq \|\mathbf{c}_0 - \mathbf{d}_0\|,$$

and hence, $\mathbf{C}_0^{(1)} = \{\frac{1}{2}(\mathbf{I} \pm \mathbf{c}_0 \cdot \boldsymbol{\sigma})\}$ and $\mathbf{D}_0^{(1)} = \{\frac{1}{2}(\mathbf{I} \pm \mathbf{d}_0 \cdot \boldsymbol{\sigma})\}$ are jointly measurable. One finds that

$$\begin{aligned} D_\infty(\mathbf{C}_0^{(1)}, \mathbf{A}) &\leq D_\infty(\mathbf{C}^{(1)}, \mathbf{A}) \leq D_1, \\ D_\infty(\mathbf{D}_0^{(1)}, \mathbf{B}) &\leq D_\infty(\mathbf{D}^{(1)}, \mathbf{B}) \leq D_2. \end{aligned}$$

We conclude that the best approximations are to be found from the subset of unbiased qubit observables, with vectors $\mathbf{c}$ and $\mathbf{d}$ in the plane spanned by $\mathbf{a}$ and $\mathbf{b}$.

(c) If (D_1, D_2) is an admissible point, then also (D'_1, D'_2) is an admissible point whenever $D_i \leq D'_i \leq \frac{1}{2}$.

Proof In view of (b) it is sufficient to show the result for admissible points which have realisations $(D_\infty(\mathbf{C}^{(1)}, \mathbf{A}), D_\infty(\mathbf{D}^{(1)}, \mathbf{B}))$. Thus let $\mathbf{C}^{(1)}, \mathbf{D}^{(1)}$ be two jointly measurable observables. Using Proposition 11.4, we note that also $\mathbf{C}^{(1)}$ and $\mathbf{D}_r^{(1)} = \{\frac{1}{2}(\mathbf{I} \pm r\mathbf{d} \cdot \boldsymbol{\sigma})\} \in \mathbf{E}^{1,rb}$ are jointly measurable for any $0 \leq r \leq 1$. Since the function

$$r \mapsto D_\infty(\mathbf{D}_r^{(1)}, \mathbf{B}) = \frac{1}{2} \|\mathbf{b} - r\mathbf{d}\|$$

is continuous, it takes all values between $D_\infty(\mathbf{D}^{(1)}, \mathbf{B})$ and $\frac{1}{2}$. We can similarly realise all values between $D_\infty(\mathbf{C}^{(1)}, \mathbf{A})$ and $\frac{1}{2}$. $\square$

(d) Observations (b) and (c) taken together entail Proposition 14.3. $\square$

(e) The admissible region is a convex set.

Proof Let (D_1, D_2) and (D'_1, D'_2) be realised by $\mathbf{C}, \mathbf{D}$ and $\mathbf{C}', \mathbf{D}'$, respectively. Let $t \in [0, 1]$. Then defining $\mathbf{C}_t, \mathbf{D}_t$ for $0 \leq t \leq 1$ and $c_{0,t} = tc_0 + (1-t)c'_0, \mathbf{c}_t = t\mathbf{c} + (1-t)\mathbf{c}'$ as well as $d_{0,t} = td_0 + (1-t)d'_0, \mathbf{d}_t = t\mathbf{d} + (1-t)\mathbf{d}'$, we obtain associated distances $D_{1,t} = D_\infty(\mathbf{C}_t, \mathbf{A})$ and $D_{2,t} = D_\infty(\mathbf{D}_t, \mathbf{B})$ which satisfy

$$D_{k,t} \leq tD_k + (1-t)D'_k, \quad k = 1, 2.$$

This together with (c) proves the claim. $\square$

(f) The set of admissible points is closed.

Proof The mapping

$$(\mathbf{c}, \mathbf{d}) \mapsto (D_\infty(\mathbf{C}^{(1)}, \mathbf{A}), D_\infty(\mathbf{D}^{(1)}, \mathbf{B})) = (\frac{1}{2}\|\mathbf{a} - \mathbf{c}\|, \frac{1}{2}\|\mathbf{b} - \mathbf{d}\|) \quad (14.56)$$

from $\mathbb{R}^3 \times \mathbb{R}^3$ to $\mathbb{R} \times \mathbb{R}$ is continuous. The set of admissible points is the image of the compact set

$$\{(c, d) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|c\| \leq 1, \|d\| \leq 1, \|c + d\| + \|c - d\| \leq 2\}, \quad (14.57)$$

hence it is itself closed and contains its boundary. This and (b) shows that for given $D_1 \in [0, \frac{1}{2}]$, there is a minimal number $D_2^*(D_1)$ such that all (D_1, D_2) with $D_2^*(D_1) \leq D_2 \leq \frac{1}{2}$ are admissible pairs while pairs with $D_2 < D_2^*(D_1)$ are not admissible. $\square$

(g) Since the admissible region is a convex set, the function $D_1 \mapsto D_2^*(D_1)$ is convex and therefore continuous. Due to (a), the curve is symmetric under reflection with respect to the line $D_1 = D_2$. We conclude that this function gives the lower boundary curve of the set of admissible points, and that the points on this curve are admissible. Together with (e) and (f), this completes the proof of Proposition 14.4.

14.7 Exercises

1. Show that the reduction of the linear map τ defined in (14.3) to $\mathcal{S}(\mathbb{C}^2)$ is a convex state automorphism induced by an antiunitary map (hence it is not completely positive), and construct an antiunitary map C such that $\tau(A) = CA^*C^*$ for all $A \in \mathcal{L}(\mathcal{H})$.
2. Prove Lemma 14.1.
3. Prove the preparation uncertainty relations (14.11), (14.12).
4. Determine the boundary of the uncertainty region (14.13).
5. Prove Lemma 14.2.
6. Find the minimum possible value for $\frac{1}{2}[F(2 - B) + B(2 - F)]$ and thus verify the lower bound of this quantity given in the first line of Eq. (14.31).
7. Work out the details of the proof of Corollary 14.1.
8. Prove Proposition 14.1.
9. Verify the statement of Remark 14.4.
10. Prove that inequality (14.42) entails the positivity of the operators in (14.40), which therefore constitute a triple joint observable.
Hint: Show that $\|g_{ijk}\|^2 = \|e_1\|^2 + \|e_2\|^2 + \|e_3\|^2 + 2(\alpha_{ijk} - 1)$. From this obtain an expression for α_{ijk} .
11. Determine the convex, symmetric boundary curve of the region of admissible points described in Proposition 14.4 and Fig. 14.1. (The solution can be found in [13].)
12. Prove the inequality (14.9), and show that it is equivalent to the Schrödinger uncertainty relation (12.11) for σ_1, σ_2 . Also deduce the inequality (14.10).
13. Using inequality (14.29), prove the following statement: any two ray effects $\mu P[\varphi], \mu P[\psi]$ are compatible if $\mu \leq \frac{2}{3}$; if $|\langle \varphi | \psi \rangle|^2 = \frac{1}{2}$ then $\mu \leq \frac{2}{3}$ is necessary for the compatibility of $\mu P[\varphi], \mu P[\psi]$.

References

1. Abbott, A.A., Alzieu, P.-L., Hall, M.J.W., Branciard, C.: Tight state-independent uncertainty relations for qubits. *Mathematics* **4**(4010008), 1–17 (2016)
2. Busch, P.: On the sharpness and bias of quantum effects. *Found. Phys.* **39**(7), 712–730 (2009)
3. Busch, P., Schmidt, H.-J.: Coexistence of qubit effects. *Quantum Inf. Process.* **9**(2), 143–169 (2010)
4. Stano, P., Reitzner, D., Heinosaari, T.: Coexistence of qubit effects. *Phys. Rev. A* **78**, 012315 (2008)
5. Gudder, S.: Coexistence of quantum effects. *Rep. Math. Phys.* **63**(2), 289–303 (2009)
6. Yu, S., Liu, N.-L., Li, L., Oh, C.H.: Joint measurement of two unsharp observables of a qubit. *Phys. Rev. A* **81**, 062116 (2010)
7. Heinosaari, T., Kiukas, J., Reitzner, D.: Coexistence of effects from an algebra of two projections. *J. Phys. A* **47**(22), 225301 (2014)
8. Busch, P.: Unsharp reality and joint measurements for spin observables. *Phys. Rev. D* (3) **33**(8), 2253–2261 (1986)
9. Pal, R., Ghosh, S.: Approximate joint measurement of qubit observables through an arthur-kelly model. *J. Phys. A* **44**(48), 485303 (2011)
10. S. Yu and C. Oh. Quantum contextuality and joint measurement of three observables of a qubit. [arXiv:1312.6470](https://arxiv.org/abs/1312.6470), 2013
11. Arthurs, E., Kelly, J.: On the simultaneous measurements of a pair of conjugate observables. *Bell Syst. Tech.* **44**, 725–729 (1965)
12. Busch, P., Lahti, P., Werner, R.F.: Heisenberg uncertainty for qubit measurements. *Phys. Rev. A* **89**, 012129 (2014)
13. Yu, S., Oh, C.: Optimal joint measurement of two observables of a qubit. [arXiv:1402.3785](https://arxiv.org/abs/1402.3785) (2014)
14. Bullock, T., Busch, P.: Measurement uncertainty: the problem of characterising optimal error bounds. [arXiv:1512.00104](https://arxiv.org/abs/1512.00104) (2015)

Chapter 15

Position and Momentum

Position and momentum are the most basic kinematic and dynamic variables of a quantum object. They reflect the structure of space and time where the object resides. In the nonrelativistic quantum theory, which is followed here, position and momentum are intertwined through the projective nature of the Weyl representation of space translations and velocity boosts. It is this intertwining which is responsible for the various no-go results concerning position and momentum, often expressed with notions like complementarity, uncertainty, and mutual disturbance.

This chapter is devoted to a study of the measurement theoretical consequences of the Weyl coupling of position and momentum. Particular attention is paid to their approximate joint measurements, Sect. 15.3, along the ideas developed in Chap. 13. To appreciate this problem, the basic properties of position and momentum will first be reviewed in Sect. 15.1. Their preparation uncertainty relations are discussed in Sect. 15.2. Section 15.4 concludes this chapter by showing a nontomographic method to determine both the position and momentum distributions from the statistics collected by a single measurement scheme. The actual measurement models will be discussed in Chap. 19.

15.1 The Weyl Pairs

The Defining Properties

Consider a spinless quantum object, with mass $m > 0$, moving in a one-dimensional Euclidean space $\mathbb{R}$, so that the Hilbert space of the object can be taken to be $\mathcal{H} = L^2(\mathbb{R})$, the position representation space. The (selfadjoint) position and momentum operators are the multiplication and differential operators $(Q\psi)(x) = x\psi(x)$ and $(P\psi)(x) = -i\hbar\psi'(x)$, respectively, with the spectral measures $X \mapsto \mathbf{Q}(X) = \mathbf{M}(X)$ and $Y \mapsto \mathbf{P}(Y) = \mathbf{M}_F(Y)$ of Sect. 8.1. The operators Q and P are the generators of

the representations $q \mapsto U(q) = e^{-i \frac{qP}{\hbar}}$ and $p \mapsto V(p) = e^{i \frac{pQ}{\hbar}}$ of space translations and velocity boosts. They are intertwined via commutation up to a factor,

$$U(q)V(p) = e^{-i \frac{qp}{\hbar}} V(p)U(q). \quad (15.1)$$

Taken together, these representations define a map

$$(q, p) \mapsto W(q, p) = e^{i \frac{qp}{2\hbar}} U(q)V(p),$$

with the property

$$W(q + q', p + p') = e^{i \frac{qp' - q'p}{2\hbar}} W(q, p)W(q', p'). \quad (15.2)$$

By Corollary 8.6, any $A \in \mathcal{L}(L^2(\mathbb{R}))$ that commutes with all the spectral projections of Q and P is a constant operator. This means that the map $(q, p) \mapsto W(q, p)$, with the multiplication law (15.2), constitutes an irreducible projective representation of the additive group $\mathbb{R}^2$ of phase space translations acting in $L^2(\mathbb{R})$. According to the von Neumann uniqueness theorem [1] this actually serves as a defining property of the pair (Q, P) .

Theorem 15.1 *Let $\mathcal{H}$ be a complex separable Hilbert space, and let A and B be any two selfadjoint operators acting in $\mathcal{H}$ such that their unitary groups $q \mapsto e^{iqA} = U^A(q)$ and $p \mapsto e^{-ipB} = V^B(p)$ satisfy the Weyl commutation relation (15.1). The pair (A, B) is unitarily equivalent to the pair (Q, P) , or to a (finite or countably infinite) direct sum of such pairs.*

The direct sum structure referred to above comes into play with spin- s objects, in which case the pair (A, B) is the $(2s + 1)$ -fold direct sum of the pair (Q, P) . We consider the case $s = 0$.

The defining property of the pair (Q, P) is the relation (15.1). In addition, position and momentum—as a Weyl pair—have the following important coupling properties:

$$QP - PQ = i\hbar I, \quad (15.3)$$

$$\inf\{\text{Var}(\mathbf{Q}_\varrho) \cdot \text{Var}(\mathbf{P}_\varrho) \mid \varrho \in \mathcal{S}(\mathcal{H})\} = \frac{\hbar^2}{4}, \quad (15.4)$$

$$\text{com}(\mathbf{Q}, \mathbf{P}) = \{0\}, \quad (15.5)$$

$$\mathbf{Q}(X) \wedge \mathbf{P}(Y) = \mathbf{Q}(X) \wedge \mathbf{P}(\mathbb{R} \setminus Y) = \mathbf{Q}(\mathbb{R} \setminus X) \wedge \mathbf{P}(Y) = 0, \quad (15.6)$$

$$\text{tr}[\mathbf{Q}(X)\mathbf{P}(Y)] = \frac{\hbar}{2\pi} \lambda(X)\lambda(Y). \quad (15.7)$$

The relation (15.3) is valid on a dense domain, and (15.6) and (15.7) hold for bounded Borel sets X and Y (Proposition 8.2). Recall that $\lambda(X)$ denotes the Lebesgue measure of X .

The commutation relation (15.3) as well as the preparation uncertainty relation (15.4) are well known and belong to the basic arsenal of quantum mechanics, the

latter being a special case of Proposition 15.1 below. We leave it as an exercise to show that the given lower bound for the uncertainty product $\text{Var}(\mathbf{Q}_\varrho) \cdot \text{Var}(\mathbf{P}_\varrho)$ can actually be reached and it is obtained exactly when ϱ is a Gaussian pure state, that is, ϱ is of the form $\varrho = |\varphi_{b,\theta}\rangle\langle\varphi_{b,\theta}|$, with

$$\varphi_{b,\theta}(x) = \left(\frac{b}{\pi}\right)^{\frac{1}{4}} e^{-\frac{1}{2}b(x-\bar{q})^2} e^{i(\theta+\bar{p}x/\hbar)}, \quad b, \theta \in \mathbb{R}, \quad (15.8)$$

where $\bar{q} = \langle\varphi_{b,\theta} | Q \varphi_{b,\theta}\rangle$ and $\bar{p} = \langle\varphi_{b,\theta} | P \varphi_{b,\theta}\rangle$.

As a consequence of Theorem 11.4, the *total noncommutativity* property (15.5) is equivalent to the statement that for no state ϱ , the map $(X, Y) \mapsto \text{tr}[\varrho \mathbf{Q}(X) \wedge \mathbf{P}(Y)]$ extends to a probability measure on $\mathcal{B}(\mathbb{R}^2)$. The relations (15.6) and (15.7) are sometimes taken as formalisations of the complementarity of these observables; again, both entail the impossibility of any joint measurement. A proof of the trace formula (15.7) is given, for instance, in [2] where it is also shown that a pair (Q, P) fulfilling this formula need not be a Weyl pair. Using the example of tilted quadratures, we will see that none of these five properties (15.3)–(15.7) is strong enough to enforce that (Q, P) is a Weyl pair.

Remark 15.1 The ‘exchange relation’ $QP - PQ = \frac{\hbar}{2\pi i}I$ discovered by Heisenberg in 1925 [3] was probably the most striking formula of the new quantum theory [3, 4]. In the case $\dim(\mathcal{H}) > 1$ such an operator equality implies that at least one of the operators Q and P is unbounded. This is the content of the Wintner–Wielandt theorem [5–7]. Another simple exercise is to show that (15.4) implies the unboundedness of both Q and P .

von Neumann’s uniqueness proof of 1931 [1] was a first characterisation of pairs of operators which satisfy the exchange relation (15.3) and also the Weyl commutation relation (15.1). Following Garrison and Wong [8], we say that a pair of selfadjoint operators (Q, P) is a *Heisenberg pair* if it satisfies the relation $QP - PQ = i\hbar I$ on a dense domain, and it is a *Weyl* or a *Schrödinger pair* if their unitary groups satisfy the relation (15.1). A Weyl pair is a Heisenberg pair but the converse is not necessarily true, the number-phase pair (N, Φ_{can}) being a well-known example (Sect. 16). A basic requirement for a Heisenberg pair (Q, P) to be a Weyl pair is that the domain of the commutator is invariant under Q and P , and the operator $Q^2 + P^2$ is essentially selfadjoint on that domain, that is, its least closed extension $(Q^2 + P^2)^{**}$ is selfadjoint. This is the content of the classic result of Rellich [9]; for a survey see, e.g., [7].

George Mackey’s paper [10] opened a new way of understanding the structure of physical quantities such as position and momentum. His imprimitivity theorem (e.g., [11]) characterises physical quantities through their behaviour under symmetry transformations; it entails that position and momentum observables, as spectral measures, are exactly those observables which are covariant under space translations and invariant under velocity boosts, and covariant under velocity boosts and invariant under space translations, respectively. $\triangleleft$

From now on we usually set $\hbar = 1$.

Commuting Functions of Q and P

Position and momentum are totally noncommutative: their commutativity domain contains only the null vector, Eq. (15.5). They do, however, have commuting spectral projections. Let $Q^g = Q \circ g^{-1}$ and $P^h = P \circ h^{-1}$ be function of Q and P , respectively, with $g : \mathbb{R} \rightarrow \mathbb{R}$, $h : \mathbb{R} \rightarrow \mathbb{R}$ being (Borel) functions. The associated selfadjoint operators are $Q^g[1] = g(Q)$ and $P^h[1] = h(P)$. The following result characterises functions g and h for which $Q^g(X)P^h(Y) = P^h(Y)Q^g(X)$ for all X and Y .

Theorem 15.2 *Let g and h be essentially bounded Borel functions such that neither $g(Q)$ nor $h(P)$ is a constant operator. The functions Q^g of position and P^h of momentum commute if and only if g and h are both periodic with minimal positive periods a, b satisfying $\frac{2\pi}{ab} \in \mathbb{N}$.*

Proof See [12–15] where the result is proven in increasing generality. $\square$

If Q^g and P^h are commuting observables, then they have the joint observable M , with $M(X \times Y) = Q^g(X)P^h(Y)$, meaning that Q^g and P^h can be measured jointly. The price for this restricted form of joint measurability of position and momentum is that they are to be coarse-grained by periodic functions g and h with appropriately related minimal periods a, b . We leave it as an exercise to show that in such a case the errors $\Delta_\alpha(Q, Q^g)$ and $\Delta_\beta(P, P^h)$ and the corresponding error bar widths are infinite.

The functions g and h can be chosen as the characteristic functions of appropriate periodic sets. This allows one to model a situation known in solid state physics, where an electron in a crystal can be confined arbitrarily closely to the atoms while at the same time its momentum is localised arbitrarily closely to the reciprocal lattice points, see e.g. [16].

Pairs of Tilted Quadratures

Position and momentum can be obtained in a smooth way from each other. Indeed, let $U_\theta = e^{i\theta H}$, $\theta \in \mathbb{R}$, be the unitary operator defined by the oscillator operator H , the unique selfadjoint extension of the essentially selfadjoint operator $\frac{1}{2}(Q^2 + P^2)$, and define $Q_\theta = U_\theta Q U_\theta^*$, so that Q_θ is the quadrature operator, with the spectral measure $X \mapsto Q_\theta(X) = U_\theta Q(X) U_\theta^*$. Clearly $U_\theta^* = U_{-\theta}$, $Q_0 = Q$ and, since the Fourier–Plancherel operator $F = U_{-\pi/2}$, one also has $Q_{\pi/2} = P$. In addition, Q_θ is the closure of the essentially selfadjoint operator $Q \cos \theta + P \sin \theta$; see Exercise 8 of Sect. 15.6.

To study the coupling properties of any two quadratures (Q_α, Q_β) , it is sufficient to consider the tilted pair (Q, Q_θ) , $\theta \in [0, 2\pi)$, since for any pair (Q_α, Q_β) one finds a unitary operator $U = U_\alpha$ such that $Q_\alpha = U Q U^*$ and $Q_\beta = U Q_{\beta-\alpha} U^*$, that is, the pair (Q_α, Q_β) is unitarily equivalent to the pair (Q, Q_θ) where $\theta = \beta - \alpha$.

Using the operator relation (see, for instance, [17, Proposition 1.1])¹

¹For mathematical accuracy we note that the unitary operator in the left hand side of (15.9) is actually defined by the unique selfadjoint extension of the essentially selfadjoint operator shown in the exponent, see Sect. 5.1. In this book we follow the common practice of omitting the closure sign where no confusion should arise.

$$e^{iy(Q \cos \theta + P \sin \theta)} = e^{iyQ \cos \theta} e^{iyP \sin \theta} e^{iy^2(\cos \theta \sin \theta)/2} \quad (15.9)$$

together with the Weyl relation, one checks that the unitary operators e^{ixQ} and e^{iyQ_θ} fulfil the Weyl relation exactly when $\theta = \pi/2$, that is, $Q_\theta = P$.

Clearly, for any pair of quadratures (Q, Q_θ) ,

$$QQ_\theta - Q_\theta Q = i \sin \theta I, \quad (15.10)$$

on the dense domain of the commutator. Moreover, since for any $\theta \in [0, 2\pi)$ for which $\sin \theta \neq 0$, the pair $(Q, \frac{1}{\sin \theta} Q_\theta)$ is a Weyl pair, one also has the relations

$$\inf\{\text{Var}(\mathbf{Q}_\varrho) \cdot \text{Var}(\mathbf{Q}_{\theta, \varrho}) \mid \varrho \in \mathcal{S}(\mathcal{H})\} = \frac{\sin^2 \theta}{4} > 0, \quad (15.11)$$

$$\text{com}(\mathbf{Q}, \mathbf{Q}_\theta) = \{0\}, \quad (15.12)$$

$$\mathbf{Q}(X) \wedge \mathbf{Q}_\theta(Y) = \mathbf{Q}(X) \wedge \mathbf{Q}_\theta(\mathbb{R} \setminus Y) = \mathbf{Q}(\mathbb{R} \setminus X) \wedge \mathbf{Q}_\theta(Y) = 0, \quad (15.13)$$

$$\text{tr}[\mathbf{Q}(X)\mathbf{Q}_\theta(Y)] = \frac{1}{2\pi} \lambda(X)\lambda(Y), \quad (15.14)$$

corresponding to the relations (15.4)–(15.7). In addition, from Theorem 15.1 one sees that, if $\theta \neq 0$ and $\theta \neq \pi$, there exists a unitary operator Y_θ on $L^2(\mathbb{R})$ such that

$$\mathbf{Q}(X) = Y_\theta \mathbf{Q}(X) Y_\theta^*, \quad U_\theta \mathbf{Q}(\sin \theta \cdot X) U_\theta^* = Y_\theta \mathbf{P}(X) Y_\theta^*$$

for any Borel set $X \subset \mathbb{R}$. Especially, Y_θ commutes with Q so that it is a multiplication operator, i.e.

$$(Y_\theta \psi)(x) = e^{i\alpha_\theta(x)} \psi(x)$$

where $\alpha_\theta : \mathbb{R} \rightarrow [0, 2\pi)$ is a measurable function, see Proposition 8.1. Note that $Y_{\pi/2} = I$, i.e. $\alpha_{\pi/2} = 0$, and $U_{\pi/2} \mathbf{Q}(X) U_{\pi/2}^* = \mathbf{P}(X)$.

Position and Momentum—Incompatibility and Measurement Disturbance

As a consequence of Proposition 4.8 and Theorem 11.1 position and momentum are incompatible; they have no joint observable and thus cannot be measured jointly. Moreover, a sharp measurement of one of these observables severely disturbs the other (cf. Proposition 11.1). For instance, if $\mathcal{M}^\mathcal{Q}$ is any position measurement, with the instrument $\mathcal{I}^\mathcal{Q}$, then the sequential observable $\mathbf{E}$ defined by $\mathcal{M}^\mathcal{Q}$ followed by a momentum measurement is, by Eq. (10.14) and Proposition 4.8, of the product form

$$\mathbf{E}(X \times Y) = \mathcal{I}^\mathcal{Q}(X)^*(\mathbf{P}(Y)) = \mathbf{Q}(X) \mathcal{I}^\mathcal{Q}(\mathbb{R})^*(\mathbf{P}(Y)),$$

where the distorted momentum effects $\mathcal{I}^\mathcal{Q}(\mathbb{R})^*(\mathbf{P}(Y))$ are functions of Q (due to the maximality of Q). The distorted momentum observable $Y \mapsto \mathcal{I}^\mathcal{Q}(\mathbb{R})^*(\mathbf{P}(Y))$ is invariant under velocity boosts, and therefore does not carry the most fundamental property of a momentum, the covariance under velocity boosts. In this sense

one may say that a measurement of position (momentum) completely destroys all the information about momentum (position) encoded in the initial state. If $(X, Y) \mapsto \mathfrak{I}^Q(X)^*(P(Y))$ is considered as an attempted joint measurement of position and momentum, combining an accurate position measurement with an inaccurate momentum measurement, the conclusion is that $Y \mapsto \mathfrak{I}^Q(\mathbb{R})^*(P(Y))$ is in fact unsuitable as an approximate momentum measurement. This statement will be made quantitative in Remark 15.3.

15.2 Preparation Uncertainty Relations for $\mathbf{Q}$ and $\mathbf{P}$

We now formulate and prove two versions of the preparation uncertainty relations for position and momentum, the first refers to the α -spreads, the second to the overall widths of probability measures.

Proposition 15.1 *For any state ϱ and $\alpha, \beta \in [1, \infty)$,*

$$\Delta_\alpha(\mathbf{Q}_\varrho)\Delta_\beta(\mathbf{P}_\varrho) \geq c_{\alpha\beta}\hbar. \quad (15.15)$$

The constant $c_{\alpha\beta}$ is connected to the ground state energy $g_{\alpha\beta}$ of the Hamiltonian $H_{\alpha\beta}$, the closure of the essentially selfadjoint operator $|Q|^\alpha + |P|^\beta$, by the equation

$$c_{\alpha\beta} = \alpha^{\frac{1}{\beta}}\beta^{\frac{1}{\alpha}} \left(\frac{g_{\alpha\beta}}{\alpha + \beta} \right)^{\frac{1}{\alpha} + \frac{1}{\beta}}. \quad (15.16)$$

The lower bound is attained exactly when ϱ arises from the ground state of the operator $H_{\alpha\beta}$ by phase space translation and dilatation. For $\alpha = \beta = 2$, $H_{\alpha\beta}$ is twice the harmonic oscillator Hamiltonian with ground state energy $g_{22} = 1$, and $c_{22} = 1/2$.

Proof This is an adaption of [18, Proposition 12]. The Hamiltonian $H_{\alpha\beta}$ appears here mainly through the quadratic form $\langle \psi | H_{\alpha\beta} | \psi \rangle$ where ψ runs over, say, the unit vectors in the Schwartz space of tempered functions. The inequality (15.15) depends only on the lower bound $g_{\alpha\beta}$ of this form.

Consider the family of operators

$$H_{\alpha\beta}(p, q, \lambda) = \lambda^\alpha |Q - qI|^\alpha + \lambda^{-\beta} |P - pI|^\beta \geq g_{\alpha\beta} I, \quad (15.17)$$

obtained from $H_{\alpha\beta}$ by shifting in phase space by (q, p) , and by a dilatation $(Q, P) \mapsto (\lambda Q, \lambda^{-1} P)$. Since these operations are unitarily implemented, the lower bound $g_{\alpha\beta}$ for all these operators is independent of p, q, λ . Now, for a given ϱ , we may assume that $\Delta_\alpha(\mathbf{Q}_\varrho)$ and $\Delta_\beta(\mathbf{P}_\varrho)$ are both finite, since these uncertainties do not vanish for any state, and one infinite factor hence renders the inequality trivial. Let q be the point for which $\Delta_\alpha(\mathbf{Q}_\varrho, \delta_q)$ attains its minimum $\Delta_\alpha(\mathbf{Q}_\varrho)$, and choose p similarly for P .

Then by taking the expectation of (15.17) with ϱ we obtain the additive uncertainty relation

$$\lambda^\alpha \Delta_\alpha(\mathbf{Q}_\varrho)^\alpha + \lambda^{-\beta} \Delta_\beta(\mathbf{P}_\varrho)^\beta \geq g_{\alpha\beta}. \quad (15.18)$$

The minimum of the left hand side with respect to λ is attained at

$$\lambda = \left(\frac{\beta \Delta_\beta(\mathbf{P}_\varrho)^\beta}{\alpha \Delta_\alpha(\mathbf{Q}_\varrho)^\alpha} \right)^{1/(\alpha+\beta)}. \quad (15.19)$$

Inserting this into (15.18) gives an expression that depends only on the uncertainty product $u = \Delta_\alpha(\mathbf{Q}_\varrho) \Delta_\beta(\mathbf{P}_\varrho)$, namely

$$u^{\alpha\beta/(\alpha+\beta)} \alpha^{-\alpha/(\alpha+\beta)} \beta^{-\beta/(\alpha+\beta)} (\alpha + \beta) \geq g_{\alpha\beta}. \quad (15.20)$$

Now solving for u gives the uncertainty inequality. Moreover, since the left hand side is still nothing but the expectation of $H_{\alpha\beta}(p, q, \lambda)$ with a suitable choice of parameters, equality holds exactly if ϱ is the ground state state of $H_{\alpha\beta}(p, q, \lambda)$. But since this operator arises by dilatation and shifts from $H_{\alpha\beta}$, its ground state must arise by the same operations from the ground state of $H_{\alpha\beta}$.

To close the proof we verify that there are eigenvalues at the bottom of the spectrum of $H_{\alpha\beta}$. For $\alpha = \beta = 2$ this is well known whereas the case $\alpha = \beta = 1$ is treated in [19]. The fact that $H_{\alpha\beta}$ has a purely discrete spectrum with eigenvalues going to infinity can be verified by the Golden–Thompson inequality according to which, for any $\lambda \geq 0$,

$$\mathrm{tr}[e^{-\lambda H_{\alpha\beta}}] \leq \mathrm{tr}[e^{-\lambda|\mathcal{Q}|^\alpha} e^{-\lambda|\mathcal{P}|^\beta}] = \int e^{-\lambda|q|^\alpha} dq \int e^{-\lambda|p|^\beta} dp < \infty \quad (15.21)$$

(see, e.g., [20, p. 94]). Hence, the positive operator on the left is trace class (and therefore compact), and consequently, the spectrum of the selfadjoint generator $-H_{\alpha\beta}$ of the semigroup $e^{-\lambda H_{\alpha\beta}}$, $\lambda > 0$, consists of a countable discrete set of eigenvalues each of finite multiplicity [21, Theorem 2.20]. (Alternatively, the latter fact is a straightforward consequence of Theorem 3.5, Definition 5.6, and Corollary 5.2.) Since each of the terms in $H_{\alpha\beta}$ already has a strictly positive expectation in any state, the lowest eigenvalue $g_{\alpha\beta}$ is strictly positive. $\square$

Remark 15.2 The inequality (15.15) is equivalent to the *additive* trade-off relation (15.18). We state this relation here for $\alpha = \beta = 2$, in which case (15.15) reduces to (12.14):

$$\lambda^2 \Delta(\mathbf{Q}_\varrho)^2 + \lambda^{-2} \Delta(\mathbf{P}_\varrho)^2 \geq 1. \quad (15.22)$$

$\triangleleft$

We now turn to the overall width uncertainty relation for Q and P. We start with recalling that by (15.6), for any two bounded intervals X and Y and for any vector state ψ one has $\mathbf{Q}_\psi(X) + \mathbf{P}_\psi(Y) < 2$. Moreover, for any such intervals, one can

construct a vector state ψ_0 for which the sum of the probabilities $\mathbf{Q}_{\psi_0}(X)$ and $\mathbf{P}_{\psi_0}(Y)$ attains its maximum value. This result follows from the work of Landau and Pollak [22] and Lenard [23] (cf. Proposition 12.3).

Proposition 15.2 *For any vector state ψ and for any bounded intervals X and Y ,*

$$\mathbf{Q}_{\psi}(X) + \mathbf{P}_{\psi}(Y) \leq 1 + \sqrt{a_0} < 2, \quad (15.23)$$

where a_0 is the largest eigenvalue of the operator $\mathbf{Q}(X)\mathbf{P}(Y)\mathbf{Q}(X)$, which is positive and trace class. There exists an optimising vector state φ_0 such that

$$\mathbf{Q}_{\varphi_0}(X) + \mathbf{P}_{\varphi_0}(Y) = 1 + \sqrt{a_0}. \quad (15.24)$$

We will say that position $\mathbf{Q}$ is ε -localised in an interval X for a given state ψ whenever $\mathbf{Q}_{\psi}(X) \geq 1 - \varepsilon$ for some (preferably small) $\varepsilon \in (0, 1)$, and similarly for momentum. Then Eq. (15.24) describes the maximum degree of approximate localisation that can be achieved in any phase space cell of given size $|X| \cdot |Y|$.

The largest eigenvalue a_0 is invariant under a scale transformation applied simultaneously to $\mathbf{Q}$ and $\mathbf{P}$; it is therefore a function of the product $|X| \cdot |Y|$ of the interval lengths $|X|$ and $|Y|$. From (15.7) one gets $\text{tr}[\mathbf{Q}(X)\mathbf{P}(Y)\mathbf{Q}(X)] = |X| \cdot |Y|/(2\pi\hbar)$, so that we obtain

$$|X| \cdot |Y| \geq 2\pi\hbar \cdot a_0. \quad (15.25)$$

If position and momentum are both approximately localised within X and Y , respectively, so that $\mathbf{Q}_{\psi}(X) \geq 1 - \varepsilon_1$ and $\mathbf{P}_{\psi}(Y) \geq 1 - \varepsilon_2$, then due to inequality (15.23), one must have $1 - \varepsilon_1 - \varepsilon_2 \leq \sqrt{a_0}$, and then (15.25) implies:

$$|X| \cdot |Y| \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2 \quad (15.26)$$

if $\mathbf{Q}_{\psi}(X) \geq 1 - \varepsilon_1$, $\mathbf{P}_{\psi}(Y) \geq 1 - \varepsilon_2$, and $\varepsilon_1 + \varepsilon_2 < 1$. We summarise these observations in the form of an overall width uncertainty relation for position and momentum.

Proposition 15.3 *For any vector state ψ and for each $\varepsilon_1, \varepsilon_2 \in (0, 1)$, if $\varepsilon_1 + \varepsilon_2 < 1$, then*

$$W_{\varepsilon_1}(\mathbf{Q}_{\psi}) \cdot W_{\varepsilon_2}(\mathbf{P}_{\psi}) \geq 2\pi\hbar \cdot c(\varepsilon_1, \varepsilon_2), \quad (15.27)$$

where

$$c(\varepsilon_1, \varepsilon_2) = (1 - \varepsilon_1 - \varepsilon_2)^2.$$

If $\varepsilon_1 + \varepsilon_2 \geq 1$, the product of widths has no positive lower bound [22]. In the case $\varepsilon_1 + \varepsilon_2 < 1$, the inequality is tight in the sense that even for fairly small values of $\varepsilon_1, \varepsilon_2$, the product of overall widths can be of the order of $2\pi\hbar$; we quote a numerical example given in [22]: if $\varepsilon_1 = \varepsilon_2 = .01$, then $|X| \cdot |Y|$ can still be as small as $6.25 \times (2\pi\hbar)$.

The bound of inequality (15.27) is not tight; it can be improved upon by a somewhat greater (but still not tight) bound [24]:

$$c'(\varepsilon_1, \varepsilon_2) = \left(\sqrt{(1 - \varepsilon_1)(1 - \varepsilon_2)} - \sqrt{\varepsilon_1 \varepsilon_2} \right)^2.$$

15.3 Approximate Joint Measurements of $\mathbf{Q}$ and $\mathbf{P}$

Covariant phase space observables constitute an important instance of approximate joint measurements of position and momentum. Let $\mathbf{G}^K$ be the phase space observable generated by a positive trace one operator K . Its margins are the unsharp position $\mu * \mathbf{Q}$ and the unsharp momentum $\nu * \mathbf{P}$, where $\mu = \mathbf{Q}_{\Pi K \Pi}$ and $\nu = \mathbf{P}_{\Pi K \Pi}$ are the position and momentum distributions defined by the state $\Pi K \Pi$ (where Π is the parity operator).

The Wasserstein deviations are in this case

$$\Delta_2(\mathbf{Q}, \mu * \mathbf{Q}) = \sqrt{\mu[2]}, \quad \Delta_2(\mathbf{P}, \nu * \mathbf{P}) = \sqrt{\nu[2]}, \quad (15.28)$$

and they obey the measurement uncertainty relation (cf. Proposition 13.1)

$$\Delta_2(\mathbf{Q}, \mu * \mathbf{Q}) \Delta_2(\mathbf{P}, \nu * \mathbf{P}) \geq \Delta(\mu) \Delta(\nu) = \Delta(\mathbf{Q}_{\Pi K \Pi}) \Delta(\mathbf{P}_{\Pi K \Pi}) \geq \frac{1}{2} \hbar.$$

The first inequality becomes an equality when K is such that $\mu[1] = \nu[1] = 0$, whereas the lower bound $\hbar/2$ is obtained if, in addition, K is a one-dimensional projection of the Gaussian form $K = |\varphi_{b,\theta}\rangle \langle \varphi_{b,\theta}|$ of Eq. (15.8), with $\bar{q} = \bar{p} = 0$.

The main result of this section states that there are no better approximate joint measurements of position and momentum than those given by the covariant phase space measurements. To formulate this result, let $\mathbf{M} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ be any phase space observable, that is, an observable defined on the phase space $\mathbb{R}^2$. We consider its Cartesian margins $\mathbf{M}_1$ and $\mathbf{M}_2$ as approximate position and approximate momentum observables. To emphasise that we also denote them as $\mathbf{M}^q$ and $\mathbf{M}^p$, respectively.

Theorem 15.3 *Let $\mathbf{M}$ be a phase space observable. Then, for any $1 \leq \alpha, \beta < \infty$,*

$$\Delta_\alpha(\mathbf{Q}, \mathbf{M}^q) \Delta_\beta(\mathbf{P}, \mathbf{M}^p) \geq c_{\alpha\beta} \hbar, \quad (15.29)$$

provided that the quantities on the left hand side are finite. The constants $c_{\alpha\beta}$ are the same as in Proposition 15.1.

The assumption that both $\Delta_\alpha(\mathbf{Q}, \mathbf{M}^q)$ and $\Delta_\beta(\mathbf{P}, \mathbf{M}^p)$ are finite rules out the indefinite product $0 \cdot \infty$, along with the case $\infty \cdot \infty$. Examples for the indefinite case can be given quite easily. It suffices to combine an ideal position measurement with a random momentum output. Although the statement given here seemingly excludes the indefinite case, it is actually the best one can say about it: If the uncertainty

relation is to express quantitatively that not both $\Delta_\alpha(\mathbf{Q}, \mathbf{M}^q)$ and $\Delta_\beta(\mathbf{P}, \mathbf{M}^p)$ can become small, then we should also have that if one is zero, the other must be infinite. But this statement is implied by Theorem 15.3, which shows that the case 0-finite does not occur (see Remark 15.3). Of course, we can also conclude that if in some process one uncertainty tends to zero the other has to diverge in keeping with the above theorem.

The proof of Theorem 15.3 is based on the following lemma which reduces the general case to a covariant one. Its proof is given in the appendix to this section. This material is based on [18].

Lemma 15.1 *Let $\mathbf{M}$ be a phase space observable and $1 \leq \alpha, \beta < \infty$. Suppose that $\Delta_\alpha(\mathbf{Q}, \mathbf{M}^q)$ and $\Delta_\beta(\mathbf{P}, \mathbf{M}^p)$ are both finite. Then there is a covariant phase space observable $\overline{\mathbf{M}}$ such that*

$$\Delta_\alpha(\mathbf{Q}, \overline{\mathbf{M}}^q) \leq \Delta_\alpha(\mathbf{Q}, \mathbf{M}^q) \quad \text{and} \quad \Delta_\beta(\mathbf{P}, \overline{\mathbf{M}}^p) \leq \Delta_\beta(\mathbf{P}, \mathbf{M}^p).$$

Remark 15.3 Assume now that the approximate joint observable $\mathbf{M} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ is such that $\Delta_\alpha(\mathbf{Q}, \mathbf{M}^q) = 0$, for instance. Then $\mathbf{M}^q = \mathbf{Q}$ (exercise), and thus $\mathbf{M}$ is of the product form $\mathbf{M}(X \times Y) = \mathbf{Q}(X)\mathbf{M}^p(Y)$, where the effect operator $\mathbf{M}^p(Y)$ is a function of the position operator Q , that is, $\mathbf{M}^p(Y) = f_Y(Q)$ for some measurable function $f_Y : \mathbb{R} \rightarrow \mathbb{R}$. Then necessarily $\Delta_\beta(\mathbf{P}, \mathbf{M}^p) = \infty$, for otherwise there would exist a covariant phase space observable with sharp position $\mathbf{M}^q(X) = \mathbf{Q}(X)$ and the distorted (noncovariant) momentum $\mathbf{M}^p(Y) = f_Y(Q)$ as the marginal observables, which is not the case. $\triangleleft$

Remark 15.4 By Eq. (13.17), the errors $\Delta_\alpha(\mathbf{Q}, \overline{\mathbf{M}}^q)$ and $\Delta_\beta(\mathbf{P}, \overline{\mathbf{M}}^p)$ coincide with the calibration errors $\Delta_\alpha^c(\mathbf{Q}, \overline{\mathbf{M}}^q)$ and $\Delta_\beta^c(\mathbf{P}, \overline{\mathbf{M}}^p)$. Though, in general, the calibration errors are smaller than the metric errors, it can be shown [18, Theorem 16], that a result corresponding to Theorem 15.3 is obtained also for the calibration errors: for any $1 \leq \alpha, \beta < \infty$, if $\mathbf{M}$ is a phase space observable such that both $\Delta_\alpha^c(\mathbf{Q}, \mathbf{M}^q)$ and $\Delta_\beta^c(\mathbf{P}, \mathbf{M}^p)$ are finite, then

$$\Delta_\alpha^c(\mathbf{Q}, \mathbf{M}^q) \Delta_\beta^c(\mathbf{P}, \mathbf{M}^p) \geq c_{\alpha\beta} \hbar. \quad (15.30)$$

$\triangleleft$

Theorem 15.3 has an important corollary concerning the joint measurability of a pair of unsharp position and momentum observables $(\mu * \mathbf{Q}, \nu * \mathbf{P})$. This result has been proved in great detail in [25] following the averaging technique developed in [19].

Corollary 15.1 *Let $\mu * \mathbf{Q}$ and $\nu * \mathbf{P}$ be any unsharp position and momentum observables with $\mu[2], \nu[2]$ finite. If there is an observable $\mathbf{M} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{M}^q = \mu * \mathbf{Q}$ and $\mathbf{M}^p = \nu * \mathbf{P}$, then there is also a covariant phase space observable $\mathbf{G}^K : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{G}_1^K = \mu * \mathbf{Q}$ and $\mathbf{G}_2^K = \nu * \mathbf{P}$.*

Proof By the proof of Lemma 15.1 the averaged covariant phase space observable $\bar{M}$ is normalised so that it only remains to check that $\bar{M}^q = \mu * Q$ and $\bar{M}^p = \nu * P$ which implies that $\mu = Q_{\Pi K \Pi}$ and $\nu = P_{\Pi K \Pi}$ for some positive trace one operator K . $\square$

We have seen that the general measurement uncertainty relation for position and momentum in terms of Wasserstein deviations follows by reduction to the special case of covariant phase space observables, which, in turn, resulted from a corresponding preparation uncertainty relation, Proposition 13.1. In the same way, such measurement uncertainty relations also hold for general phase space observables and error bar widths [26]. This follows again by reduction to the case of covariant phase space observables, stated in Proposition 13.4, which itself is a consequence of the associated preparation uncertainty relation for overall widths, Proposition 15.3.

Theorem 15.4 *Let M be a phase space observable with margins $M_1 = M^q$, $M_2 = M^p$. Then for $\varepsilon_1, \varepsilon_2 \in (0, 1)$ with $\varepsilon_1 + \varepsilon_2 < 1$,*

$$\mathcal{W}_{\varepsilon_1}(M^q, Q) \mathcal{W}_{\varepsilon_2}(M^p, Q) \geq 2\pi\hbar(1 - \varepsilon_1 - \varepsilon_2)^2. \quad (15.31)$$

provided that the quantities on the left hand side are finite.

Note that it is the finiteness of the Wasserstein deviations and error bar widths between the margins M^q, M^p of M and Q, P , respectively, that ensures the existence of a covariant phase space observable $\bar{M}$ for which the approximation is no worse. As shown in Proposition 13.6, finite Wasserstein distance always implies finite error bar width. It is a simple exercise to identify (even covariant) phase space observables with finite error bar widths but infinite values of the Wasserstein deviations. Hence the criterion of finite error bar widths is more liberal, admitting a wider class of phase space observables M for which the existence of a covariant $\bar{M}$ as a good joint approximation for position and momentum can be achieved.

15.4 Measuring Q and P with a Single Measurement Scheme

Position and momentum do not admit any joint measurement, and any approximate joint measurement of Q and P satisfies a trade-off relation of the form (15.29). Still it is possible to determine the position and momentum distributions Q_ϱ and P_ϱ from the statistics collected by a single measurement. Consider an approximate joint measurement of Q and P realised as a joint measurement of the unsharp observables $\mu * Q$ and $\nu * P$. By Corollary 15.1, the measures μ and ν are Fourier-related and obtained from a common generating state K . Thus, for any initial state ϱ , the actual measured statistics is $(\mu * Q_\varrho, \nu * P_\varrho)$. If one is able to solve (Q_ϱ, P_ϱ) from $(\mu * Q_\varrho, \nu * P_\varrho)$, then the measurement serves, for that state, also as an indirect (Q, P) -measurement.

In this section the statistical method of moments will be used to solve this problem for a large class of initial states ϱ . Interestingly, this can be achieved without a full state reconstruction (which we study in Chap. 18). We review a few basic results on the method of moments.

How Much Statistics Is Needed to Determine a Distribution?

Consider a generic probability measure $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$. Its expectation $\mu[1]$ never determines the distribution μ , but if its variance is zero, that is, $\mu[2] = \mu[1]^2$, then μ is a point measure concentrated at $\mu[1]$. As a rule one needs all the moment information on $\mu[n]$, $n \in \mathbb{N}$, to determine μ , and even this is not always sufficient. The following question is the uniqueness part of the classic Hamburger moment problem:

When does the moment sequence $(\mu[n])_{n \in \mathbb{N}}$ determine μ ?

A measure μ is said to be *determinate* if it is determined by its moment sequence $(\mu[n])_{n \in \mathbb{N}}$. The possible occurrence of nonuniqueness is a typical quantum mechanical situation, as is evident in the following example.

Example 15.1 Let $\varphi_1, \varphi_2 \in L^2(\mathbb{R})$ be any two compactly supported smooth functions of unit norm and with disjoint supports, as they occur, for instance, in a double-slit experiment. Then, for any $n \in \mathbb{N}$, $\varphi_1, \varphi_2 \in \mathcal{D}(Q^n) \cap \mathcal{D}(P^n)$. For any unit vector $\psi_\delta = 2^{-1/2}(\varphi_1 + e^{i\delta}\varphi_2)$, $\delta \in \mathbb{R}$, the moments of the position and momentum distributions $|\psi_\delta|^2$ and $|\hat{\psi}_\delta|^2$ do not depend on δ . While the position distribution does not depend on δ , the momentum distribution is δ -dependent (which is responsible for the interference effect observable in a double-slit experiment). Indeed, the Fourier transform $\hat{\psi}_\delta$ of ψ_δ is continuous and, by the uniqueness theorem of analytic functions, its support is the whole real line, $\text{supp } \hat{\psi}_\delta = \text{supp } \hat{\varphi}_1 = \text{supp } \hat{\varphi}_2 = \mathbb{R}$. Therefore, the momentum density

$$p \mapsto |\hat{\psi}_\delta(p)|^2 = \frac{1}{2} \left[|\hat{\varphi}_1(p)|^2 + |\hat{\varphi}_2(p)|^2 + 2 \operatorname{Re} \left(\overline{\hat{\varphi}_1(p)} \hat{\varphi}_2(p) e^{i\delta} \right) \right]$$

depends on δ , which shows that the moments $\int p^n |\hat{\psi}_\delta(p)|^2 dp = \langle \psi_\delta | P^n \psi_\delta \rangle$, $n \in \mathbb{N}$, do not determine the distribution $p \mapsto |\hat{\psi}_\delta(p)|^2$. $\triangleleft$

A useful sufficient condition for the sequence $(\mu[n])_{n \in \mathbb{N}}$ to determine μ is the following: if the measure μ is exponentially bounded, that is, there is an $a > 0$ such that $\int e^{a|x|} d\mu(x) < \infty$, then μ is uniquely determined by its moments $\mu[n]$, $n \in \mathbb{N}$ (see e.g. [27, Theorems II.4.3 and II.5.2]). Exponential boundedness of μ admits a simple operational characterisation.

Proposition 15.4 *The probability measure μ is exponentially bounded if and only if for some $C, R > 0$ the moment sequence $(\mu[n])_{n \in \mathbb{N}}$ satisfies*

$$|\mu[n]| \leq C R^n n! \tag{15.32}$$

for all $n \in \mathbb{N}$.

Proof Assume that for some $a > 0$ and $C > 0$, $\int e^{a|x|} d\mu = C$. Then by the Lebesgue dominated convergence theorem we also have $|\mu[n]| \leq C(\frac{1}{a})^n n!$ for all $n \in \mathbb{N}$. The converse implication is proved in [28, Proposition 1.5]. $\square$

Indirect $(\mathbf{Q}, \mathbf{P})$ -Measurement

Consider the measured statistics $(\mu * \mathbf{Q}_\varrho, \nu * \mathbf{P}_\varrho)$, and assume that the convolving measures μ and ν are such that all their moments $\mu[n]$ and $\nu[n]$, $n \in \mathbb{N}$, are finite. By Proposition 8.3, all the moments of $\mu * \mathbf{Q}_\varrho$, $\nu * \mathbf{P}_\varrho$ can then be computed for all states ϱ whose moments $\mathbf{Q}_\varrho[k]$ and $\mathbf{P}_\varrho[k]$, $k \in \mathbb{N}$, are finite. They are

$$\begin{aligned} (\mu * \mathbf{Q}_\varrho)[k] &= \sum_{n=0}^k \binom{k}{n} \mu[k-n] \mathbf{Q}_\varrho[n], \\ (\nu * \mathbf{P}_\varrho)[k] &= \sum_{n=0}^k \binom{k}{n} \nu[k-n] \mathbf{P}_\varrho[n], \end{aligned}$$

from which one obtains the recursion formulas

$$\begin{aligned} \mathbf{Q}_\varrho[k] &= (\mu * \mathbf{Q}_\varrho)[k] - \sum_{n=0}^{k-1} \binom{k}{n} \mu[k-n] \mathbf{Q}_\varrho[n], \\ \mathbf{P}_\varrho[k] &= (\nu * \mathbf{P}_\varrho)[k] - \sum_{n=0}^{k-1} \binom{k}{n} \nu[k-n] \mathbf{P}_\varrho[n]. \end{aligned}$$

Assume now that the moment sequences $(\mathbf{Q}_\varrho[k])_{k \in \mathbb{N}}$ and $(\mathbf{P}_\varrho[k])_{k \in \mathbb{N}}$ fulfil the operational condition of exponential boundedness. Then the distributions $\mathbf{Q}_\varrho$ and $\mathbf{P}_\varrho$ are determinate. Thus in this case the statistics $\mu * \mathbf{Q}_\varrho$, $\nu * \mathbf{P}_\varrho$ determine uniquely the position and momentum distributions $\mathbf{Q}_\varrho$ and $\mathbf{P}_\varrho$.

The probability measures μ and ν depend, in concrete applications, on the adjustable parameters of the measurement scheme applied, as will be seen in Sect. 19.3. Therefore, without any essential limitations, one may always choose those parameters such that the measures are exponentially bounded. Let $\varrho = |\varphi\rangle\langle\varphi|$ be a pure state, with φ a linear combination of Hermite functions. For such states, the probability measures $\mathbf{Q}_\varrho$ and $\mathbf{P}_\varrho$ are exponentially bounded so that they are uniquely determined by their respective moment sequences $(\langle\varphi| \mathbf{Q}^k \varphi\rangle)_{k \in \mathbb{N}}$ and $(\langle\varphi| \mathbf{P}^k \varphi\rangle)_{k \in \mathbb{N}}$. Furthermore, since the linear span of the Hermite functions is dense in $L^2(\mathbb{R})$, the distributions (densities) $|\varphi|^2$ and $|\hat{\varphi}|^2$ suffice to determine the whole observables $\mathbf{Q}$ and $\mathbf{P}$. Finally, since the convolution of any two exponentially bounded probability measures is also exponentially bounded, one may conclude that under the above assumptions, even the actual statistics $\mu * \mathbf{Q}_\varrho$, $\nu * \mathbf{P}_\varrho$ are uniquely determined by the moment sequences $(\mu * \mathbf{Q}_\varrho[k])_k$ and $(\nu * \mathbf{P}_\varrho[k])_k$, which, in turn, uniquely determine the distributions $\mathbf{Q}_\varrho$ and $\mathbf{P}_\varrho$.

15.5 Appendix

This appendix develops details of the proof of Theorem 15.3. A key tool is the Markov–Kakutani fixed point theorem dealing with a commuting set of continuous affine mappings on a compact convex subset of a topological vector space. As a preparation, we present some auxiliary material of more general interest. We begin with some functional analytic considerations.

Compactness in a Space of Operators

Let Ω be a locally compact Hausdorff space and $C_0(\Omega)$ the Banach space (commutative C^* -algebra) of continuous complex functions on Ω vanishing at infinity. On the space $\mathcal{L}(C_0(\Omega), \mathcal{L}(\mathcal{H}))$ of all bounded linear maps from $C_0(\Omega)$ to $\mathcal{L}(\mathcal{H})$ we consider the topology τ_s of simple (i.e. pointwise) weak* convergence. This means that τ_s is the locally convex topology generated by the seminorms $\Phi \mapsto |\operatorname{tr}[T\Phi(f)]|$ where $f \in C_0(\Omega)$, $T \in \mathcal{T}(\mathcal{H})$. The topology τ_s may also be seen as being induced by the product topology τ_p in the space Π of all maps from $C_0(\Omega)$ to the space $\mathcal{L}(\mathcal{H})$ which is equipped with the weak* topology $\sigma(\mathcal{L}(\mathcal{H}), \mathcal{T}(\mathcal{H}))$. Since the τ_p -limit of a net of linear maps is linear, the subspace L of linear (not necessarily bounded) maps in Π is τ_p -closed. Similarly, the subset L^+ of L , consisting of the positive linear maps $\Phi \in L$ (for which $\Phi(f) \geq 0$ whenever $f \geq 0$) is τ_p -closed (see Theorem 3.14). We finally consider the set $L_1^+ = \{\Phi \in L^+ \mid \|\Phi f\| \leq 1 \text{ whenever } \|f\| \leq 1\}$.

Lemma 15.2 *The set L_1^+ is τ_s -compact.*

Proof Clearly $L_1^+ \subset \mathcal{L}(C_0(\Omega), \mathcal{L}(\mathcal{H}))$. Since the closed unit ball B of $\mathcal{L}(\mathcal{H})$ is weak*-compact in $\mathcal{L}(\mathcal{H})$ by the Banach–Alaoglu theorem, any net in L_1^+ converging with respect to τ_s has its limit also (in the closed set L^+ and, moreover) in L_1^+ , and so L_1^+ is τ_s -closed. Now let D denote the closed unit ball of $C_0(\Omega)$. Since B is weak*-compact, by Tychonoff’s theorem the function space B^D is compact with respect to the product topology. But the mapping $\Phi \mapsto \Phi|_D$ is a homeomorphism from L_1^+ (with the topology τ_s) onto the set of restrictions $R = \{\Phi|_D \mid \Phi \in L_1^+\}$ (with the topology induced by the product topology of B^D). It is thus enough to show that R is closed in B^D . Let (ψ_i) be a net in R converging to some $\psi \in B^D$. Each $\psi_i = \Psi_i|_D$ for a unique $\Psi_i \in L_1^+$. For any nonzero $f \in C_0(\Omega)$ we have $\Psi_i(f) = \|f\| \Psi_i(\|f\|^{-1}f) = \|f\| \psi_i(\|f\|^{-1}f) \rightarrow \|f\| \psi(\|f\|^{-1}f)$. Denote $\Psi(f) = \|f\| \psi(\|f\|^{-1}f) = \lim_i \Psi_i(f)$. Since Ψ is the pointwise limit of the net (Ψ_i) , we know that $\Psi \in L_1^+$. To complete the proof, we still show that $\psi = \Psi|_D$. For any nonzero $f \in D$ we have $\psi(f) = \lim_i \psi_i(f) = \lim_i \Psi_i(f) = \Psi(f)$. $\square$

Recall from Sect. 4.10 the definition of a normalised positive linear map $\Phi : C_o(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$. Such a map Φ belongs to L_1^+ as can be seen for example from Proposition 5.5 by using the integral representation in Theorem 4.4.

Remark 15.5 The subset of L_1^+ consisting of its normalised elements is in general not τ_s -compact. For example, take $\Omega = \mathbb{N}$ and let $\Phi_n(f) = f(n)I$ for all $n \in \mathbb{N}$ and $f \in C_0(\mathbb{N})$. Then the sequence (Φ_n) of normalised positive linear operators tends to the

zero operator with respect to τ_s . In the sequel, however, we consider some interesting compact subclasses of normalised positive operator measures (equivalently, positive linear maps). In their treatment the following lemma along with its consequence Proposition 15.5 will be instrumental. $\triangleleft$

Lemma 15.3 *Let $(\mu_i)_{i \in \mathcal{I}}$ be a net of positive linear functionals on $C_0(\Omega)$ converging in the weak* topology to some (clearly positive linear functional) μ . Suppose that $g \in C_0(\Omega)$ is a strictly positive function and $C > 0$ a constant such that $\mu_i(f) \leq C$ whenever $f \in C_0(\Omega)$ satisfies $0 \leq f \leq g^{-1}$.*

(a) *Suppose that $\sup\{\mu_i(f) \mid f \in C_0(\Omega), 0 \leq f \leq 1\} = \alpha$ for all $i \in \mathcal{I}$. If $\epsilon > 0$, there is some $f \in C_0(\Omega)$, $0 \leq f \leq 1$, such that $\mu(f) \geq \alpha - \epsilon$.*

(b) *If $\|\mu_i\| = 1$ for all $i \in \mathcal{I}$, then $\|\mu\| = 1$.*

Proof (a) Let $\epsilon > 0$. Choose a compact set $K \subset \Omega$ such that $g(x) < \frac{\epsilon}{C}$ whenever $x \in \Omega \setminus K$. If $f \in C_0(\Omega)$ satisfies $0 \leq f \leq 1$ and $\text{supp}(f) \subset \Omega \setminus K$, we have $\frac{C}{\epsilon} f \leq g^{-1}$, and so $\frac{C}{\epsilon} \mu_i(f) \leq C$, i.e. $\mu_i(f) \leq \epsilon$. By regularity and Urysohn's lemma we get, using the same notation for the functional and the measure corresponding to it in the Riesz–Markov–Kakutani representation theorem, that $\mu_i(\Omega \setminus K) \leq \epsilon$, so that $\mu_i(K) \geq \alpha - \epsilon$. Now choose a function $h \in C_0(\Omega)$ satisfying $0 \leq h \leq 1$ and $h(x) = 1$ for all $x \in K$. Then $\mu(h) = \lim_i \mu_i(h) \geq \alpha - \epsilon$.

(b) As it is clear that $\|\mu\| \leq 1$, (a) implies our claim. $\square$

We prepare the operator version of the above result with another lemma.

Lemma 15.4 *Let $\mathcal{T}$ be a set of positive trace class operators such that the intersection of their kernels is $\{0\}$. If $A \in \mathcal{L}(\mathcal{H})$ is such that $A \geq 0$ and $\text{tr}[AT] = 0$ for all $T \in \mathcal{T}$, then $A = 0$.*

Proof Take any unit vector η and extend $\{\eta\}$ to an orthonormal basis K of $\mathcal{H}$. If $T \in \mathcal{T}$, then $0 = \text{tr}[AT] = \text{tr}[A^{\frac{1}{2}}TA^{\frac{1}{2}}] = \sum_{\xi \in K} \left\langle \xi \mid A^{\frac{1}{2}}TA^{\frac{1}{2}}\xi \right\rangle$, which implies that each term vanishes. In particular, $0 = \left\langle \eta \mid A^{\frac{1}{2}}TA^{\frac{1}{2}}\eta \right\rangle = \|T^{\frac{1}{2}}A^{\frac{1}{2}}\eta\|^2$. Thus $T^{\frac{1}{2}}A^{\frac{1}{2}} = 0$, implying that $TA^{\frac{1}{2}} = T^{\frac{1}{2}}T^{\frac{1}{2}}A^{\frac{1}{2}} = 0$, for all $T \in \mathcal{T}$, and so by our assumption on $\mathcal{T}$ we have $A = (A^{\frac{1}{2}})^2 = 0$. $\square$

Proposition 15.5 *Let $(\Phi_i)_{i \in \mathcal{I}}$ be a net of positive linear maps from $C_0(\Omega)$ to $\mathcal{L}(\mathcal{H})$. Suppose that this net converges with respect to τ_s to some (clearly positive) linear map $\Phi : C_0(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$. Let $\mathcal{T}$ be a set of positive trace class operators such that the intersection of their kernels is $\{0\}$. Suppose that for each $T \in \mathcal{T}$ there is a strictly positive function $g_T \in C_0(\Omega)$ with a constant $C_T > 0$ such that $\text{tr}[T\Phi_i(f)] \leq C_T$ for all $i \in \mathcal{I}$ whenever $f \in C_0(\Omega)$ satisfies $0 \leq f \leq g_T^{-1}$. If Φ_i is normalised for all $i \in \mathcal{I}$, then Φ is also normalised.*

Proof Consider a fixed $T \in \mathcal{T}$. Let $\epsilon > 0$. Now let $S = \sup\{\Phi(f) \mid f \in C_0(\Omega), 0 \leq f \leq 1\}$. From Lemma 15.3 (a) and Lemma 6.1 we see that there is a function $f \in$

$C_0(\Omega)$, $0 \leq f \leq 1$, such that $\text{tr}[T\Phi_i(f)] > \text{tr}[T] - \epsilon$ for all $i \in \mathcal{I}$. Thus by Theorem 3.14 (a) we have $\text{tr}[TS] \geq \text{tr}[T\Phi(f)] = \lim_i \text{tr}[T\Phi_i(f)] > \text{tr}[T] - \epsilon$. It follows that $\text{tr}[TS] \geq \text{tr}[T]$. On the other hand, whenever $f \in C_0(\Omega)$ and $0 \leq f \leq 1$, we get $\langle \xi | \Phi(f)\xi \rangle = \lim_i \langle \xi | \Phi_i(f)\xi \rangle \leq \langle \xi | \xi \rangle$ for all $\xi \in \mathcal{H}$, so that $S \leq I$. Since $T \geq 0$, by Theorem 3.14 (a) we thus get $\text{tr}[TS] \leq \text{tr}[T]$, and so $\text{tr}[TS] = \text{tr}[T]$. Since $I - S \geq 0$ and $\text{tr}[(I - S)T] = 0$ for all $T \in \mathcal{T}$, Lemma 15.4 implies that $S = I$. $\square$

A Compact Subset of Approximating Observables

We now specify $\Omega = \mathbb{R}^2$. For any positive operator measure $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ we define the positive linear map $\Phi_{\mathbf{E}} : C_0(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$ via the formula $\Phi_{\mathbf{E}}(f) = \int f d\mathbf{E}$. Thus $\Phi_{\mathbf{E}}(f) = L(f, \mathbf{E})$ in our usual notation. Since any complex measure defined on $\mathcal{B}(\mathbb{R}^2)$ is known to be regular (see Proposition 4.9), Theorem 4.4 shows that in this way we get a bijection $\mathbf{E} \mapsto \Phi_{\mathbf{E}}$ from the set of positive operator measures onto the set of positive linear operators. Via this bijection we transfer the topology τ_s of simple weak* convergence (induced on the set of positive operators $\Phi : C_0(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$) to the set of positive operator measures $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$. We retain the notation τ_s for this transferred topology, too. When convenient, we even identify a positive operator measure $\mathbf{E}$ with its image in this bijection. This usage should not cause confusion.

Let $\mathbf{A} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{B} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ be any two normalised positive operator measures. They will serve as target observables which we aim to approximate with the Cartesian margins $\mathbf{E}^1$ and $\mathbf{E}^2$ of some observables $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$. To specify the degree of approximation we fix the exponents $1 \leq \alpha, \beta < \infty$ and for fixed uncertainty levels $\Delta_{\mathbf{A}}$ and $\Delta_{\mathbf{B}}$ we let $\mathcal{N}$ denote the set of observables $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ for which $\Delta_{\alpha}(\mathbf{A}, \mathbf{E}^1) \leq \Delta_{\mathbf{A}}$ and $\Delta_{\beta}(\mathbf{B}, \mathbf{E}^2) \leq \Delta_{\mathbf{B}}$. Thus in our identification of positive linear maps with positive operator measures, $\mathcal{N}$ will be seen as a subclass of L_1^+ .

We show next that under some relatively mild assumptions on the target observables $\mathbf{A}, \mathbf{B}$ the set $\mathcal{N}$ is a τ_s -closed subset of the τ_s -compact set L_1^+ . To that end, we recall that by the Kantorovich duality (13.11) we have, for any observable $\mathbf{E} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ and any state ϱ ,

$$\Delta_{\alpha}(\mathbf{A}_{\varrho}, \mathbf{E}_{\varrho}^1)^{\alpha} = \sup \left(\int f d\mathbf{E}_{\varrho}^1 - \int g d\mathbf{A}_{\varrho} \right),$$

where the supremum is taken over all nonnegative continuous functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ with compact support satisfying $f(y) - g(x) \leq |x - y|^{\alpha}$. A similar statement holds for $\Delta_{\beta}(\mathbf{B}_{\varrho}, \mathbf{E}_{\varrho}^2)$.

Thus, in particular, for each $\mathbf{E} \in \mathcal{N}$, for any $\varrho \in \mathcal{S}(\mathcal{H})$ and for all functions f and g with the above properties,

$$\int f d\mathbf{E}_{\varrho}^1 - \int g d\mathbf{A}_{\varrho} \leq \Delta_{\mathbf{A}}^{\alpha}. \quad (15.33)$$

Finally, we recall that for any state $\varrho \in \mathcal{S}(\mathcal{H})$, $\Delta_\alpha(\mathbf{A}_\varrho, \delta_0)^\alpha = \int |x|^\alpha d\mathbf{A}_\varrho(x)$, and similarly for $\mathbf{B}$, β . An essential ingredient in the compactness proof of $\mathcal{N}$ is that for a sufficiently rich set of states the preparation uncertainties $\Delta_\alpha(\mathbf{A}_\varrho, \delta_0)$ and $\Delta_\beta(\mathbf{B}_\varrho, \delta_0)$ of the target observables are finite. In particular, this is the case if the probability measures $\mathbf{A}_\varrho$ and $\mathbf{B}_\varrho$ involved are compactly supported or exponentially bounded.

Lemma 15.5 *Let $(\mathbf{E}_i)$ be a net in $\mathcal{N}$ converging with respect to τ_s to some $\tilde{\mathbf{E}} \in L_1^+$. Assume that there is a subset $\mathcal{T}$ of states such that the intersection of their kernels is $\{0\}$ and for each $\varrho \in \mathcal{T}$ the preparation uncertainties $\Delta_\alpha(\mathbf{A}_\varrho, \delta_0)$ and $\Delta_\beta(\mathbf{B}_\varrho, \delta_0)$ are finite.*

- (a) *The positive operator measure $\tilde{\mathbf{E}}$ is normalised.*
- (b) $\tilde{\mathbf{E}} \in \mathcal{N}$.

Proof (a) Let $h(x, y) = |x|^\alpha + |y|^\beta + 1$ and $g = h^{-1}$. Then for any $\mathbf{E} \in \mathcal{N}$ the triangle inequality (13.3) implies $\Delta_\alpha(\mathbf{E}_\varrho^1, \delta_0) \leq \Delta_\alpha(\mathbf{E}_\varrho^1, \mathbf{A}_\varrho) + \Delta_\alpha(\mathbf{A}_\varrho, \delta_0) \leq \Delta_\mathbf{A} + \Delta_\alpha(\mathbf{A}_\varrho, \delta_0)$, and similarly for the second component. Therefore, for any $\mathbf{E} \in \mathcal{N}$ and $\varrho \in \mathcal{T}$ we have

$$\int h(x, y) d\mathbf{E}_\varrho(x, y) \leq (\Delta_\mathbf{A} + \Delta_\alpha(\mathbf{A}_\varrho, \delta_0))^\alpha + (\Delta_\mathbf{B} + \Delta_\beta(\mathbf{B}_\varrho, \delta_0))^\beta + 1 = C_\varrho.$$

This holds, in particular, for the elements of the net $(\mathbf{E}_i)$ in $\mathcal{N}$ converging to $\tilde{\mathbf{E}} \in L_1^+$. Using Proposition 15.5 we thus see that the limit $\tilde{\mathbf{E}}$ is normalised.

(b) Let $f, g \in \mathcal{C}_0(\mathbb{R})$ be nonnegative real functions with compact support and ϱ a state. Choose a sequence of continuous functions $e_n : \mathbb{R} \rightarrow [0, 1]$ with compact support, converging pointwise to 1. By the preceding discussion, in particular (15.33), we have

$$\int f d(\mathbf{E}_i)_\varrho^1 - \int g d\mathbf{A}_\varrho \leq \Delta_\mathbf{A}^\alpha,$$

implying

$$\int f(x)e_n(y) d(\mathbf{E}_i)_\varrho(x, y) \leq \int f d(\mathbf{E}_i)_\varrho^1 \leq \Delta_\mathbf{A}^\alpha + \int g d\mathbf{A}_\varrho.$$

Since $\mathbf{E}_i \rightarrow \tilde{\mathbf{E}}$ in τ_s , we get

$$\int f(x)e_n(y) d\tilde{\mathbf{E}}_\varrho(x, y) \leq \Delta_\mathbf{A}^\alpha + \int g d\mathbf{A}_\varrho,$$

and so

$$\int f d\tilde{\mathbf{E}}_\varrho^1 = \int f(x) d\tilde{\mathbf{E}}_\varrho(x, y) = \lim_{n \rightarrow \infty} \int f(x)e_n(y) d\tilde{\mathbf{E}}_\varrho(x, y) \leq \Delta_\mathbf{A}^\alpha + \int g d\mathbf{A}_\varrho$$

by the dominated convergence theorem. A similar argument works for the second component. As $\tilde{\mathbf{E}}$ is normalised (by virtue of (a)), then $\tilde{\mathbf{E}} \in \mathcal{N}$. $\square$

Proposition 15.6 *Assume that $\mathcal{T}$ is as in Lemma 15.5. Then the set $\mathcal{N}$ is τ_s -compact.*

Proof It follows from the preceding lemma that $\mathcal{N}$ is a closed subset of the set L_1^+ , which is τ_s -compact by Lemma 15.2. $\square$

Remark 15.6 The above proposition extends immediately to approximate measurements of any finite set of observables $A^i, i = 1, \dots, n$, with the assumption that there is a subset $\mathcal{T}$ of states such that the intersection of their kernels is $\{0\}$ and for each $\varrho \in \mathcal{T}$ all the preparation uncertainties $\Delta_{\alpha_i}(A_{\varrho}^i, \delta_0)$ are finite. $\triangleleft$

Applying the Markov–Kakutani Fixed Point Theorem

We now take $A = Q$ and $B = P$. For the set $\mathcal{T}$ we may choose, for instance, the oscillator eigenstates, that is, the one-dimensional projections defined by the Hermite functions in the $L^2(\mathbb{R})$ -representation. (See e.g. the paragraph preceding this appendix.) For the fixed exponents α, β and the uncertainty levels Δ_Q, Δ_P the set $\mathcal{N}$ of approximating joint observables is thus compact.

For any observable $E : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ and $(q, p) \in \mathbb{R}^2$, we use the Weyl operator $W(q, p)$ to define

$$E^{(q,p)}(Z) = W(q, p)^* E(Z + (q, p)) W(q, p) \quad (15.34)$$

whenever $Z \in \mathcal{B}(\mathbb{R}^2)$. Then $E^{(q,p)} \in \mathcal{N}$ whenever $E \in \mathcal{N}$ (exercise). We denote by $\mathcal{E}_{(q,p)} : \mathcal{N} \rightarrow \mathcal{N}$ the mapping $E \mapsto E^{(q,p)}$. Then $E \in \mathcal{N}$ is covariant if and only if it is a fixed point of every $\mathcal{E}_{(q,p)}$.

In the proof of the next result we use the following formulation of the Markov–Kakutani fixed point theorem (see Theorem V.10.6 and the remark following it in [29, p. 456]): If K is a compact convex subset of a topological vector space and $\mathcal{F}$ is a commuting set of continuous affine mappings $F : K \rightarrow K$, then there is a point $p \in K$ such that $Fp = p$ for all $F \in \mathcal{F}$. In our application we consider $K = \mathcal{N}$, which is clearly a convex and, by the preceding proposition, τ_s -compact subset of $\mathcal{L}(C_0(\mathbb{R}^2), \mathcal{L}(L^2(\mathbb{R})))$. For $\mathcal{F}$ we take the family of the maps $\mathcal{E}_{(q,p)}$ defined above.

Proposition 15.7 *There is at least one covariant $E \in \mathcal{N}$.*

Proof A glance at the formula (15.2) yields the equation $\mathcal{E}_{(q+q', p+p')} = \mathcal{E}_{(q,p)} \mathcal{E}_{(q',p')}$ which implies that the maps $\mathcal{E}_{(q,p)}$ form a commutative family of (clearly affine) bijections of $\mathcal{N}$. We show that each $\mathcal{E}_{(q,p)}$ is τ_s -continuous. Suppose $T \in \mathcal{T}(\mathcal{H})$, and $f \in C_0(\mathbb{R}^2)$. Using the notation $f_{(q,p)}(x, y) = f((x, y) - (q, p))$, Proposition 4.14, and the fact that $U^* L(f, E) U = L(f, U^* E(\cdot) U)$ for any unitary operator U , we obtain the equality $\text{tr}[T L(f, \mathcal{E}_{(q,p)}(E))] = \text{tr}[T W(q, p)^* L(f_{(q,p)}, E) W(q, p)] = \text{tr}[W(q, p) T W(q, p)^* L(f_{(q,p)}, E)]$. Here

$$W(q, p) T W(q, p)^* \in \mathcal{T}(\mathcal{H})$$

and $f_{(q,p)} \in C_0(\mathbb{R}^2)$, and using the definition of the topology τ_s we see that $\mathcal{E}_{(q,p)}$ is continuous. Therefore the Markov–Kakutani fixed point theorem implies our claim. $\square$

Now the proof of Theorem 15.3 is obtained by combining this result with the fact that the covariant phase space observables satisfy the inequality (15.29).

15.6 Exercises

1. Show that the infimum (15.4) of the uncertainty product $\text{Var}(\mathbf{Q}_\theta) \cdot \text{Var}(\mathbf{P}_\theta)$ is reached exactly for the Gaussian states (15.8).
2. Show that the equality $AB - BA = cI$, $A, B \in \mathcal{L}(\mathcal{H})$, $c \in \mathbb{C}$, implies that $\dim(\mathcal{H}) = 1$.
3. Assume that A and B are selfadjoint operators such that their uncertainty product $\text{Var}(\mathbf{A}_\theta) \cdot \text{Var}(\mathbf{B}_\theta)$ has a state independent strictly positive lower bound. Show that both A and B are unbounded.
4. Show that the Fourier–Plancherel operator F equals $U_{-\pi/2}$.
5. Show that $Q \cos \theta + P \sin \theta \subset Q_\theta$.
6. Let $\mathcal{H} = L^2(\mathbb{R})$. Show that the map $(t, q, p) \mapsto U(t, q, p)$, with $U(t, q, p)f(x) = e^{i(t+pq/2+qx)}f(x+p)$, is a strongly continuous unitary operator valued function on $\mathbb{R}^3$.
7. For a fixed $(t, q, p) \in \mathbb{R}^3$, show that $s \mapsto U_s(t, q, p) = U(st, sq, sp)$, is a one-parametric unitary group with the property $U_s(t, q, p)(C_c^\infty(\mathbb{R})) \subset C_c^\infty(\mathbb{R})$.
8. Let A be the selfadjoint operator defined by $s \mapsto U_s(t, q, p)$ through the Stone Theorem 5.6. Show that for any $f \in C_c^\infty(\mathbb{R})$, $(Af)(x) = (t + qQ + pP)f(x)$, $x \in \mathbb{R}$, meaning that the operator $qQ + pP$ is essentially selfadjoint on $C_c^\infty(\mathbb{R})$. Show also that for any $f \in C_c^\infty(\mathbb{R})$, $(U_s(t, q, p)f)(x) = e^{i(st+s^2pq/2+sqx)}f(x+sp) = e^{is(t+qP+pP)}f(x)$.
9. Use the above items 6, 7, 8 to prove (15.9).
10. Prove (15.3).
11. Prove (15.11).
12. Prove (15.12).
13. Prove (15.13).
14. Prove (15.14).
15. Prove the equivalence of the inequalities (15.15) and (15.18).
16. Consider a pair of unsharp position and momentum $\mu * \mathbf{Q}, \nu * \mathbf{P}$. Show that (15.6) is still valid if μ, ν have bounded supports.
17. Give an example of a smeared position observable $\mu * \mathbf{Q}$ with finite error bar width $\mathcal{W}_{\varepsilon_1}(\mu * \mathbf{Q}, \mathbf{Q})$ for all $\varepsilon_1 \in (0, 1)$ and infinite Wasserstein deviation $\Delta_2(\mu * \mathbf{Q}, \mathbf{Q})$.
18. Let $\mathbf{G}^K$ be the covariant phase space observable which minimises the error in Theorem 15.3. Show that K is not a Gaussian state for $\alpha = \beta = 1$.
Hint. Let L_N be the linear span of the Hermite functions h_n , $n = 0, \dots, N$, for some fixed $N \geq 2$. Let $\psi_0 \in L_N$ be a function (of unit length) which minimises the number $\langle \psi | S\psi \rangle$, $\psi \in L_N$, and show that the choice $K = |\psi_0\rangle\langle\psi_0|$ gives a smaller value for the product $\Delta_1(\mathbf{Q}, \mathbf{G}_1^K) \cdot \Delta_1(\mathbf{P}, \mathbf{G}_2^K)$ than the choice $K = |h_0\rangle\langle h_0|$.

19. Assume that the variance of a probability measure $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ is zero. Show that μ is a point measure at $\mu[1]$.
20. Show that the convolution of any two exponentially bounded probability measures is exponentially bounded.

References

1. von Neumann, J.: Die Eindeutigkeit der Schrödingerschen Operatoren. *Math. Ann.* **104**(1), 570–578 (1931)
2. Cassinelli, G., Varadarajan, V.S.: On Accardi's notion of complementary observables. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **5**(2), 135–144 (2002)
3. Heisenberg, W.: Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Physik* **33**, 879–893 (1925)
4. Born, M., Jordan, P.: Zur Quantenmechanik. *Z. Physik* **34**(1), 858–888 (1925)
5. Wintner, A.: The unboundedness of quantum-mechanical matrices. *Phys. Rev.* **71**, 738–739 (1947)
6. Wielandt, H.: Über die Unbeschränktheit der Operatoren der Quantenmechanik. *Math. Ann.* **121**, 21 (1949)
7. Putnam, C.R.: Commutation properties of hilbert space operators and related topics. In: *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band, vol. 36*. Springer, New York, Inc., New York (1967)
8. Garrison, J.C., Wong, J.: Canonically conjugate pairs, uncertainty relations, and phase operators. *J. Math. Phys.* **11**, 2242–2249 (1970)
9. Rellich, F.: Der Eindeutigkeitssatz für die Lösungen der quantenmechanischen Vertauschungsrelationen. *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. Math.-Phys.-Chem. Abt.* **107–115**, 1946 (1946)
10. Mackey, G.W.: Imprimitivity for representations of locally compact groups. I. *Proc. Natl. Acad. Sci. U. S. A.* **35**, 537–545 (1949)
11. Mackey, G.W.: Unitary group representations in physics, probability, and number theory. *Advanced Book Classics*. Advanced Book Program, 2nd edn. Addison-Wesley Publishing Company, Redwood City (1989)
12. Busch, P., Lahti, P.J.: To what extent do position and momentum commute? *Phys. Lett. A* **115**(6), 259–264 (1986)
13. Busch, P., Schonbek, T.P., Schroeck Jr., F.E.: Quantum observables: compatibility versus commutativity and maximal information. *J. Math. Phys.* **28**(12), 2866–2872 (1987)
14. Reiter, H., Thirring, W.: Are x and p incompatible observables? *Found. Phys.* **19**(8), 1037–1039 (1989)
15. Ylinen, K.: Commuting functions of the position and momentum observables on locally compact abelian groups. *J. Math. Anal. Appl.* **137**(1), 185–192 (1989)
16. Zak, J.: The kq -representation in the dynamics of electrons in solids. In: Henry Ehrenreich, F.S., Turnbull, D. (eds.) *Solid State Physics*, pp. 1–62. Academic Press, New York (1972)
17. Taylor, M.E.: Noncommutative harmonic analysis. *Mathematical Surveys and Monographs*. American Mathematical Society, Providence (1986)
18. Busch, P., Lahti, P., Werner, R.F.: Measurement uncertainty relations. *J. Math. Phys.* **55**, 042111 (2014)
19. Werner, R.: The uncertainty relation for joint measurement of position and momentum. *Quantum Inf. Comput.* **4**(6–7), 546–562 (2004)
20. Simon, B.: *Trace Ideals and their Applications*. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge (1979)
21. Davies, E.B.: *One-Parameter Semigroups*. London Mathematical Society Monographs, vol. 15. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London (1980)

22. Landau, H.J., Pollak, H.O.: Prolate spheroidal wave functions, Fourier analysis and uncertainty. II. *Bell Syst. Tech. J.* **40**, 65–84 (1961)
23. Lenard, A.: The numerical range of a pair of projections. *J. Funct. Anal.* **10**, 410–423 (1972)
24. Uffink, J.: *The Uncertainty Principle*. Ph.D. thesis, University of Utrecht (1990)
25. Carmeli, C., Heinonen, T., Toigo, A.: On the coexistence of position and momentum observables. *J. Phys. A* **38**(23), 5253–5266 (2005)
26. Busch, P., Pearson, D.: Universal joint-measurement uncertainty relation for error bars. *J. Math. Phys.* **48**(8), 082103 (2007)
27. Freud, G.: *Orthogonal Polynomials*. Pergamon Press, Oxford (1971)
28. Simon, B.: The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.* **137**(1), 82–203 (1998)
29. Dunford, N., Schwartz, J.T.: *Linear Operators. Part I*. Wiley Classics Library. Wiley, New York (1988). Reprint of the 1958 original, A Wiley-Interscience Publication

Chapter 16

Number and Phase

In classical physics the phase of an electromagnetic field is well defined both theoretically and by interference experiments. Diffraction of light, holography, and many other phase dependent phenomena are well understood. As we will see in Sect. 19.3, it is also easy to describe classically eight-port homodyne detection and other direct measurements of the phase difference of signal and local strong laser beams with fixed phases. A problem arises when, for instance, the signal field in homodyne detection is so weak that one must take into account quantum effects [1].

The quantum theory of eight-port homodyne detection shows that when a local laser is strong and its phase is fixed one can measure a phase space observable whose angle margin observable is covariant under the shifts generated by the number operator (when the reference state is diagonal in the number basis). Using this covariance condition together with the choice of the range for the phase variable one can define a whole family of phase shift covariant observables. None of these observables is projection valued, which shows that there is no phase shift covariant selfadjoint phase operator. The quantum phase problem thus highlights the need to extend the notion of observable beyond the traditional identification with selfadjoint operators.

The problem of defining quantum phase was sketched out in Dirac's paper of 1927 [2] and has since been addressed in numerous articles and monographs [3, 4], resulting in a variety of alternative descriptions. For an overview of the literature, one may refer to the reference lists of [4–6].

In this chapter, we present a simple solution to the quantum phase problem in the form of covariant phase observables, and study their properties. The order of presentation is logical rather than historical. Readers interested in the history of the problem may consult, for example, Lynch's review article [5], Pegg's and Barnett's tutorial review [6], or the special issue T48 of *Physica Scripta* [7]. Profound introductions to the treatment of observables as covariant objects and the history of this approach are found in the monographs of Helstrom [8] and Holevo [9].

16.1 Covariant Observables

Let $\mathcal{H}$ be a Hilbert space with an orthonormal basis $\mathbf{b} = \{b_n\}_{n \in Z}$, where $Z \subset \mathbb{Z}$, $Z \neq \emptyset$, and let $N_{\mathbf{b}} = \sum_{n \in Z} n |b_n\rangle\langle b_n|$ be the associated ‘number operator’. We also use the notation $\mathbf{N}_{\mathbf{b}}$ for the number observable, that is, the spectral measure of $N_{\mathbf{b}}$. Note that $[0, 2\pi) \ni \theta \mapsto e^{i\theta N_{\mathbf{b}}}$ is a strongly continuous unitary representation of the group $\mathbb{T} = \{e^{i\theta} \mid \theta \in [0, 2\pi)\}$. In this chapter, we are especially interested in a physically relevant situation where $Z = \mathbb{N}_0$ and vectors b_n are the number states $|n\rangle$ associated with a single-mode optical field (see Sect. 19.3).

Define $e_n(\theta) = e^{-in\theta}$, $n \in \mathbb{Z}$, $\theta \in \mathbb{R}$, so that $\mathbf{e} = \{e_n\}_{n \in \mathbb{Z}}$ is an orthonormal basis of $L^2(\mu)$ where $d\mu(\theta) = (2\pi)^{-1}d\theta$ is the normalised Lebesgue measure on $[0, 2\pi)$. Let $\mathbf{M}$ be the canonical spectral measure on $L^2(\mu)$, i.e. $(\mathbf{M}(X)\psi)(\theta) = \chi_X(\theta)\psi(\theta)$. Physically, $\mathbf{M}$ is a sharp localisation observable of a particle moving on the circle $\mathbb{T}$. In this context, $\hbar N_{\mathbf{e}} = -i\hbar d/d\theta$ is the momentum operator.

Let $\Phi : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ be an observable. If

$$e^{i\theta N_{\mathbf{b}}} \Phi(X) e^{-i\theta N_{\mathbf{b}}} = \Phi(X \dot{+} \theta)$$

for all $\theta \in [0, 2\pi)$ and $X \in \mathcal{B}([0, 2\pi))$ we say that Φ is $\mathbb{T}$ -covariant with respect to $\mathbf{b}$, that is, with respect to the shifts generated by the number operator $N_{\mathbf{b}}$. Here $\dot{+}$ means addition modulo 2π . For example, since

$$\mathbf{M}(X) = \sum_{n, m \in \mathbb{Z}} \int_X e^{i(n-m)\theta} d\mu(\theta) |e_n\rangle\langle e_m|,$$

(weakly) one sees immediately that $\mathbf{M}$ is $\mathbb{T}$ -covariant with respect to $\mathbf{e}$. Physically this means that $\mathbf{M}$ is covariant with respect to the position shifts generated by the momentum operator $\hbar N_{\mathbf{e}}$. In the case of a single mode field, i.e., $b_n = |n\rangle$, $n \geq 0$, the corresponding $\mathbb{T}$ -covariant observable is called a *phase observable* since it is covariant with respect to the phase shifts generated by the number operator $N = N_{\mathbf{b}}$ (Sect. 19.3). We let $\mathbf{N}$ denote the spectral measure of N and call it the number observable.

Theorem 16.1 *An observable $\Phi : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ is $\mathbb{T}$ -covariant with respect to $\mathbf{b}$ if and only if there exists a unit vector sequence $\{\eta_n\}_{n \in \mathbb{Z}}$ in a Hilbert space $\mathcal{K}$ such that*

- (i) *vectors η_n , $n \in \mathbb{Z}$, span $\mathcal{K}$,*
- (ii) *$\Phi(X) = W^*[\mathbf{M}(X) \otimes I]W$ for all $X \in \mathcal{B}([0, 2\pi))$,*

where $W = \sum_{n \in \mathbb{Z}} |e_n \otimes \eta_n\rangle\langle b_n|$ is an isometry from $\mathcal{H}$ to $L^2(\mu) \otimes \mathcal{K}$. In this case, Φ is projection valued if and only if $Z = \mathbb{Z}$ and $\dim \mathcal{K} = 1$.

Proof If $\Phi(X) = W^*[\mathbf{M}(X) \otimes I]W$ as above, it is easy to verify that Φ is $\mathbb{T}$ -covariant with respect to $\mathbf{b}$ since $\mathbf{M}$ is $\mathbb{T}$ -covariant with respect to $\mathbf{e}$ and $(e^{i\theta N_{\mathbf{e}}} \otimes I)W = We^{i\theta N_{\mathbf{b}}}$.

Conversely, let Φ be $\mathbb{T}$ -covariant with respect to $\mathbf{b}$. Let $\varrho = \sum_{n \in \mathbb{Z}} c_n |b_n\rangle\langle b_n|$ be a state where $c_n > 0$ for all $n \in \mathbb{Z}$ and $\sum_{n \in \mathbb{Z}} c_n = 1$. Now Φ is absolutely continuous with respect to the probability measure $\Phi_\varrho(X) = \text{tr}[\varrho\Phi(X)]$. Since $\Phi_\varrho(X) = \Phi_\varrho(X + \theta)$, i.e., Φ_ϱ is shift-invariant, one must have $\Phi_\varrho = \mu$ due to the uniqueness of the (normalised) Haar measure of $\mathbb{T}$. This implies that Φ is μ -continuous and, hence, for all $n, m \in \mathbb{Z}$,

$$\mu_{nm}(X) = \langle b_n | \Phi(X) b_m \rangle = \int_X f_{nm}(\theta) d\mu(\theta)$$

where $f_{nm} \in L^1(\mu)$. From the covariance assumption, one sees that

$$\mu_{nm}(X + \theta) = e^{i(n-m)\theta} \mu_{nm}(X)$$

and, by denoting $k = n - m$, we get

$$\int_0^{2\pi} e^{-il\alpha} d\mu_{nm}(\alpha) = \int_0^{2\pi} e^{-il(\beta+\theta)} d\mu_{nm}(\beta + \theta) = e^{i(k-l)\theta} \int_0^{2\pi} e^{-il\beta} d\mu_{nm}(\beta).$$

Hence,

$$\int_0^{2\pi} e^{-il\alpha} f_{nm}(\alpha) d\mu(\alpha) = \begin{cases} c_{nm} \in \mathbb{C}, & l = k, \\ 0, & l \neq k. \end{cases}$$

Since complex measures on $\mathbb{T}$ are fully determined by their Fourier coefficients, one must have

$$\mu_{nm}(X) = c_{nm} \int_X e^{i(n-m)\theta} d\mu(\theta).$$

Especially, $c_{nn} = 1$ since $\mu_{nn} = \mu$ by the uniqueness of the Haar measure.

The matrix $(c_{nm})_{n,m \in \mathbb{Z}}$ is positive semidefinite, since otherwise there exist complex numbers $d_n, n \in \mathbb{Z}$, such that $d_n \neq 0$ only for finitely many n 's and

$$\sum_{n,m \in \mathbb{Z}} \overline{d_n} c_{nm} d_m < 0;$$

then, by continuity of $\theta \mapsto g(\theta) = \sum_{n,m \in \mathbb{Z}} \overline{d_n} e^{-in\theta} c_{nm} d_m e^{-im\theta}$,

$$\langle \varphi | \Phi([0, \epsilon)) \varphi \rangle = \int_0^\epsilon g(\theta) d\mu(\theta) < 0$$

for a small $\epsilon > 0$, where $\varphi = \sum_{m \in \mathbb{Z}} d_m b_m$, leading to a contradiction with the positivity of Φ . Let $\mathcal{K}$ and $(\eta_n)_{n \in \mathbb{Z}} \subset \mathcal{K}$ be a (separable) Hilbert space and a vector sequence associated with a minimal Kolmogorov decomposition of $(c_{nm})_{n,m \in \mathbb{Z}}$, that is, $c_{nm} = \langle \eta_n | \eta_m \rangle$ and η_n -vectors span $\mathcal{K}$. Now $\|\eta_n\|^2 = c_{nn} = 1$.

By defining an isometry $W = \sum_{n \in Z} |e_n \otimes \eta_n\rangle \langle b_n|$ we get, for all $X \in \mathcal{B}([0, 2\pi))$, that $\Phi(X) = W^*[\mathbf{M}(X) \otimes I]W$.

Since characteristic functions χ_X span $L^2(\mu)$, one finds a minimal Kolmogorov decomposition of a $\mathbb{T}$ -covariant observable Φ :

$$\Phi(X) = \sum_{n,m \in Z} \int_X \langle \psi_n(\theta) | \psi_m(\theta) \rangle d\mu(\theta) |b_n\rangle \langle b_m|$$

where $\psi_n = \eta_n e_n \in L^2_{\mathcal{K}}(\mu) \cong L^2(\mu) \otimes \mathcal{K}$. From Remark 8.1 and Corollary 8.5 we see that Φ is projection valued if and only if $\{\psi_n\}_{n \in Z}$ is an orthonormal basis of $L^2_{\mathcal{K}}(\mu)$ if and only if $Z = \mathbb{Z}$ and $\mathcal{K}$ is one-dimensional. $\square$

From the above theorem, one sees that any $\mathbb{T}$ -covariant observable Φ can be written in the form

$$\Phi(X) = \sum_{n,m \in Z} c_{nm} \int_X e^{i(n-m)\theta} d\mu(\theta) |b_n\rangle \langle b_m|$$

(weakly) where the numbers $c_{nm} = \langle \eta_n | \eta_m \rangle$ form a unique *structure matrix* of Φ . Obviously, the η_n -vectors, which give a fixed $(c_{nm})_{n,m \in Z}$, are not unique, but a certain uniqueness can be reached as follows: let $\mathcal{K}$ and $\mathcal{K}'$ be Hilbert spaces, $(\eta_n)_{n \in Z} \subset \mathcal{K}$ and $(\eta'_n)_{n \in Z} \subset \mathcal{K}'$ sequences which span $\mathcal{K}$ and $\mathcal{K}'$, respectively, and $c_{nm} = \langle \eta_n | \eta_m \rangle = \langle \eta'_n | \eta'_m \rangle$. Then there exists a unitary operator $U : \mathcal{K} \rightarrow \mathcal{K}'$ such that $\eta'_n = U\eta_n$ (see Exercise 1). This shows that the rank of Φ is

$$\text{rank } \Phi = \dim \mathcal{K} \leq \dim \mathcal{H} = \text{card}(Z).$$

Hence, one can always consider $\mathcal{K}$ as a closed subspace of $\mathcal{H}$. Moreover, $(L^2(\mu) \otimes \mathcal{K}, X \mapsto \mathbf{M}(X) \otimes I, W)$ is a minimal diagonal Naimark dilation for Φ . Note that $\text{rank } \mathbf{M} = 1$. In contrast, if $\text{rank } \Phi = 1$ then $\mathcal{K} \cong \mathbb{C}$ and $\eta_n = e^{i\alpha_n}$, $\alpha_n \in [0, 2\pi)$. Thus, any rank-1 Φ is of the form

$$\Phi(X) = V \Phi_{\text{can}}(X) V^*$$

where the unitary operator $V = \sum_{n \in Z} e^{-i\alpha_n} |b_n\rangle \langle b_n|$ commutes with the representation $\theta \mapsto e^{i\theta N_b}$ and

$$\Phi_{\text{can}}(X) = \sum_{n,m \in Z} \int_X e^{i(n-m)\theta} d\mu(\theta) |b_n\rangle \langle b_m|$$

is the *canonical $\mathbb{T}$ -covariant observable (with respect to $\mathbf{b}$)*. Note that sharp $\mathbb{T}$ -covariant observables are always of rank 1 and a rank-1 $\mathbb{T}$ -covariant observable is sharp if and only if $Z = \mathbb{Z}$.

Remark 16.1 Let Φ and Φ' be $\mathbb{T}$ -covariant observables with respect to $\mathbf{b}$. If $\Phi(X) = V\Phi'(X)V^*$ for all X , where the unitary operator

$$V = \sum_{n \in \mathbb{Z}} e^{-i\alpha_n} |b_n\rangle\langle b_n|$$

commutes with the representation $\theta \mapsto e^{i\theta N_b}$, we say that Φ and Φ' are *unitarily equivalent*. (Now V only scales the basis $\mathbf{b}$ and $N_{V\mathbf{b}} = N_b$.) This terminology is slightly ambiguous, since there may exist a unitary operator V_1 which is not diagonal and $\Phi(X) = V_1\Phi'(X)V_1^*$ but Φ and Φ' are not unitarily equivalent in the above sense. Indeed, let $\mathbf{b} = \{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ and $z \in \mathbb{C}$, $0 < |z| < 1$, and define $\mathbb{T}$ -covariant observables Φ' and Φ by setting

$$\begin{aligned}\Phi'(X) &= \mu(X) I + z \int_X e^{-i\theta} d\mu(\theta) |0\rangle\langle 1| + \bar{z} \int_X e^{i\theta} d\mu(\theta) |1\rangle\langle 0|, \\ \Phi(X) &= \mu(X) I + z \int_X e^{-i\theta} d\mu(\theta) |2\rangle\langle 3| + \bar{z} \int_X e^{i\theta} d\mu(\theta) |3\rangle\langle 2|\end{aligned}$$

for all $X \in \mathcal{B}([0, 2\pi))$, and

$$V_1 = |2\rangle\langle 0| + |0\rangle\langle 2| + |1\rangle\langle 3| + |3\rangle\langle 1|.$$

Then $\Phi(X) = V_1\Phi'(X)V_1^*$. It is easy to see that Φ and Φ' cannot be unitarily equivalent in the above sense.

Remark 16.2 Since the structure matrix $(c_{nm})_{n,m \in \mathbb{Z}}$ of any $\mathbb{T}$ -covariant observable is unique and

$$c_{nm} = i(n - m)\langle b_n | \Phi[1] b_m \rangle, \quad n \neq m,$$

where

$$\Phi[1] = \int_0^{2\pi} \theta d\Phi(\theta)$$

is the selfadjoint first moment operator of Φ , the structure of Φ is fully determined already by the first moment operator $\Phi[1]$ —the higher moments are not needed to determine Φ . This leads us to an interesting conclusion: *for a selfadjoint operator A , there may exist two (unequal) observables E_a and E_b such that $E_a[1] = A = E_b[1]$ and E_a and E_b are uniquely determined by A (by requiring some additional properties for E_a, E_b).* In the case of $A = \Phi[1]$ such observables are Φ (which is covariant) and the spectral measure of $\Phi[1]$ (which is projection valued). Hence, it is reasonable to call $\Phi[1]$ a *phase operator* although the phase observable associated with it is the semispectral measure Φ and not the spectral measure of $\Phi[1]$.

16.2 Canonical Phase

In this section, we consider the case of a single-mode electromagnetic field and the associated canonical phase observable, and study its properties.

Let $\mathcal{H}$ be a Hilbert space spanned by the (photon) number states $|n\rangle$, $n \geq 0$. Retaining the notations of the previous section, we now have $Z = \mathbb{N}_0$, $b_n = |n\rangle$, and $N = N_b = \sum_{n \in \mathbb{N}} n |n\rangle \langle n|$. In this context, we say that the canonical $\mathbb{T}$ -covariant observable Φ_{can} is the *canonical phase observable*. We show that Φ_{can} is (essentially) the only phase observable which generates number shifts.

Number Shifts

Since the number operator N generates the phase shifts and phase observables are phase shift covariant, it is reasonable to ask whether there exists a phase observable which generates the number shifts $n \mapsto n + k$. It must be stressed that number shifts $k \in \mathbb{N}_0$ form only a semigroup instead of a group, so that the situation differs slightly from phase shifts which form a group $\mathbb{T}$.

Let Φ be a phase observable with $(c_{nm})_{n,m \in \mathbb{N}_0}$ and define the cyclic moment operators $V_\Phi^{(k)} = \int_0^{2\pi} e^{ik\theta} d\Phi(\theta)$ for all $k \in \mathbb{Z}$. Now, for all $k \in \mathbb{N}_0$,

$$V_\Phi^{(k)} = \left(V_\Phi^{(-k)}\right)^* = \sum_{n \in \mathbb{N}_0} c_{n,n+k} |n\rangle \langle n+k|$$

so that Φ is determined by knowing the cyclic moments for $k \in \mathbb{N}$. Now if Φ is of rank 1, that is, $c_{nm} = e^{i(\alpha_n - \alpha_m)}$, it follows that $k \mapsto V_\Phi^{(k)}$ is a representation of the number shift semigroup $\mathbb{N}_0$. Indeed, $V_\Phi^{(0)} = I$ and

$$V_\Phi^{(k+l)} = V_\Phi^{(k)} V_\Phi^{(l)}, \quad k, l \in \mathbb{N}_0.$$

Let $\mathbf{N}$ be the spectral measure of N so that for any $Y \subset \mathbb{N}_0$

$$\mathbf{N}(Y) = \sum_{n \in Y} |n\rangle \langle n|.$$

We determine the phase observables Φ that generate the number shifts, that is,

$$V_\Phi^{(k)} \mathbf{N}(Y + k) \left(V_\Phi^{(k)}\right)^* = \mathbf{N}(Y)$$

for all $Y \subset \mathbb{N}_0$ and $k \in \mathbb{N}_0$. By direct calculation, one see that this happens exactly when

$$|c_{nm}| = 1.$$

But, by Cauchy–Schwarz inequality, $1 = |c_{nm}| = |\langle \eta_n | \eta_m \rangle| \leq \|\eta_n\| \|\eta_m\| = 1$, so that $\eta_n = c \eta_m$, $c \in \mathbb{T}$, and $\text{rank } \Phi = \dim \text{lin}\{\eta_n \mid n \in \mathbb{N}_0\} = 1$. Hence, we have the following result.

Theorem 16.2 *A phase observable Φ generates number shifts if and only if Φ is unitarily equivalent to Φ_{can} .*

The above theorem shows that N and Φ_{can} form a *canonical pair* in the sense that

N generates phase shifts and transforms covariantly under the number shifts;
 Φ_{can} generates numbers shifts and transforms covariantly under the phase shifts.

Similarly as in the case of position and momentum, we find phase observables which are *invariant* with respect to the number shifts $k \mapsto V^{(k)} = \sum_{n \in \mathbb{N}_0} |n\rangle \langle n+k|$.

Theorem 16.3 *Let Φ be a phase observable and (c_{nm}) its structure matrix. The following items are equivalent:*

(i) *For all $k \in \mathbb{N}_0$ and $X \in \mathcal{B}([0, 2\pi))$,*

$$V^{(k)} \Phi(X) (V^{(k)})^* = \Phi(X).$$

(ii) *For all $k, n, m \in \mathbb{N}_0$,*

$$c_{n+k, m+k} = c_{nm}.$$

(iii) *There exists a probability measure $\nu : \mathcal{B}([0, 2\pi)) \rightarrow [0, 1]$ such that*

$$\Phi(X) = \int_0^{2\pi} \nu(X \dot{+} (-\theta)) d\Phi_{\text{can}}(\theta), \quad X \in \mathcal{B}([0, 2\pi)),$$

that is, Φ is a smearing of Φ_{can} .

Proof It is an easy exercise to show that (i) is equivalent to (ii) and (iii) implies (i). If (ii) holds then $c_{n, n+k} = c_{0, k} = c_k$ and $c_{n+k, n} = \overline{c_k} = c_{-k}$ for all $n, k \in \mathbb{N}_0$. For any $s \geq 1$, define a continuous 2π -periodic function

$$\theta \mapsto \sigma_s(\theta) = \frac{1}{s+1} \sum_{n, m=0}^s \underbrace{c_{nm}}_{=c_{m-n}} e^{i(n-m)\theta} \geq 0$$

for which

$$\int_0^{2\pi} e^{ik\theta} \sigma_s(\theta) d\mu(\theta) = c_k \frac{\max\{0, s+1-|k|\}}{s+1} \rightarrow c_k \quad \text{as } s \rightarrow \infty,$$

The measures ν_s lie in the unit ball of the Banach space of regular Borel measures $\mathcal{B}(\mathbb{T}) \rightarrow \mathbb{C}$ which is the dual space of the separable Banach space of continuous functions $\mathbb{T} \rightarrow \mathbb{C}$ by the Riesz representation theorem. The Banach–Alaoglu theorem says that the above *metrisable* [10, p.426] unit ball is compact with respect to the weak*-topology and hence the sequence $s \mapsto \nu_s$ has a convergent subsequence $p \mapsto \nu_{s_p}$ (note $s_p < s_{p+1}$):

$$\lim_{p \rightarrow \infty} \int_0^{2\pi} f(\theta) d\nu_{s_p}(\theta) = \int_0^{2\pi} f(\theta) d\nu(\theta)$$

for all continuous 2π -periodic functions $f : \mathbb{R} \rightarrow \mathbb{C}$ where $\nu : \mathcal{B}([0, 2\pi]) \rightarrow \mathbb{C}$ is a regular Borel measure with Fourier coefficients c_k (put $f(\theta) = e^{-ik\theta}$ above). Obviously, ν must be a probability measure (consider nonnegative functions f above). Now

$$\int_0^{2\pi} \nu(X \dot{+} (-\theta)) d\langle n | \Phi_{\text{can}}(\theta) | m \rangle = c_{m-n} \int_X e^{i(n-m)\theta} d\mu(\theta) = \langle n | \Phi(X) | m \rangle$$

and (iii) follows. $\square$

From the above proof one sees that if a probability measure ν on the circle is given, one can define a number shift invariant phase observable with the structure matrix elements $c_{nm}^\nu = \int_0^{2\pi} e^{i(m-n)\theta} d\nu(\theta)$. It is a smearing of Φ_{can} ; see item c) of the preceding theorem. Similarly, one can define a smearing of any phase Φ with (c_{nm}) —it is a phase observable with the matrix elements $c_{nm} c_{nm}^\nu$ (they form the Schur product matrix of the positive semidefinite structure matrices (c_{nm}) , (c_{nm}^ν) which is positive semidefinite by the Schur product theorem).

Extremality

Let $\mathcal{O}$ be the convex set of all observables $\mathcal{B}([0, 2\pi]) \rightarrow \mathcal{L}(\mathcal{H})$, and let $\mathcal{O}_{\text{ph}}$ be the convex set of phase observables. Let Φ be a phase observable determined by vectors $\eta_n \in \mathcal{K}$ which span $\mathcal{K}$. Note that, if Φ is an extreme element of $\mathcal{O}$ it is automatically an extreme point of $\mathcal{O}_{\text{ph}} \subset \mathcal{O}$ (this follows also from the next theorem by putting $D(\theta) = D$).

Theorem 16.4 (a) Φ is an extreme point of $\mathcal{O}$ if and only if, for any weakly measurable map $[0, 2\pi) \ni \theta \mapsto D(\theta) \in \mathcal{L}(\mathcal{K})$, the condition

$$\int_0^{2\pi} e^{i(n-m)\theta} \langle \eta_n | D(\theta) \eta_m \rangle d\mu(\theta) = 0, \quad n, m \in \mathbb{N}_0,$$

implies $D(\theta) = 0$ for μ -almost all $\theta \in [0, 2\pi)$.

(b) Φ is an extreme point of $\mathcal{O}$ if Φ is unitarily equivalent to Φ_{can} .

(c) Φ is an extreme point of $\mathcal{O}_{\text{ph}}$ if and only if, for any $D \in \mathcal{L}(\mathcal{K})$,

$$\langle \eta_n | D \eta_n \rangle = 0, \quad n \in \mathbb{N}_0,$$

implies $D = 0$.

(d) There exist extreme points of $\mathcal{O}_{\text{ph}}$ of any rank.

Proof (a) Let Φ be a phase observable with the vectors $\eta_n \in \mathcal{K}$, and let

$$\Phi(X) = \sum_{n, m \in \mathbb{N}_0} \int_X \langle \psi_n(\theta) | \psi_m(\theta) \rangle d\mu(\theta) |n\rangle \langle m|,$$

where $\psi_n = \eta_n e_n \in L^2_{\mathcal{K}}(\mu)$, be its minimal Kolmogorov decomposition. By Theorem 9.5, Φ is extremal if and only if, for any measurable field of operators $[0, 2\pi) \ni \theta \mapsto D(\theta) \in \mathcal{L}(\mathcal{K})$, the condition

$$\int_0^{2\pi} \langle \psi_n(\theta) | D(\theta) \psi_m(\theta) \rangle d\mu(\theta) = \int_0^{2\pi} e^{i(n-m)\theta} \langle \eta_n | D(\theta) \eta_m \rangle d\mu(\theta) = 0, \quad n, m \in \mathbb{N}_0,$$

implies $D(\theta) = 0$ for μ -almost all $\theta \in [0, 2\pi)$.

(b) In the case of Φ_{can} one can choose $\mathcal{K} = \mathbb{C}$, $\eta_n = 1$ and hence $D(\theta) = \lambda(\theta) \in \mathbb{C}$. Now the extremality condition of (a) reads

$$\int_0^{2\pi} e^{i(n-m)\theta} \lambda(\theta) d\mu(\theta) = 0, \quad n, m \in \mathbb{N}_0,$$

so that $\lambda(\theta) = 0$ since the complex measure $X \mapsto \int_X \lambda(\theta) d\mu(\theta)$ is uniquely determined by its Fourier coefficients. If $\Phi(X) = V \Phi_{\text{can}}(X) V^*$ for some unitary V , then, for all $a \in (0, 1)$, $E_1, E_2 \in \mathcal{O}$,

$$\Phi = aE_1 + (1-a)E_2$$

equals

$$\Phi_{\text{can}} = aV^*E_1V + (1-a)V^*E_2V$$

so that $V^*E_1V = V^*E_2V$ and $E_1 = E_2$ showing that Φ is extremal.

(c) Assume that there exists a nonzero $D \in \mathcal{L}(\mathcal{K})$ such that $\langle \eta_n | D \eta_m \rangle = 0$ for all $n \in \mathbb{N}_0$. Without restricting generality, we may assume that $D^* = D$ and $\|D\| \leq 1$. Hence, $D^\pm = I \pm D \geq 0$ and we define observables $\Phi^\pm : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{K})$ by

$$\Phi^\pm(X) = \sum_{n, m \in \mathbb{N}_0} \langle \eta_n | D^\pm \eta_m \rangle \int_X e^{i(n-m)\theta} d\mu(\theta) |n\rangle \langle m|.$$

Now $\Phi(X) = \frac{1}{2}[\Phi^+(X) + \Phi^-(X)]$ and $\Phi^+ \neq \Phi^-$ since $\langle \eta_n | D^+ \eta_m \rangle \neq \langle \eta_n | D^- \eta_m \rangle$ for some n, m (note that $D^+ \neq D^-$ and vectors η_n span $\mathcal{K}$).

If Φ is not extremal, that is, there exist distinct phase observables $\Phi^\pm$ (with the structure matrices $(c_{nm}^\pm)_{n, m \in \mathbb{N}_0}$) such that $\Phi(X) = \frac{1}{2}[\Phi^+(X) + \Phi^-(X)]$, one can define positive semidefinite sesquilinear forms

$$F^\pm(W\varphi, W\psi) = \sum_{n, m \in \mathbb{N}_0} c_{nm}^\pm \langle \varphi | n \rangle \langle m | \psi \rangle$$

where $\varphi, \psi \in \text{lin}\{|n\rangle \mid n \in \mathbb{N}_0\}$ and $W = \sum_{n \in \mathbb{N}_0} |\eta_n\rangle \langle n|$. As $\langle \psi | \Phi^\pm([0, \epsilon)) \psi \rangle \leq 2\langle \psi | \Phi([0, \epsilon)) \psi \rangle$ for all $\epsilon \in [0, 2\pi]$ and $\psi \in \mathcal{H}$ it follows that $F^\pm(W\psi, W\psi) \leq 2\langle W\psi | W\psi \rangle$ and

$$|F^\pm(W\varphi, W\psi)| \leq 2\|W\varphi\| \|W\psi\|, \quad \varphi, \psi \in \text{lin}\{|n\rangle \mid n \in \mathbb{N}_0\},$$

by the Cauchy–Schwarz inequality (which holds for any sesquilinear form). Hence, the forms $F^\pm$ are well-defined and bounded defining operators $D^\pm \in \mathcal{L}(\mathcal{K})$ such that $F^\pm(W\varphi, W\psi) = \langle W\varphi | D^\pm W\psi \rangle$. From $\Phi^+ \neq \Phi^-$ we have $D = D^+ - D^- \neq 0$ and $\langle \eta_n | D \eta_n \rangle = c_{nn}^+ - c_{nn}^- = 1 - 1 = 0$.

(d) Let $r \in \{1, 2, \dots\} \cup \{\infty\}$ be arbitrary, and set $\mathcal{K} = \mathbb{C}^r$, if $r < \infty$, and $\mathcal{K} = \mathcal{H}$ when $r = \infty$. Let $\{\xi_n \in \mathcal{K} \setminus \{0\} \mid n \in \mathbb{N}_0\}$ be a countable dense set in $\mathcal{K}$ (which exists since $\mathcal{K}$ is separable). Define $\eta_n = \xi_n / \|\xi_n\|$. Obviously, the vectors η_n span $\mathcal{K}$. By density, for all $D \in \mathcal{L}(\mathcal{K})$, the condition $\langle \eta_n | D \eta_n \rangle = 0$, $n \in \mathbb{N}_0$, implies $D = 0$, so that the phase observable Φ determined by the η_n -vectors above is a rank- r extremal. $\square$

Operations

In this subsection, we show that a measurement of the canonical phase

$$\Phi_{\text{can}}(X) = \sum_{s,t \in \mathbb{N}_0} \int_X e^{i(s-t)\theta} d\mu(\theta) |s\rangle \langle t|$$

can be used to get measurement outcome statistics of any Φ_{can} -continuous observable.

Let $E : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H}')$ be any Φ_{can} -continuous observable, that is, $\Phi_{\text{can}}(X) = 0$ implies $E(X) = 0$ for all X . Hence, E is of the form

$$E(X) = \sum_{n,m=1}^{\dim \mathcal{H}'} \int_X \langle \psi'_n(\theta) | \psi'_m(\theta) \rangle d\mu(\theta) |\varphi_n\rangle \langle \varphi_m|$$

where $\{\varphi_n\}_{n=1}^{\dim \mathcal{H}'}$ is an orthonormal basis of a (separable) Hilbert space $\mathcal{H}'$ and vectors $\psi'_n \in L^2_{\mathcal{H}'}(\mu) \cong L^2(\mu) \otimes \mathcal{H}'$ form an orthonormal set (see Sect. 8.6). Now vectors $e_n \otimes \varphi_m$, $n \in \mathbb{Z}$, $m \in \{1, \dots, \dim \mathcal{H}'\}$ constitute an orthonormal basis of $L^2(\mu) \otimes \mathcal{H}'$ and we can write

$$\psi'_n = \sum_{s \in \mathbb{Z}} e_s \otimes v_n^s$$

where $v_n^s \in \mathcal{H}'$ and $\sum_{s \in \mathbb{Z}} \langle v_n^s | v_m^s \rangle = \delta_{nm}$.

Let $P_M = \sum_{n=-M}^M |e_n\rangle \langle e_n| \otimes I$ be a projection on $L^2(\mu) \otimes \mathcal{H}'$. Now $\{P_M\}_{M \geq 1}$ is an increasing sequence and $P_M \rightarrow I \otimes I$ strongly when $M \rightarrow \infty$. Hence, for all $\varphi, \psi \in L^2(\mu) \otimes \mathcal{H}'$ and $X \in \mathcal{B}([0, 2\pi))$,

$$\langle \varphi | P_M (\mathbf{M}(X) \otimes I) P_M \psi \rangle \rightarrow \langle \varphi | (\mathbf{M}(X) \otimes I) \psi \rangle$$

when $M \rightarrow \infty$.

Let $\mathcal{X}$ be a finite subset of $\mathcal{B}([0, 2\pi))$, and let $N < \dim \mathcal{H}' + 1$ and $\epsilon > 0$. One can pick an $M > 0$ such that

$$\left| \int_X \langle \psi'_n(\theta) | \psi'_m(\theta) \rangle d\mu(\theta) - \int_X \langle \psi_n^M(\theta) | \psi_m^M(\theta) \rangle d\mu(\theta) \right| < \epsilon$$

for all $X \in \mathcal{X}$ and $n, m \leq N$ where $\psi_n^M = P_M \psi'_n = \sum_{s=-M}^M e_s \otimes v_n^s$. Define a positive operator valued measure

$$\mathbf{E}_M(X) = \sum_{n,m=1}^{\dim \mathcal{H}'} \underbrace{\int_X \langle \psi_n^M(\theta) | \psi_m^M(\theta) \rangle d\mu(\theta)}_{\langle \psi'_n | P_M(M(X) \otimes I) P_M \psi'_m \rangle} |\varphi_n\rangle \langle \varphi_m| \leq I$$

which is not necessarily normalised. However, one can consider $\mathbf{E}_M$ as an approximation of $\mathbf{E}$ when M is large.

It is an easy exercise to show that

$$\Theta_M(\Phi_{\text{can}}(X)) = \mathbf{E}_M(X), \quad X \in \mathcal{B}([0, 2\pi))$$

where $\Theta_M : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}')$ is a completely positive linear map defined as

$$\Theta_M(B) = Y_M^*(B \otimes I)Y_M, \quad B \in \mathcal{L}(\mathcal{H}),$$

where

$$Y_M = \sum_{n=1}^{\dim \mathcal{H}'} \sum_{s=0}^{2M} |s\rangle \otimes v_n^{s-M} \langle \varphi_n|$$

for which $Y_M^* Y_M \leq I$. Moreover, one gets a Kraus decomposition

$$\Theta_M(B) = \sum_{k=1}^{\dim \mathcal{H}'} A_k^* B A_k, \quad B \in \mathcal{L}(\mathcal{H}),$$

where $A_k = \sum_{s=0}^{2M} \sum_{n=1}^{\dim \mathcal{H}'} \langle \varphi_k | v_n^{s-M} \rangle |s\rangle \langle \varphi_n|$ and $\sum_k A_k^* A_k \leq I$. Hence, Θ_M is the dual of a completely positive operation

$$\Theta_{M*} : \mathcal{T}(\mathcal{H}') \rightarrow \mathcal{T}(\mathcal{H}), \quad \varrho \mapsto \sum_{k=1}^{\dim \mathcal{H}'} A_k \varrho A_k^*.$$

In summary, to obtain the measurement outcome statistics of the observable $\mathbf{E} : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H}')$ in a state $\varrho \in \mathcal{T}(\mathcal{H}')$ one can pick a large M and measure Φ_{can} in a state $\varrho_M = \Theta_{M*}(\varrho)/\text{tr}[\Theta_{M*}(\varrho)]$ to get

$$\mathrm{tr}[\varrho \mathbf{E}(X)] \approx \mathrm{tr}[\varrho \mathbf{E}_M(X)] = \mathrm{tr}[\varrho_M \Phi_{\mathrm{can}}(X)] \mathrm{tr}[\Theta_{M*}(\varrho)]$$

for a finite number of sets X which can constitute, e.g., an arbitrarily dense discretisation of the interval $[0, 2\pi)$. In this sense, a canonical phase resembles any spectral measure for which a similar (but stronger) property is known to be true [11, Theorem 3].

Note that, generally, one must use the ‘discretised’ version $\mathbf{E}_M$ of $\mathbf{E}$ above, since there is a Φ_{can} -continuous observable $\mathbf{E}$ such that $\Theta(\Phi_{\mathrm{can}}(X)) = \mathbf{E}(X)$, $X \in \mathcal{B}([0, 2\pi))$, does not hold for any channel $\Theta : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}')$. Indeed, let $\mathcal{H}' = L^2(\mu)$ and $\mathbf{E} = \mathbf{M}$ be the canonical spectral measure (associated with a minimal Naimark dilation of Φ_{can}). Let X be such that $0 \neq \mathbf{M}(X) \neq I$. Since $\mathbf{M}(X)$ is a projection there exists a unit vector $\psi \in L^2(\mu)$ so that $\mathbf{M}(X)\psi = \psi$. But $\Theta(\Phi_{\mathrm{can}}(X)) = \mathbf{M}(X)$ implies $\mathrm{tr}[\Phi_{\mathrm{can}}(X)\Theta_*(|\psi\rangle\langle\psi|)] = \langle\psi|\mathbf{M}(X)\psi\rangle = 1$ and, by the spectral decomposition $\Theta_*(|\psi\rangle\langle\psi|) = \sum_{i=1}^{\infty} \lambda_i |\zeta_i\rangle\langle\zeta_i|$, $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, $\zeta_i \in \mathcal{H}$, $\langle\zeta_i|\zeta_j\rangle = \delta_{ij}$, one gets $\sum_i \lambda_i \langle\zeta_i|\Phi_{\mathrm{can}}(X)\zeta_i\rangle = 1$, that is, for any $\lambda_i \neq 0$, $\langle\zeta_i|\Phi_{\mathrm{can}}(X)\zeta_i\rangle = 1 = 1 - \langle\zeta_i|\Phi_{\mathrm{can}}(X^c)\zeta_i\rangle$ or, equivalently, $\|\sqrt{\Phi_{\mathrm{can}}(X^c)}\zeta_i\|^2 = \langle\zeta_i|\Phi_{\mathrm{can}}(X^c)\zeta_i\rangle = 0$. Thus $\zeta_i - \Phi_{\mathrm{can}}(X)\zeta_i = \Phi_{\mathrm{can}}(X^c)\zeta_i = \sqrt{\Phi_{\mathrm{can}}(X^c)}\sqrt{\Phi_{\mathrm{can}}(X^c)}\zeta_i = 0$ and $\Phi_{\mathrm{can}}(X)$ has an eigenvalue 1; but this is impossible by virtue of Theorem 8.2.

Remark 16.3 Let $\mathcal{H}' = \mathcal{H}$ and assume that $\mathbf{E} : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ is a phase observable determined by a unit vector sequence $\{\eta_n\}_{n \in \mathbb{N}_0} \subset \mathcal{H}$, i.e.,

$$\mathbf{E}(X) = \sum_{n,m=0}^{\infty} \int_X \langle\psi'_n(\theta)|\psi'_m(\theta)\rangle \mu(\theta) |n\rangle\langle m|$$

where $\psi'_n = e_n \otimes \eta_n$ and hence $v_n^s = \delta_{ns} \eta_n$. Now

$$Y_M = \sum_{n=0}^{\infty} \sum_{s=0}^{2M} |s\rangle \otimes v_n^{s-M} \langle n| = \sum_{n=0}^M |n+M\rangle \otimes \eta_n \langle n|.$$

Define an isometry

$$Y = \sum_{n=0}^{\infty} |n\rangle \otimes \eta_n \langle n|$$

and the corresponding channel

$$\Theta(B) = Y^*(B \otimes I)Y, \quad B \in \mathcal{L}(\mathcal{H}),$$

whose predual operator $\Theta_* : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ is trace-preserving. It is easy to verify that, for all $X \in \mathcal{B}([0, 2\pi))$,

$$\lim_{M \rightarrow \infty} \Theta_M(\Phi_{\mathrm{can}}(X)) = \mathbf{E}(X) = \Theta(\Phi_{\mathrm{can}}(X))$$

ultraweakly. Hence, we have a stronger result: to obtain the measurement outcome statistics of a phase observable $\mathbf{E}$ in a state $\varrho \in \mathcal{T}(\mathcal{H})$ one can measure Φ_{can} in a state $\Theta_*(\varrho)$ to get

$$\text{tr}[\varrho \mathbf{E}(X)] = \text{tr}[\Theta_*(\varrho) \Phi_{\text{can}}(X)]$$

for all $X \in \mathcal{B}([0, 2\pi))$.

16.3 Phase Space Phase Observables

In Sect. 8.4, we defined covariant phase space observables

$$\mathbf{G}^K(Z) = \int_Z W(q, p) K W(q, p)^* \frac{dq dp}{2\pi}, \quad Z \in \mathcal{B}(\mathbb{R}^2),$$

where $K \geq 0$ and $\text{tr}[K] = 1$. In Sect. 19.3, we will show that any $\mathbf{G}^K$ can be measured by double homodyne detection. Using (modified) polar coordinates $r = \sqrt{(q^2 + p^2)}/2$ and $\theta = \arctan(p/q)$ one can express $\mathbb{R}^2$ of the form $[0, \infty) \times [0, 2\pi)$ and the Lebesgue measure $dq dp$ becomes $2r dr d\theta = dr^2 d\theta$. Defining $W(re^{i\theta}) = W(\sqrt{2}r \cos \theta, \sqrt{2}r \sin \theta)$ we get the angle margin observable

$$\Phi^K(X) = \int_X \left[\int_0^\infty W(re^{i\theta}) K W(re^{i\theta})^* dr^2 \right] \frac{d\theta}{2\pi}, \quad X \in \mathcal{B}([0, 2\pi)).$$

Similarly, one can define the radial margin observable $\mathbf{R}^K : \mathcal{B}([0, \infty)) \rightarrow \mathcal{L}(\mathcal{H})$ of $\mathbf{G}^K$.

Let $K_\lambda = \sum_{s \in \mathbb{N}_0} \lambda_s |s\rangle \langle s|$ where $\lambda = \{\lambda_s\}_{s \in \mathbb{N}_0}$, $\lambda_s \geq 0$ and $\sum_{s \in \mathbb{N}_0} \lambda_s = 1$. Since

$$\mathbf{G}^{K_\lambda}(Z) = \sum_{s \in \mathbb{N}_0} \lambda_s \mathbf{G}^{|s\rangle}(Z), \quad Z \in \mathcal{B}(\mathbb{R}^2),$$

(weakly) and $e^{i\alpha N} W(re^{i\theta}) = W(re^{i(\theta+\alpha)}) e^{i\alpha N}$ it follows easily that Φ^{K_λ} is a phase observable. Indeed, it can be shown [12, Theorem 4.1] that also the converse holds: if Φ^K is phase shift covariant then $K = K_\lambda$ for some λ .

From the formula

$$\langle n | W(re^{i\theta}) | s \rangle = e^{i(n-s)\theta} f_n^s(r^2) e^{-r^2/2}$$

where

$$f_n^s(x) = (-1)^{\max\{0, s-n\}} \sqrt{\frac{(\min\{n, s\})!}{(\max\{n, s\})!}} x^{|s-n|/2} L_{\min\{n, s\}}^{|s-n|}(x)$$

and

$$L_k^\alpha(x) = \sum_{l=0}^k (-1)^l \binom{k+\alpha}{k-l} \frac{x^l}{l!}$$

is the associated Laguerre polynomial, one sees that the elements of the structure matrix $(c_{m,n}^{[s]})$ of the phase observable $\Phi^{[s]}$ are of the form

$$c_{nm}^{[s]} = \int_0^\infty f_n^s(x) f_m^s(x) e^{-x} dx \in \mathbb{R}.$$

Hence, the η_n -vectors of $(c_{nm}^{[s]})$ are the unit vectors f_n^s , $n \in \mathbb{N}_0$, of $L^2(\mathbb{R}_+, e^{-x} dx)$, where $\mathbb{R}_+$ is the set of positive reals. All the functions f_n^s , $n \in \mathbb{N}_0$, are linearly independent, and therefore $\text{rank } \Phi^{[s]} = \infty$.

Now the structure matrix elements of any Φ^{K_λ} are

$$c_{nm}^{K_\lambda} = \sum_{s=0}^\infty \lambda_s c_{nm}^{[s]} = \langle f_n^\lambda | f_m^\lambda \rangle \in \mathbb{R}$$

where $f_n^\lambda = \sum_{s=0}^\infty \sqrt{\lambda_s} f_n^s \otimes |s\rangle$ is a unit vector of $L^2(\mathbb{R}_+, e^{-x} dx) \otimes \mathcal{H}$, and it follows that $\text{rank } \Phi^{K_\lambda} = \infty$.

Proposition 16.1 *Let Φ be a phase observable. If $\text{rank } \Phi > 1$ and $c_{nm} \in \mathbb{R}$ for all $n, m \in \mathbb{N}_0$, then Φ is not an extreme point of $\mathcal{O}_{\text{ph}}$. Especially, Φ^{K_λ} is never extremal.*

Proof Let η_n -vectors determine the structure of Φ and choose two linearly independent vectors, say, η_m and η_n . Define a nonzero bounded operator $D = |\eta_m\rangle\langle\eta_n| - |\eta_n\rangle\langle\eta_m|$ on $\mathcal{K}$. For every $j \in \mathbb{N}_0$, we have

$$\langle\eta_j|D\eta_j\rangle = c_{jm}c_{nj} - c_{jn}c_{mj} = c_{jm}c_{nj} - \overline{c_{nj}c_{jm}} = c_{nj}c_{jm} - c_{nj}c_{jm} = 0$$

so that Φ is not extreme by Theorem 16.4. □

To conclude, the experimentally realisable phase observables Φ^{K_λ} are not extremals and, thus, they are affected by the classical noise related to fluctuations in the measuring procedure. This suggests that better phase measurement schemes could be found in future. As a step in this direction, it has been shown that Φ_{can} can be measured via eight-port homodyne detection, if the input state ϱ and the reference state $|0\rangle$ are first entangled by using a two-mode unitary operator (a unitary channel) [13].

16.4 Number-Phase Complementarity

Let $\Phi : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ be a phase observable. Since the condition

$$|n\rangle\langle n|\Phi(X) = \Phi(X)|n\rangle\langle n|$$

for all $n \in \mathbb{N}_0$ and $X \in \mathcal{B}([0, 2\pi))$ implies that $\Phi(X) = \mu(X)I$, it follows that the number observable $\mathbf{N}$ and any nontrivial phase observable Φ are incompatible (Proposition 4.8). As in the case of position and momentum observables, Sect. 15.1, one may ask if number and phase are also complementary observables. Following [14, Section IV.2.3] we say that a phase $\Phi : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ and the number $\mathbf{N}$ are *complementary* if for any finite set $\{n_1, \dots, n_k\} \subset \mathbb{N}_0$ and for any $X \in \mathcal{B}([0, 2\pi))$, for which $0 \neq \Phi(X) \neq I$, the greatest lower bound

$$\left(\sum_{i=1}^k |n_i\rangle\langle n_i| \right) \wedge \Phi(X) = 0.$$

For any n and X , if A is an effect operator below $|n\rangle\langle n|$ and $\Phi(X)$, then $A = \lambda|n\rangle\langle n|$ and $\lambda \leq \ell(X)/2\pi$, and thus $|n\rangle\langle n| \wedge \Phi(X) = \frac{\ell(X)}{2\pi}|n\rangle\langle n|$. It is then a simple exercise to check that, for instance, any elementary phase, that is, a phase with only two nonzero nondiagonal matrix elements $c_{st} = \overline{c_{ts}}$ (cf. Remark 16.1), is not complementary with number. It is an *open question* whether any phase observable Φ and $\mathbf{N}$ are complementary.

Complementary observables are also probabilistically complementary: we say that $\Phi : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{N}$ are *probabilistically complementary* if

$$\sum_{i=1}^k |\langle \psi | n_i \rangle|^2 = 1 \text{ implies that } 0 < \langle \psi | \Phi(X) \psi \rangle < 1,$$

$$\langle \psi | \Phi(X) \psi \rangle = 1 \text{ implies that } 0 < \sum_{i=1}^k |\langle \psi | n_i \rangle|^2 < 1,$$

for all $\psi \in \mathcal{H}$, for all nonempty sets $\{n_1, \dots, n_k\} \subset \mathbb{N}_0$, and for any $X \in \mathcal{B}([0, 2\pi))$ such that $0 \neq \Phi(X) \neq I$ (i.e. $0 < \mu(X) < 1$).

Proposition 16.2 *For each state $\varrho \in \mathcal{S}(\mathcal{H})$ and a phase observable Φ ,*

$$\text{tr}[\varrho\Phi(X)] < 1$$

for all $X \in \mathcal{B}([0, 2\pi))$ such that $\mu(X) < 1$. Especially, any phase Φ and $\mathbf{N}$ are probabilistically complementary.

Proof Let $X \in \mathcal{B}([0, 2\pi))$ be such that $\mu(X) < 1$ and assume that there exists a $\psi \in \mathcal{H}$ for which $\langle \psi | \Phi_{\text{can}}(X) \psi \rangle = 1$. As before, we see that this implies that $\Phi_{\text{can}}(X)\psi =$

ψ which is impossible, see Theorem 8.2. Hence, one must have $\langle \psi | \Phi_{\text{can}}(X) \psi \rangle < 1$ for all $\psi \in \mathcal{H}$ and X such that $\mu(X) < 1$.

Let then Φ be a phase observable and Θ a channel for which $\Theta(\Phi_{\text{can}}(X)) = \Phi(X)$ (see Remark 16.3). Let ϱ be a state, and let

$$\Theta_*(\varrho) = \sum_{k=1}^{\infty} p_k |\psi_k\rangle \langle \psi_k|,$$

$p_k \geq 0$, $\sum_k p_k = 1$, $\psi_k \in \mathcal{H}$, $\|\psi_k\| = 1$, be the spectral decomposition of the state $\Theta_*(\varrho)$. Now, for all X such that $\mu(X) < 1$,

$$\text{tr}[\varrho \Phi(X)] = \text{tr}[\Theta_*(\varrho) \Phi_{\text{can}}(X)] = \sum_{k=1}^{\infty} p_k \underbrace{\langle \psi_k | \Phi_{\text{can}}(X) \psi_k \rangle}_{< 1} < 1.$$

The last claim is trivial. □

The above proposition shows that we cannot find a state which is localised in some X (essentially) other than $[0, 2\pi)$, that is, a phase observable cannot be localised.

There is yet another intuitive notion of complementarity, to which we refer as value complementarity. The idea is that if one of the observables assumes a sharp value, then the other should be uniformly distributed. There are technical problems in formalising this idea because continuous quantities do not have eigenvalues and for unbounded value sets a uniform distribution cannot be easily defined. Here is our attempted definition: a phase Φ and $\mathbf{N}$ are *value complementary* if the following conditions are satisfied:

1. for any number eigenstate $|n\rangle$, the phase distribution $X \mapsto \langle n | \Phi(X) | n \rangle$ is uniform;
2. for any sequence of vector states (φ_r) , if the phase distributions $X \mapsto \langle \varphi_r | \Phi(X) \varphi_r \rangle$ approach a delta distribution (on the circle) centered at some $\alpha \in [0, 2\pi)$, then the number distributions $n \mapsto |\langle \varphi_r | n \rangle|^2$ get increasingly uniform.

The probability distribution of any phase observable Φ is uniform in every number state $|n\rangle$, indeed, $\langle n | \Phi(X) | n \rangle = \mu(X)$. Thus the value complementarity of Φ and $\mathbf{N}$ depends only on the second of the above conditions.

Proposition 16.3 *Let $\{\psi_i\}_{i \in \mathbb{N}}$ be a sequence of unit vectors for which the probability measures $X \mapsto \langle \psi_i | \Phi(X) \psi_i \rangle$ tend (with $i \rightarrow \infty$) to a point measure δ_α , $\alpha \in [0, 2\pi)$. Then the number probabilities $|\langle \psi_i | n \rangle|^2$ tend to zero for all n . Especially, any phase Φ and $\mathbf{N}$ are value complementary.*

Proof Let Φ be a phase observable with the structure matrix $(c_{nm})_{n,m \in \mathbb{N}_0}$, and put $\Phi_\psi(X) = \langle \psi | \Phi(X) \psi \rangle$, for any unit vector ψ . Let $\{\psi_i\}_{i \in \mathbb{N}} \subset \mathcal{H}$ be a sequence of unit vectors such that

$$\lim_{i \rightarrow \infty} \Phi_{\psi_i}([0, x)) = \delta_\alpha([0, x)) = \begin{cases} 0 & \text{when } 0 < x < \alpha \\ 1 & \text{when } \alpha < x \leq 2\pi \end{cases}$$

where δ_α is a Dirac measure concentrated on the point $\alpha \in [0, 2\pi)$. This implies that $\lim_{i \rightarrow \infty} \int_0^{2\pi} e^{ik\theta} d\Phi_{\psi_i}(\theta) = e^{ik\alpha}$ for all $k \in \mathbb{N}$ (see e.g. [15, Theorem 26.3]). Then

$$\left| \int_0^{2\pi} e^{ik\theta} d\Phi_{\psi_i}(\theta) \right|^2 = \left| \sum_{n=0}^{\infty} c_{n,n+k} \overline{\langle n | \psi_i \rangle} \langle n+k | \psi_i \rangle \right|^2 \leq \sum_{m=k}^{\infty} |\langle m | \psi_i \rangle|^2 \rightarrow 1$$

for all $k \in \mathbb{N}$ if and only if

$$\sum_{n=0}^p |\langle n | \psi_i \rangle|^2 \rightarrow 0$$

for all $p \in \mathbb{N}$. □

Note that, in the context of the above proposition, $\langle \psi_i | N \psi_i \rangle \rightarrow \infty$. This situation, where the number and, thus, the energy gets large and the phase is arbitrarily well defined, corresponds to the classical limit for a single mode optical field.

For any phase Φ and $\epsilon \in (0, 1)$,

$$X \mapsto \Phi_\epsilon(X) = \epsilon \mu(X) I + (1 - \epsilon) \Phi(X)$$

is a phase observable which is not complementary with number. Indeed, $\Phi_\epsilon(X) \geq \epsilon |k\rangle \langle k|$ for all $k \geq 0$ so that $\epsilon |k\rangle \langle k|$ is below $\Phi_\epsilon(X)$ and $|k\rangle \langle k|$. This shows that every phase observable Φ is “arbitrarily close” to a phase observable Φ_ϵ which is not complementary with the number.

This observation, which generalises to every canonical pair, implies that the complementarity of such pairs, given that it holds, is not strictly testable. This is also true for probabilistic complementarity in the sense that finite statistics can never confirm strictly whether a given event has probability equal to one. Nevertheless, complementarity indicates a relation between two observables which is robust under small imprecisions: if a pair of observables is complementary, then a “nearby” noncomplementary pair will only allow “small” positive joint lower bounds between their positive operators.

16.5 Other Phase Theories

The difficulty in finding a selfadjoint phase operator has led to several phase theories (for an overview, see e.g. [3, 5, 7]) including the following attempts.

1. Phase observables are defined as phase shift covariant observables [1, 8, 9, 12, 14].
2. Selfadjoint phase operators acting on $\mathcal{H}$ are defined by some quantisation rules [2, 3, 16–27].
3. Phase observables are defined as phase distributions [5, 7, 24, 28–31].
4. Selfadjoint phase operators are defined in some subspaces of $\mathcal{H}$ [6, 32].
5. Selfadjoint phase operators are defined in some extensions of $\mathcal{H}$ [33–35].
6. Phase observables are defined as phase difference operators of two modes acting on $\mathcal{H} \otimes \mathcal{H}$ [16, 20–22, 36–39].

The common factor in most of these theories is that they present a phase observable as a selfadjoint operator. None of those theories satisfies all the requirements needed for a satisfactory phase theory. These requirements are (see e.g. [6, 24, 40, 41]):

1. The Hilbert space of a phase observable is spanned by the number states.
2. The set of values of a phase observable is a phase interval, which can be chosen to be $[0, 2\pi)$.
3. The phase distribution in a number state is uniform.
4. The number observable generates phase shifts.
5. The phase observable generates number shifts.

These requirements force a phase observable to be the canonical phase which is not a projection measure. Hence, if we want to represent a phase observable as a selfadjoint operator (or a projection measure) we have to give up some of the above requirements.

It can be shown [42] that all these theories have connections to the covariant theory. For example, Dirac's original 'phase quantisation' (see, Sect. 19.3) can be seen just as a polar decomposition

$$a = \int_0^{2\pi} e^{i\theta} d\Phi_{\text{can}}(\theta) \sqrt{N}$$

where the associated phase observable is the canonical phase Φ_{can} . Indeed, the canonical phase is associated with the polar decompositions of *all* (unilateral) shift operators [43]. In many cases the suggested single-mode phase operator is a (generalised) operator integral of some function of the phase angle with respect to a covariant phase observable (or a covariant sequilinear form measure which has a similar structure as the covariant phase observables [42]). Also, two mode theories seem to give covariant phase observables of the first mode when the phase of the second (reference) mode is fixed (or uniform). These notions justify the covariant approach to the quantum phase problem.

There are still some open problems left. For instance, direct measurements of the canonical phase have not yet been reported; for a recent model, see [13]. Moreover, it is not known if the canonical phase and the number are complementary observables in the strict sense of the measurement theory. Another interesting open question is whether the canonical phase minimises the minimum (phase) variance in any state [42].

As we have seen, the canonical phase has many properties which distinguish it from other phase observables and allow us to call it ‘canonical’—a comprehensive list of special properties of the canonical phase is gathered in p. 51 of [42]. In addition to the previous ones, Φ_{can} is determined by the Susskind–Glogover [21] phase states $|\theta\rangle = \sum_{n \in \mathbb{N}_0} e^{in\theta} |n\rangle \in V^\times$ where $V^\times$ is the algebraic dual of the vector space $V = \text{lin}\{|n\rangle \mid n \in \mathbb{N}_0\}$. Indeed,

$$\Phi_{\text{can}}(X) = \sum_{n,m \in \mathbb{N}_0} \int_X e^{i(n-m)\theta} d\mu(\theta) |n\rangle \langle m| = \int_X |\theta\rangle \langle \theta| d\mu(\theta).$$

Note that vectors $|\theta\rangle$ give the London [28] phase distribution $\theta \mapsto |\langle \theta | \psi \rangle|^2$ (formally) for all $\psi \in \mathcal{H}$. Another rank-1 phase observables are unitarily equivalent with Φ_{can} , but if the rank of a phase observable Φ is greater than 1, then one needs more than one set of (generalised) phase states $|k; \theta\rangle$ to determine its structure, that is,

$$\Phi(X) = \sum_{n,m \in \mathbb{N}_0} \langle \eta_n | \eta_m \rangle \int_X e^{i(n-m)\theta} d\mu(\theta) |n\rangle \langle m| = \sum_{k=1}^{\text{rank } \Phi} \int_X |k; \theta\rangle \langle k; \theta| d\mu(\theta)$$

where $|k; \theta\rangle = \sum_{n \in \mathbb{N}_0} \langle \eta_n | k \rangle e^{in\theta} |n\rangle \in V^\times$ and vectors

$$\eta_n \in \mathcal{K} = \overline{\text{lin}\{|k\rangle \mid k < \text{rank } \Phi + 1\}}$$

are as in Theorem 16.1. Next we show that Φ can be discretised.

Let $s, t \in \mathbb{N}_0$, $\theta_{s,t} = 2\pi t/(s+1)$, and $\mathcal{H}_s = \text{lin}\{|n\rangle \mid n = 0, 1, \dots, s\}$ an $(s+1)$ -dimensional subspace of $\mathcal{H}$. Define a discrete observable $\Phi^s : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H}_s) \subset \mathcal{L}(\mathcal{H})$ by

$$\begin{aligned} \Phi^s(X) &= \frac{1}{s+1} \sum_{n,m=0}^s \langle \eta_n | \eta_m \rangle \sum_{\theta_{s,t} \in X} e^{i(n-m)\theta_{s,t}} |n\rangle \langle m| \\ &= \frac{1}{s+1} \sum_{\theta_{s,t} \in X} \sum_{k=1}^{\text{rank } \Phi} |s, k; \theta_{s,t}\rangle \langle s, k; \theta_{s,t}| \end{aligned}$$

where $|s, k; \theta\rangle = \sum_{n=0}^s \langle \eta_n | k \rangle e^{in\theta} |n\rangle \in \mathcal{H}_s$. The support of Φ^s is the finite set $\{\theta_{s,t} \mid t = 0, 1, \dots, s\}$ which forms a partition of the circle into $s+1$ pieces. It is easy to show that, for all $0 \leq a < b \leq 2\pi$,

$$\lim_{s \rightarrow \infty} \Phi^s([a, b]) = \Phi([a, b])$$

ultraweakly in $\mathcal{L}(\mathcal{H})$. Hence, to get the statistics of a measurement of Φ in a state ϱ , one can pick a large s and measure Φ^s in a state $\tilde{\varrho}_s = \varrho_s / \text{tr}[\varrho_s]$ where $\varrho_s = \sum_{n,m=0}^s \langle n | \varrho | m \rangle |n\rangle \langle m|$ to get

$$\text{tr}[\varrho \Phi([a, b])] \approx \text{tr}[\tilde{\varrho}_s \Phi^s([a, b])] \text{tr}[\varrho_s].$$

It is interesting to note that Φ^s is covariant with respect to the discrete phase shifts $\theta_{s,t} \mapsto \theta_{s,t} + \theta_{s,u}$ generated by a truncated number operator $N^s = \sum_{n=0}^s n |n\rangle \langle n|$ [41]. Moreover, Φ^s is a spectral measure for all $s \in \mathbb{N}_0$ if and only if Φ is unitarily equivalent with Φ_{can} [41]. In this sense, Φ_{can} is approximately a covariant spectral measure. Note that, in the case of the canonical phase, unit vectors

$$\frac{1}{\sqrt{s+1}} |s, 1; \theta_{s,t}\rangle = \frac{1}{\sqrt{s+1}} \sum_{n=0}^s e^{in\theta_{s,t}} |n\rangle$$

are called as Pegg-Barnett phase states and the Pegg-Barnett formalism [6] can be viewed as a discretisation process of the canonical phase.

16.6 Exercises

1. Calculate the cyclic moments $V_{\Phi}^{(k)}$ and the first moment $\int \theta d\Phi(\theta)$ of a phase observable Φ with (c_{nm}) . Find a necessary and sufficient property for (c_{nm}) such that $V_{\Phi}^{(k+l)} = V_{\Phi}^{(k)} V_{\Phi}^{(l)}$ holds. Does the phase space phase observable $\Phi^{(0)}$ satisfy this condition?
2. Prove that (i) equals (ii) and (iii) implies (i) in Theorem 16.3.
3. Show that $\Phi'(X) = \int_0^{2\pi} \varrho(X \dot{+} (-\theta)) d\Phi(\theta)$ defines a phase observable Φ' with the structure matrix $(c_{nm} f_{m-n})$, where (c_{nm}) is a structure matrix of a phase Φ and f_k is the k th Fourier coefficient of a probability measure ϱ on the circle.
4. Let $\xi \in \mathbb{C}$, $\|\xi\| \leq 1$. Define $\eta_{2n}^{\xi} = |0\rangle$ and $\eta_{2n+1}^{\xi} = \xi |0\rangle + \sqrt{1 - |\xi|^2} |1\rangle$ for all $n \in \mathbb{N}_0$ and a matrix $(c_{n,m}^{\xi})$ by $c_{n,m}^{\xi} = \langle \eta_n^{\xi} | \eta_m^{\xi} \rangle$. Thus, $(c_{n,m}^{\xi})$ is the structure matrix of the *chess-board phase observable* Φ_{ξ} [43]. What is the rank of Φ_{ξ} ? Show that Φ_{ξ} is not an extreme point of $\mathcal{O}_{\text{ph}}$ when $\|\xi\| \neq 1$. Moreover, show that, if $\|\xi\| = 1$, then Φ_{ξ} is an extremal of $\mathcal{O}$.
5. Show that $\Theta_M(\Phi_{\text{can}}(X)) = E_M(X)$ and $\lim_{M \rightarrow \infty} E_M(X) = E(X)$ (ultraweakly) where Θ_M , E_M , and E are defined in Sect. 16.2.
6. Show that the channel Θ_* defined in Remark 16.3 preserves pure states (resp. is bijective) if and only if E is unitarily equivalent with Φ_{can} .
7. Prove that $\Phi^{K_{\lambda}}$ is a phase observable if K_{λ} is diagonal in the number basis. Determine its structure matrix.

8. Show that $|n\rangle\langle n|\Phi(X) = \Phi(X)|n\rangle\langle n|$ for all X and n implies $\Phi(X) = \mu(X)I$, and that the commutativity domain of $\mathbf{N}$ and Φ_{can} is $\{0\}$.
9. Prove that $a = \int_0^{2\pi} e^{i\theta} d\Phi_{\text{can}}(\theta) \sqrt{N}$ where a is the lowering operator.
10. Show that, for all $0 \leq a < b \leq 2\pi$,

$$\lim_{s \rightarrow \infty} \Phi^s([a, b]) = \Phi([a, b])$$

(ultraweakly) in $\mathcal{L}(\mathcal{H})$ where Φ^s is the discretisation of a phase observable Φ .

References

1. Noh, J.W., Fougères, A., Mandel, L.: Measurement of the quantum phase by photon counting. *Phys. Rev. Lett.* **67**, 1426–1429 (1991)
2. Dirac, P.A.M.: The quantum theory of the emission and absorption of radiation. *Proc. R. Soc. Lond. Ser. A* **114**(767), 243–265 (1927)
3. Dubin, D.A., Hennings, M.A., Smith, T.B.: *Mathematical Aspects of Weyl Quantization and Phase*. World Scientific Publishing Co., Inc., River Edge (2000)
4. Perinová, V., Luks, A., Perina, J. *Phase in Optics*. World Scientific Series in Contemporary Chemical Physics, vol. 15. World Scientific (1998)
5. Lynch, R.: The quantum phase problem: a critical review. *Phys. Rep.* **256**(6), 367–436 (1995)
6. Pegg, D.T., Barnett, S.M.: Tutorial review: quantum optical phase. *J. Mod. Opt.* **44**(2), 225–264 (1997)
7. Schleich, W.P., Barnett, S.M.: Quantum phase and phase dependent measurements. *Phys. Scripta* **1993**(T48), 3 (1993)
8. Helstrom, C.: *Quantum Detection and Estimation Theory*. Academic Press, New York (1976)
9. 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
10. Dunford, N., Schwartz, J.T.: *Linear Operators. Part I*. Wiley Classics Library. Wiley, New York (1988). Reprint of the 1958 original, A Wiley-Interscience Publication
11. Pellonpää, J.-P.: Complete characterization of extreme quantum observables in infinite dimensions. *J. Phys. A* **44**(8), 085304 (2011)
12. Lahti, P., Pellonpää, J.-P.: Covariant phase observables in quantum mechanics. *J. Math. Phys.* **40**(10), 4688–4698 (1999)
13. Pellonpää, J.-P., Schultz, J.: Measuring the canonical phase with phase-space measurements. *Phys. Rev. A* **88**, 012121 (2013)
14. Busch, P., Grabowski, M., Lahti, P.: *Operational Quantum Physics*. Lecture Notes in Physics Monographs, vol. 31. Springer, 2nd corrected edition (1997)
15. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
16. Carruthers, P., Nieto, M.M.: Phase and angle variables in quantum mechanics. *Rev. Mod. Phys.* **40**, 411–440 (1968)
17. Galindo, A.: Phase and number. *Lett. Math. Phys.* **8**(6), 495–500 (1984)
18. Garrison, J.C., Wong, J.: Canonically conjugate pairs, uncertainty relations, and phase operators. *J. Math. Phys.* **11**, 2242–2249 (1970)
19. Lerner, E.C., Huang, H.W., Walters, G.E.: Some mathematical properties of oscillator phase operators. *J. Math. Phys.* **11**, 1679–1684 (1970)
20. Lévy-Leblond, J.-M.: Who is afraid of nonhermitian operators? A quantum description of angle and phase. *Ann. Physics* **101**(1), 319–341 (1976)

21. Susskind, L., Glogower, J.: Quantum mechanical phase and time operator. *Physics* **1**, 49–62 (1964)
22. Volkin, H.C.: Phase operators and phase relations for photon states. *J. Math. Phys.* **14**, 1965–1976 (1973)
23. Werner, R.: Wigner quantisation of arrival time and oscillator phase. *J. Phys. A* **21**(24), 4565–4575 (1988)
24. Royer, A.: Phase states and phase operators for the quantum harmonic oscillator. *Phys. Rev. A* **53**, 70–108 (1996)
25. Rocca, F., Sirugue, M.: Phase operator and condensed systems. *Commun. Math. Phys.* **34**, 111–121 (1973)
26. Turski, Ł.A.: The velocity operator for many-boson systems. *Physica* **57**, 432–444 (1972)
27. Paul, H.: Phase of a microscopic electromagnetic field and its measurement. *Fortschr. Physik* **22**, 657–689 (1974)
28. London, F.: Winkelvariable und kanonische Transformationen in der Undulationsmechanik. *Z. Physik* **40**(3–4), 193–210 (1927)
29. Freyberger, M., Schleich, W.: Photon counting, quantum phase, and phase-space distributions. *Phys. Rev. A* **47**, R30–R33 (1993)
30. Freyberger, M., Vogel, K., Schleich, W.: From photon counts to quantum phase. *Phys. Lett. A* **176**(1–2), 41–46 (1993)
31. Vogel, W., Schleich, W.: Phase distribution of a quantum state without using phase states. *Phys. Rev. A* **44**, 7642–7646 (1991)
32. Popov, V.N., Yarunin, V.S.: A photon phase operator. *Teoret. Mat. Fiz.* **89**(3), 395–401 (1991)
33. Newton, R.G.: Quantum action-angle variables for harmonic oscillators. *Ann. Phys.* **124**(2), 327–346 (1980)
34. Ozawa, M.: Phase operator problem and macroscopic extension of quantum mechanics. *Ann. Phys.* **257**(1), 65–83 (1997)
35. Vaccaro, J.A.: Phase operators on Hilbert space. *Phys. Rev. A* **51**, 3309–3317 (1995)
36. Luis, A., Sánchez-Soto, L.L.: Phase-difference operator. *Phys. Rev. A* **48**, 4702–4708 (1993)
37. Ban, M.: Phase operator in quantum optics. *Phys. Lett. A* **176**(1–2), 47–53 (1993)
38. Torgerson, J.R., Mandel, L.: Is there a unique operator for the phase difference of two quantum fields? *Phys. Rev. Lett.* **76**, 3939–3942 (1996)
39. Yu, S.: Quantized phase difference. *Phys. Rev. Lett.* **79**, 780–783 (1997)
40. Leonhardt, U.: *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge (1997)
41. Lahti, P., Pellonpää, J.-P.: The Pegg-Barnett formalism and covariant phase observables. *Phys. Scripta* **66**(1), 66 (2002)
42. Pellonpää, J.-P.: Covariant phase observables in quantum mechanics. Ph.D. thesis, University of Turku (2002)
43. Lahti, P., Pellonpää, J.-P.: Characterizations of the canonical phase observable. *J. Math. Phys.* **41**(11), 7352–7381 (2000)

Chapter 17

Time and Energy

Space and time have a double role in quantum mechanics. On the one hand, they form the arena for physical events; physical systems occur in spacetime. On the other hand, the symmetries of the underlying spacetime specify the description of the system and dictate its basic properties. In the relativistic case spacetime symmetries are expressed with respect to the Poincaré group, while in the nonrelativistic case the relevant symmetry group is the Galilei group. In this book we generally work in the frame of nonrelativistic quantum mechanics, in which case space and time can be separated. This chapter presents an analysis of covariant time observables in the case of classical Hamiltonian mechanics and analogous considerations in quantum mechanics. The division into periodic and nonperiodic systems is central.

17.1 The Concept of Time in Quantum Mechanics

As already pointed out in Sect. 15.1, position and momentum observables, as spectral measures, are uniquely determined by their behaviour under the symmetry transformations of the spatial translations $\mathbb{R}^3 \ni \mathbf{x} \mapsto \mathbf{x} + \mathbf{q} \in \mathbb{R}^3$ and the velocity boosts $\mathbb{R}^3 \ni \mathbf{u} \mapsto \mathbf{u} + \mathbf{v} \in \mathbb{R}^3$ given by the Weyl representation, position as a translation covariant and velocity boost invariant, momentum as a velocity boost covariant and translation invariant sharp observable. The selfadjoint operators generating spatial translations and velocity boosts along a given coordinate axis are identified as the momentum and position operators, respectively. In addition, Theorem 8.3 shows that any translation covariant and velocity boost invariant, resp. velocity boost covariant and translation invariant, observable is a smearing (by convolution with a probability measure) of the sharp position, resp. momentum.¹

¹The one-dimensional case of this theorem extends easily to the three-dimensional case.

Similarly, time translations $t \mapsto t + t'$, both as a semigroup $(\mathbb{R}^+, +)$ and a group $(\mathbb{R}, +)$, give rise to a symmetry expressed as a dynamical (semi)group $t \mapsto s_t \in \mathcal{L}(\mathcal{T}(\mathcal{H}))$, with respect to which one may define time shift covariant observables T for which $\mathsf{T}_{s_t(\varrho)}(X) = \mathsf{T}_\varrho(X + t)$ for each state $\varrho = s_0(\varrho)$ and all (Borel) subsets X of $\mathbb{R}^+$, or $\mathbb{R}$. As with position and momentum, this condition encodes the minimal requirement that a measurement of T at time $t > 0$ gives a result in $[t_1, t_2]$ with the same probability as a measurement of T at $t = 0$ gives a result in the shifted interval $[t_1 + t, t_2 + t]$.

Even in the case of a dynamical group, where $t \mapsto s_t$ can be implemented by a unitary representation $t \mapsto U_t$ defined by a selfadjoint operator H , see Sect. 9.4, one can solve the condition $\text{tr}[e^{itH} \varrho e^{-itH} \mathsf{T}(X)] = \text{tr}[\varrho \mathsf{T}(X + t)]$ only under additional constraints on the time shift-generating operator H , the Hamiltonian or energy operator. Clearly, such a T depends on H .² Moreover, if a solution exists, the spectral structure of H typically excludes a spectral measure solution, that is, there is no (selfadjoint) time operator—the so-called Pauli theorem [4]. In addition, though energy $\mathsf{H} = \mathsf{E}^H$ generates time shifts, time T (if exists) cannot, in general, generate energy shifts (as the spectrum of H has typically a complicated structure including unevenly distributed eigenvalues). In this sense, time-energy pairs are usually not fully analogous to the position-momentum pair.

These well-known complications have hampered the development of a systematic theory of time observables in quantum mechanics, even to the extent that one still finds statements in the research and textbook literature to the effect that “time is not an observable (but just a parameter—a ‘c-number’)”. However, it is undeniable that the time of the occurrence of various types of events or the duration of certain processes are measurable features of physical systems and should therefore be amenable to a theoretical description as an observable of the system. Examples are times of decay or emission of unstable particles, arrival time of a particle, or a time delay caused by a scattering or tunneling.

The literature on the problem of time in quantum mechanics is overwhelmingly large and a survey of it is beyond the scope of this monograph. For a recent exposition of the theory of time measurement in quantum mechanics, with numerous realistic examples, the reader is referred to the two-part monograph *Time in Quantum Mechanics* [5, 6]. In this chapter we shall concentrate only on some basic results on covariant time observables in quantum mechanics. Due to the complexity of the problem we shall start (Sect. 17.2) with a brief account of time in nonrelativistic classical mechanics. With this background the study of covariant time observables (Sect. 17.3) is divided into nonperiodic and periodic cases.

²In Galilei relativistic quantum mechanics the structure of the evolution generator, the energy operator, is determined to a large extent by the symmetry requirements [1–3]. Here we recall only that for a spinless free particle moving in $\mathbb{R}^3$, $H = \frac{1}{2m} \mathbf{P}^2$ whereas it is of the form $H = \frac{1}{2m} \sum_{i=1}^3 (\mathbf{P} - \mathbf{A}(\mathbf{Q}))^2 + V(\mathbf{Q})$ if the particle is moving in an external field described by vector and scalar potentials $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ (measurable functions), respectively. Here $\mathbf{P} = (P_1, P_2, P_3)$, $\mathbf{Q} = (Q_1, Q_2, Q_3)$.

A time observable which is covariant under the time shifts cannot commute with the shift generator, the energy operator. By Propositions 4.8 and 11.1 this excludes any joint measurement of time and energy. This raises questions concerning the meaning and validity of time-energy uncertainty relations, both in view of preparations and approximate joint measurements. Apart from some obvious applications of the preparations uncertainty relations of Sect. 12.3, these questions are largely open. Even in the case where time is not treated formally as an observable but as a parameter characteristic of the dynamics of the system, there are only a few conceptually clear and quantitatively meaningful formulations of time-energy uncertainty relations. Instances of such studies are the classic lifetime-energy relation as given, for instance, in [7, Sect. 4.2.11], or a more recent arrival time-energy relation of [8], where the arrival of a particle at a detector is modelled by an absorbing term added to the Hamiltonian. For a survey of some such relations we refer to the review [9].

17.2 Time in Nonrelativistic Classical Mechanics

In nonrelativistic mechanics, (universal) time is modelled as points t of $\mathbb{R}$. Let (Π, H) be a classical Hamiltonian system, that is, Π is a phase space (a $2n$ -dimensional smooth symplectic manifold) and $H : \Pi \rightarrow \mathbb{R}$ a (time-independent) Hamiltonian function of a mechanical system. Let $(q^1, \dots, q^n; p^1, \dots, p^n)$ be a (local) canonical chart of Π , and let $\dot{q}^i$ and $\dot{p}^i$ be the corresponding derivative coordinates (of the tangent bundle of Π). Now the motion of the system is a (smooth) curve $\gamma : (t_1, t_2) \rightarrow \Pi$ which satisfies the Hamilton equations

$$\dot{q}^i = \frac{\partial H}{\partial p^i}, \quad \dot{p}^i = -\frac{\partial H}{\partial q^i}, \quad i = 1, \dots, n.$$

A (local) time observable can be defined as a (smooth) function $T : U \rightarrow \mathbb{R}$, where $U \subset \Pi$ is open, which is a solution of the following *classical covariance system*:

$$\{T, H\} = 1 \tag{17.1}$$

on U . Here $\{f, g\}$ is the Poisson bracket of (smooth) real functions on Π locally expressed as

$$\{f, g\} = \sum_{i=1}^n \left(\frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p^i} - \frac{\partial f}{\partial p^i} \frac{\partial g}{\partial q^i} \right).$$

Now

$$\frac{d(T \circ \gamma)(t)}{dt} = (\{T, H\} \circ \gamma)(t) = 1$$

so that $T(\gamma(t)) = t + t_0$ for all $t \in (t_1, t_2)$ (such that $\gamma(t) \in U$) where t_0 is some constant. This means that the value of the time observable T on any trajectory of the system at the time t is $t + t_0$. Hence, one can use T as a *clock*. Also we can say that T transforms covariantly with respect to the time shifts generated by H .

Example 17.1 Let $\Pi = \mathbb{R}^2$ with the identity map as a canonical chart. Here we consider Π as the phase space of a falling particle: the coordinate q (resp. p) describes the (vertical) position (resp. momentum) of the particle. Let

$$H(q, p) = \frac{1}{2m}p^2 + mgq$$

be the Hamiltonian of the particle, with the constant $g \neq 0$ representing the gravitational acceleration. Note that $p = m\dot{q}$ where m is the mass of the particle. All solutions of the Hamilton equations are of the form

$$\gamma(t) = (q(t), p(t)) = (-gt^2/2 + \dot{q}_0 t + q_0, -mgt + m\dot{q}_0) \quad (17.2)$$

where q_0 and $\dot{q}_0$ are constants. If $T(q, p) = p/(-mg)$ then $\{T, H\} = 1$ on Π . Hence, one can use the momentum (and thus velocity) of a constantly accelerated particle as a clock. Of course, for a measurement of velocity, one needs another clock so that constantly accelerated particles are not very useful for time measurements.

There are other time observables too. For example, one could drop a particle from a height $q_0 > 0$ with the initial speed $\dot{q}_0 \in \mathbb{R}$ (at $t = 0$) and determine the time $t_{\text{arr}} > 0$ when the particle arrives at $q = 0$. Solving $\gamma(t_{\text{arr}}) = (0, p(t_{\text{arr}}))$ from Eq. (17.2) one gets the arrival time

$$t_{\text{arr}} = \frac{1}{g} \left(\dot{q}_0 + \sqrt{(\dot{q}_0)^2 + 2gq_0} \right).$$

Now the phase space is the open submanifold $\Pi_{\text{arr}} = \{(q, p) \in \mathbb{R}^2 \mid q > 0\}$ and the arrival time observable is

$$T_{\text{arr}}(q, p) = \frac{1}{mg} \left(p + \sqrt{2mH(q, p)} \right)$$

which gives $T_{\text{arr}}(q_0, m\dot{q}_0) = t_{\text{arr}}$. Note that $H > 0$ and $\{T_{\text{arr}}, H\} = -1$ when restricted to Π_{arr} . $\triangleleft$

Example 17.2 If $g = 0$ then Example 17.1 reduces to the free particle case with $H(q, p) = \frac{1}{2m}p^2$. Then $\gamma(t) = (\dot{q}_0 t + q_0, m\dot{q}_0)$, and $T(q, p) = mq/p$, $p \neq 0$, satisfies (locally) $\{T, H\} = 1$. Thus, the position of a free particle can be used as a clock (if $(p \circ \gamma)(t) = m\dot{q}_0 \neq 0$, that is, when the particle is actually moving). Indeed, similarly as light rays, free particles can be used as signals in the fundamental definitions of simultaneous events, time order, synchronisation of clocks, etc. In this case, the time of arrival at $q = 0$ is $T_{\text{arr}}(q, p) = -T(q, p) = -mq/p$ restricted

to the open submanifold determined by the constraint $T_{\text{arr}}(q, p) > 0$, i.e. $qp < 0$. Sometimes, one is only interested in particles coming from the left to the detector at $q = 0$ so that one must restrict $T_{\text{arr}}(q, p)$ to the open submanifold with $q < 0$ and $p > 0$. $\triangleleft$

Example 17.3 A problem of the clocks introduced in the preceding example is that they can be used for time measurements only for relatively short time intervals: the particle moves away and does not come back. A better way to measure time is to use periodic motion, that is, *oscillators* like an atom clock. In the one dimensional motion as above the potential function $V(q)$ must have a local minimum so that we can expand $V(q)$ into the Taylor series around the minimum point. By assuming that oscillations are small around the minimum, one can model the system as a harmonic oscillator:

$$H(q, p) = \frac{1}{2m}p^2 + \frac{k}{2}q^2,$$

where $k > 0$. Any solution of the Hamilton equations is of the form

$$\gamma(t) = (A \sin(\omega t + \theta_0), A\omega m \cos(\omega t + \theta_0)),$$

with $\omega = \sqrt{k/m}$, $A \geq 0$, and $\theta_0 \in [0, 2\pi)$. For

$$T(q, p) = \frac{1}{\omega} \left[\arctan \left(\omega m \frac{q}{p} \right) + \frac{\pi}{2} \right], \quad p \neq 0,$$

one has (locally) $\{T, H\} = 1$ so that T is a clock (if $A \neq 0$). The situation differs from the preceding example since now the value space of T is a bounded interval $(0, T_0/2)$ where $T_0 = 2\pi/\omega$ and *one must also count the number of (half) oscillations*. Note that the observable

$$T_A(q, p) = \frac{1}{\omega} \arcsin \left(\frac{q}{A} \right)$$

can also be used for time measurements for fixed motions, that is, for curves with a fixed amplitude A . Hence, the oscillator's position alone can be used for time measurements if the amplitude is known.

For example, consider the simple pendulum which consists of a small plumb bob of mass m swinging at the end of a light string of length l . The motion is along a circular arc defined by the angle φ . Let $q = l\varphi$. Now $p = m\dot{q}$ and

$$H(q, p) = \frac{1}{2m}p^2 + mgl[1 - \cos(q/l)] \approx \frac{1}{2m}p^2 + \frac{mg/l}{2}q^2$$

for small oscillations $q \approx 0$. Hence, $\omega = \sqrt{g/l}$ and, when $g \approx 9.81 \text{ m/s}^2$ and $l \approx 1.00 \text{ m}$, one gets a half-period of $T_0/2 = \pi/\omega \approx 1.00 \text{ s}$. This system is commonly known as the seconds pendulum. $\triangleleft$

17.3 Covariant Time Observables in Nonrelativistic Quantum Mechanics

Let $\mathcal{H}$ be a separable Hilbert space and $H : \mathcal{D}(H) \rightarrow \mathcal{H}$ a selfadjoint operator generating a unitary evolution of the system: $s_t(\varrho) = U_t \varrho U_t^*$, with $U_t = e^{itH/\hbar}$, $t \in \mathbb{R}$. Replacing $H/\hbar$ by H we may assume that $\hbar = 1$ (and easily put $\hbar$ back into the equations when needed). Now we have two cases: a time observable may be periodic or not. If it is periodic, then its value space is an interval, say, $[0, T_0)$, $T_0 > 0$, equipped with the addition modulo T_0 . This is the case when H is such that $U_{t+T_0} = U_t$ for all $t \in \mathbb{R}$. Actually in the periodic case, without restricting generality, the value space can be chosen to be the circle $\mathbb{T}$. In the nonperiodic case, one can choose $\mathbb{R}$ for the value space.

As shown in Sect. 8.7, H can be diagonalised, that is, we may (and will) assume that

$$\mathcal{H} = \int_{\sigma(H)}^{\oplus} \hat{\mathcal{H}}_{\hat{d}(\epsilon)} d\mu_H(\epsilon) \subset L^2_{\hat{\mathcal{H}}}(\mu_H) \cong L^2(\mu_H) \otimes \hat{\mathcal{H}}$$

where $\sigma(H)$ is the spectrum of H , $\hat{d}(\epsilon)$ is the dimension of the multiplicity space $\hat{\mathcal{H}}_{\hat{d}(\epsilon)} \subset \hat{\mathcal{H}}$ of $\epsilon \in \sigma(H)$, and μ_H is a positive measure such that the spectral measure $\mathbf{H}$ of H is absolutely continuous with respect to μ_H . Moreover, $(e^{itH}\psi)(\epsilon) = e^{it\epsilon}\psi(\epsilon)$ for all $t \in \mathbb{R}$, $\psi \in \mathcal{H}$, and μ_H -almost all $\epsilon \in \sigma(H)$. Recall that μ_H is a weighted counting measure on the discrete part of the spectrum. If μ_H is absolutely continuous with respect to the Lebesgue measure $d\epsilon$ we may (and will) assume that $d\mu_H(\epsilon) = d\epsilon$.

Nonperiodic Systems

We say that an observable $\mathbf{T} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is a *covariant time observable* (with respect to H) if it is time shift covariant,

$$U_t^* \mathbf{T}(\tau) U_t = \mathbf{T}(\tau + t)$$

for all $t \in \mathbb{R}$ and $\tau \in \mathcal{B}(\mathbb{R})$. Covariant time observables have been applied to describe the measurement of arrival and tunneling times of massive particles [10, 11].

One can define the symmetric operator $T = \mathbf{T}[1] = \int_{\mathbb{R}} t d\mathbf{T}(t)$ and integrate to get

$$U_t^* T U_t = T - tI.$$

Formally writing $e^{itH} = I + itH + \dots$ the above equation leads to the Heisenberg commutation relation $[T, H] = iI$ which corresponds to the classical equation (17.1). We recall (Remark 15.1) that, if T is selfadjoint, $[T, H] = iI$ can hold only when $\dim \mathcal{H} = \infty$ and either T or H is unbounded. Moreover, unless $\mathbf{T}$ is sharp, it cannot be the spectral measure of T .

Note that, by using the Fourier transform $\hat{T}(\epsilon) = \int_{\mathbb{R}} e^{-i\epsilon t} dT(t) \in \mathcal{L}(\mathcal{H})$, $\epsilon \in \mathbb{R}$, of T , the covariance condition reads

$$U_t^* \hat{T}(\epsilon) U_t = e^{i\epsilon t} \hat{T}(\epsilon).$$

This is the Weyl form of the covariance condition.

It is common in the literature to use an alternative covariance condition $U_t T(\tau) U_t^* = T(\tau + t)$ or, equivalently, $U_t^* T(\tau) U_t = T(\tau - t)$, which corresponds to the change $H \rightarrow -H$ in our definition above. Indeed, we used this type of definition in Chap. 16 for covariant phase observables (see Example 17.8).

Assume that there is a covariant time observable $T : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$. This implies that H cannot have any eigenvalues. Indeed, suppose that $E \in \mathbb{R}$ is an eigenvalue of H and let $\psi_E \in \mathcal{D}(H)$ be a corresponding normalised eigenvector, that is, $H\psi_E = E\psi_E$. Then the covariance condition $U_t^* T(\tau) U_t = T(\tau + t)$ yields

$$\langle \psi_E | T(\tau + t) \psi_E \rangle = \langle U_t \psi_E | T(\tau) U_t \psi_E \rangle = \langle e^{itE} \psi_E | T(\tau) e^{itE} \psi_E \rangle = \langle \psi_E | T(\tau) \psi_E \rangle$$

so that, by virtue of the uniqueness of the Haar measure of $\mathbb{R}$, the probability measure $\tau \mapsto \langle \psi_E | T(\tau) \psi_E \rangle$ must be the Lebesgue measure $d\tau$ multiplied by a constant. This is impossible since the Lebesgue measure is not finite. Thus, $\sigma(H) = \sigma_c(H)$. In addition, it can be shown that the continuous spectrum is absolutely continuous, that is, for any $\psi \in \mathcal{H}$ the measure $\mathbf{H}_\psi$ is absolutely continuous with respect to $d\epsilon$ [12, Corollary 1]. In other words, we may put $d\mu_H(\epsilon) = d\epsilon$.

Theorem 17.1 *An observable $T : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ is a covariant time observable (with respect to H) if and only if $\mathbf{H}$ is absolutely continuous with respect to $d\epsilon$ and there exists a weakly measurable family of isometries $W(\epsilon) : \hat{\mathcal{H}}_{d(\epsilon)} \rightarrow \hat{\mathcal{H}}$ such that*

$$\langle \varphi | T(\tau) \psi \rangle = \frac{1}{2\pi} \int_{\tau} \int_{\sigma(H)} \int_{\sigma(H)} e^{i(\epsilon' - \epsilon)t} \langle W(\epsilon) \varphi(\epsilon) | W(\epsilon') \psi(\epsilon') \rangle d\epsilon d\epsilon' dt$$

for all absolutely integrable $\varphi, \psi \in \mathcal{H}$ and $\tau \in \mathcal{B}(\mathbb{R})$. Moreover, T is a spectral measure if and only if $\sigma(H) = \mathbb{R}$ and, for (almost) all $\epsilon \in \mathbb{R}$, $d(\epsilon) = c \in \mathbb{N} \cup \{\infty\}$, $c < \dim \mathcal{H} + 1$, and any $W(\epsilon)$ is a unitary operator from $\hat{\mathcal{H}}_c$ onto a c -dimensional subspace of $\hat{\mathcal{H}}$ (which can be chosen to be $\hat{\mathcal{H}}_c$).

Proof First we show that T is absolutely continuous with respect to the Lebesgue measure $d\tau$. Let $\{e_n\}_{n=1}^{\dim \mathcal{H}}$ be an orthonormal basis of $\mathcal{H}$, and $\varrho = \sum_{n=1}^{\dim \mathcal{H}} \lambda_n |e_n\rangle\langle e_n|$, $\lambda_n > 0$, $\sum_n \lambda_n = 1$. Now T is absolutely continuous with respect to the measure $\tau \mapsto T_\varrho(\tau) = \text{tr}[\varrho T(\tau)]$. Then $T(\tau + t) = U_t^* T(\tau) U_t$ implies that T_ϱ is quasi-invariant, that is, the shifted measures $\tau \mapsto T_\varrho(\tau + t)$, $t \in \mathbb{R}$, are mutually absolutely continuous. Hence, T_ϱ is absolutely continuous with respect to the Lebesgue measure.

Let (Q, J) be a Naimark dilation of T where $Q : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{K})$ is the canonical spectral measure acting on $\mathcal{K} = L^2_{\mathcal{H}}(\mathbb{R})$ and $J : \mathcal{H} \rightarrow \mathcal{K}$ is an isometry such that $T(\tau) = J^* Q(\tau) J$. Let $\mathcal{F} : \mathcal{K} \rightarrow \mathcal{K}$ be the (unitary) Fourier–Plancherel operator

defined by $(\mathcal{F}\psi)(t) = (2\pi)^{-1/2} \int_{\mathbb{R}} e^{-it\epsilon} \psi(\epsilon) d\epsilon$ where $\psi \in \mathcal{K}$ is absolutely integrable. Define an observable $\mathbf{P} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{K})$ by $\mathbf{P}(\tau) = \mathcal{F}^* \mathbf{Q}(\tau) \mathcal{F}$, $\tau \in \mathcal{B}(\mathbb{R})$. Obviously,

$$\mathbf{T}(\tau) = W^* \mathbf{P}(\tau) W$$

where $W = \mathcal{F}^* J : \mathcal{H} \rightarrow \mathcal{K}$ is an isometry. Let $\mathcal{K}'$ be the closure of

$$\text{lin}\{\mathbf{P}(\tau) W \psi \mid \tau \in \mathcal{B}(\mathbb{R}), \psi \in \mathcal{H}\},$$

so that $W\mathcal{H} \subset \mathcal{K}'$ and $\tau \mapsto \mathbf{P}'(\tau) = \mathbf{P}(\tau)|_{\mathcal{K}'} \in \mathcal{L}(\mathcal{K}')$ is a spectral measure. Define a strongly continuous unitary representation $\mathbb{R} \ni t \mapsto \hat{U}_t \in \mathcal{L}(\mathcal{K}')$ by

$$\hat{U}_t \mathbf{P}'(\tau) W \psi = \mathbf{P}'(\tau - t) W U_t \psi$$

for all $t \in \mathbb{R}$, $\tau \in \mathcal{B}(\mathbb{R})$, and $\psi \in \mathcal{H}$. It is easy to show that $\mathbf{P}'(\tau + t) = \hat{U}_t^* \mathbf{P}'(\tau) \hat{U}_t$ so that one must have $(\hat{U}_t \psi)(\epsilon) = e^{it\epsilon} \psi(\epsilon)$ for all $\psi \in \mathcal{K}'$.

Since $W U_t = \hat{U}_t W$ the isometry $W : \mathcal{H} \rightarrow \mathcal{K}'$ is decomposable [13], $W = \int_{\mathbb{R}}^{\oplus} W(\epsilon) d\epsilon$, and we get

$$\begin{aligned} \langle \varphi | \mathbf{T}(\tau) \psi \rangle &= \frac{1}{2\pi} \int_{\tau} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i(\epsilon' - \epsilon)t} \langle (W\varphi)(\epsilon) | (W\psi)(\epsilon') \rangle d\epsilon d\epsilon' dt \\ &= \frac{1}{2\pi} \int_{\tau} \int_{\sigma(H)} \int_{\sigma(H)} e^{i(\epsilon' - \epsilon)t} \langle W(\epsilon)\varphi(\epsilon) | W(\epsilon')\psi(\epsilon') \rangle d\epsilon d\epsilon' dt, \end{aligned}$$

where $\varphi, \psi \in \mathcal{H}$ are absolutely integrable and $\tau \in \mathcal{B}(\mathbb{R})$. The rest of the proof is left as an exercise. (For the final claim we note that if $\mathbf{T}$ is projection valued then $W : \mathcal{H} \rightarrow \mathcal{K}'$ is a unitary operator [14].) $\square$

From Theorem 17.1 one sees that, for given H (with the absolutely continuous spectrum), there actually exists an infinite number of covariant time observables; it is an easy exercise to show that any family of isometries $W(\epsilon)$ defines a covariant time observable. One can study their optimality (e.g. extremality) properties and try to find the most optimal time observables. For example, one can fully characterise covariant time observables which are extremals in the convex set of covariant time observables and even in the set of all observables, see [15].

It is interesting to note that if the spectrum of H is bounded from below then there are no sharp time observables or time operators in the sense of the spectral theorem. This happens, e.g., when H is the Hamiltonian $H = (2m)^{-1} P + V(Q)$ of a particle subjected to a compactly supported bounded positive potential $V : \mathbb{R} \rightarrow \mathbb{R}$. However, if the potential V does not have these properties, then there may exist sharp time observables. The next example indicates such a case.

Example 17.4 In this example, we study a quantum analog of the time for a freely falling particle as in Example 17.1. Let $\mathcal{H} = L^2(\mathbb{R})$ be the Hilbert space of a particle

(of mass m) falling along a vertical q -axis, and let $\mathbf{Q}$ and $\mathbf{P}$ be the usual position and momentum observables so that

$$H = \frac{1}{2m} P^2 + mgQ$$

and $\sigma(H) = \sigma_c(H) = \mathbb{R}$, $d\mu_H(\epsilon) = d\epsilon$. It is an easy exercise to show that $\mathbf{T}_{\text{can}}(\tau) = \mathbf{P}(-mg\tau)$ is the canonical covariant time observable with respect to H . The rank of $\mathbf{T}_{\text{can}}$ is 1. Moreover, the time observable $\mathbf{T}_{\text{can}}$ is a spectral measure (and thus extremal) with the time operator $T_{\text{can}} = \mathbf{T}_{\text{can}}[1] = P/(-mg)$ which corresponds to the classical time $T(q, p) = p/(-mg)$. Clearly, $[T_{\text{can}}, H] = iI$ in a dense domain. It is another easy exercise to show that the pair (T_{can}, H) is indeed a Schrödinger pair so that it is unitarily equivalent to the canonical pair (Q, P) . Especially, the time-energy pair $(\mathbf{T}_{\text{can}}, H)$ share the preparation and measurement uncertainty relations of the canonical position and momentum observables studied in Chap. 15.

Suppose then that the initial (vector) state of the particle at $t = 0$ is the minimum uncertainty state $\varphi_{b,0}(q) = (b/\pi)^{1/4} e^{-\frac{1}{2}b(q-\bar{q})^2}$ in the position representation, where $b > 0$ and $\bar{q} \in \mathbb{R}$ is the distance parameter, see Eq. (15.8). Especially, $\langle \varphi_{b,0} | Q \varphi_{b,0} \rangle = \bar{q}$ and $\langle \varphi_{b,0} | P \varphi_{b,0} \rangle = 0$ so that the particle ‘falls from rest’ and $\bar{q}$ is the ‘initial position’ of the particle. At laboratory time t , the state is $\varphi'_{b,0} = e^{-itH} \varphi_{b,0}$ and the equations $\langle \varphi'_{b,0} | Q \varphi'_{b,0} \rangle = \bar{q} - \frac{1}{2}gt^2$, $\langle \varphi'_{b,0} | P \varphi'_{b,0} \rangle = -mgt$, $\langle \varphi'_{b,0} | H \varphi'_{b,0} \rangle = mg\bar{q}$, and $\langle \varphi'_{b,0} | T_{\text{can}} \varphi'_{b,0} \rangle = t$ correspond to the classical equation (17.2). In the momentum representation, $\hat{\varphi}_{b,0}(p) = (b\pi)^{-1/4} e^{-i\bar{q}p} e^{-p^2/(2b)}$ so that the probability that the time measurement yields a value smaller than t' is

$$\begin{aligned} \langle e^{-it'H} \varphi_{b,0} | \mathbf{T}_{\text{can}}((-\infty, t']) e^{-itH} \varphi_{b,0} \rangle &= \langle \varphi_{b,0} | \mathbf{T}_{\text{can}}((-\infty, t' - t]) \varphi_{b,0} \rangle \\ &= \frac{1}{\sqrt{b\pi}} \int_{-mg(t'-t)}^{\infty} e^{-p^2/b} dp = \frac{1}{2} [1 + \text{erf}(b^{-1/2}mg(t' - t))] \end{aligned}$$

which is $\frac{1}{2}$ when $t' = t$, almost zero when $t' < t$, and almost 1 when $t' > t$ (if $b \approx 0$). Moreover, the probability does not depend on $\bar{q}$. To conclude, $\mathbf{T}_{\text{can}}$ can be used as a quantum clock, realised by a freely falling particle. $\triangleleft$

Example 17.5 As the second example we consider a free particle moving in one dimension. Then $H = (2m)^{-1}P^2$ with the (absolutely continuous) spectrum $\sigma(H) = \{\epsilon \in \mathbb{R} | \epsilon \geq 0\}$, and we assume that $\mathcal{H}$ is the momentum representation space, that is, $(\mathbf{P}(X)\psi)(p) = \chi_X(p)\psi(p)$. Obviously, p and $-p$ give the same generalised eigenvalue $\epsilon = p^2/(2m)$ so that the multiplicities $d(\epsilon) = 2$ and the energy representation space is $L^2_{\mathbb{C}^2}([0, \infty)) \cong \mathcal{H}$, where the isomorphism is given by

$$\mathcal{H} \ni \psi \mapsto \left[\epsilon \mapsto \sqrt{\frac{m}{2\epsilon}} \left(\psi(\sqrt{2m\epsilon})e_1 + \psi(-\sqrt{2m\epsilon})e_2 \right) \right] \in L^2_{\mathbb{C}^2}([0, \infty))$$

where $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

From Theorem 17.1 one sees that any covariant time observable T is determined by isometries $W(\epsilon) : \mathbb{C}^2 \rightarrow \hat{\mathcal{H}}$ where $\hat{\mathcal{H}}$ is an infinite dimensional Hilbert space. Denote $\eta_i(\epsilon) = W(\epsilon)e_i \in \hat{\mathcal{H}}, i = 1, 2$, so that the isometry condition $W(\epsilon)^*W(\epsilon) = I$ equals $\langle \eta_i(\epsilon) | \eta_j(\epsilon) \rangle = \delta_{ij}$ and T is determined by a measurable family of orthonormal vectors $\{\eta_1(\epsilon), \eta_2(\epsilon)\}_{\epsilon \geq 0}$. We note that there is no projection valued time observable.

When one chooses $\eta_i(\epsilon) = \psi_i, i = 1, 2, \langle \psi_i | \psi_j \rangle = \delta_{ij}$, one obtains the canonical time observable [16, p. 133]:

$$\langle \varphi | T_{\text{can}}(\tau) \psi \rangle = \frac{1}{2\pi} \int_{\tau} \int_0^{\infty} \int_0^{\infty} e^{i(\epsilon' - \epsilon)t} \langle \varphi(\epsilon) | \psi(\epsilon') \rangle d\epsilon d\epsilon' dt$$

for all $\varphi, \psi \in L^1_{\mathbb{C}^2}([0, \infty)) \cap L^2_{\mathbb{C}^2}([0, \infty))$ and $\tau \in \mathcal{B}(\mathbb{R})$. It resembles the canonical phase observable introduced in Chap. 16 except that the spectrum of the shift generator is $[0, \infty)$ instead of $\mathbb{N}_0$ and the value space is $\mathbb{R}$ instead of $\mathbb{T}$. Hence, the canonical time has essentially the same properties as the canonical phase. For example, it can be shown that the canonical time is extremal in the convex set of all observables, see [15].

The first moment operator of T_{can} is symmetric and densely defined and it coincides with the Aharonov–Bohm time operator

$$T_{AB} = \frac{m}{2}(QP^{-1} + P^{-1}Q)$$

on a dense subspace of $L^2(\mathbb{R})$. It corresponds to the classical (local) time $T(q, p) = mq/p$ of Example 17.2 [16]. $\triangleleft$

Example 17.6 In this example, we consider a time-energy uncertainty relation of a covariant time observable and a Hamiltonian H with a simple, bounded, absolutely continuous spectrum $[0, 2\pi]$. One can think of a particle moving in one spatial dimension, with its momentum confined to the interval $[0, p_0]$, where $p_0^2/2m = 2\pi$.

Let $\mathcal{H}$ be the Hilbert space $L^2(0, 2\pi) = L^2([0, 2\pi])$ in which H acts as the multiplication operator $(H\psi)(\epsilon) = \epsilon\psi(\epsilon), \epsilon \in [0, 2\pi]$. We choose a shift-covariant family of unit vectors $\varphi_t, t \in \mathbb{R}$, as follows: $\varphi_t(\epsilon) = (e^{-itH}\varphi_0)(\epsilon) = e^{-i\epsilon t}/\sqrt{2\pi}$. We can then define a (rank-1) time-shift covariant observable via

$$T(\tau) = \int_{\tau} |\varphi_t\rangle\langle\varphi_t| dt = \int_{\tau} e^{-itH} |\varphi_0\rangle\langle\varphi_0| e^{itH} dt \quad (17.3)$$

for all $\tau \in \mathcal{B}(\mathbb{R})$, and

$$\langle \varphi | T(\tau) \psi \rangle = \frac{1}{2\pi} \int_{\tau} \int_0^{2\pi} \int_0^{2\pi} e^{i(\epsilon' - \epsilon)t} \overline{\varphi(\epsilon)} \psi(\epsilon') d\epsilon d\epsilon' dt.$$

(Note that the isometries $W(\epsilon) = 1$ of Theorem 17.1 are now trivial.)

A *formal* time operator is obtained from the first moment operator of $\mathbf{T}$:

$$(T\psi)(\epsilon) = \int_{\mathbb{R}} t \varphi_t(\epsilon) \langle \varphi_t | \psi \rangle dt = i \frac{d}{d\epsilon} \psi(\epsilon); \quad (17.4)$$

this is well-defined for functions $\psi \in L^2(0, 2\pi)$ which are absolutely continuous and such that the derivative $\psi' \in L^2(0, 2\pi)$. In order for this operator to be symmetric, the domain must be further restricted by appropriate boundary conditions. It is well known that the condition $\psi(2\pi) = c\psi(0)$ makes $id/d\epsilon$ a selfadjoint operator $T^{(c)}$ for any c of modulus 1 [17]. Each such $T^{(c)}$ is a selfadjoint extension of the differential operator understood as a symmetric operator $T^{(0)}$ with the boundary condition $\psi(0) = \psi(2\pi) = 0$. Note that the spectrum of $T^{(c)}$ is $\mathbb{Z} + \arg(c)/(2\pi)$, with eigenvectors $e^{-i \arg(c)H/(2\pi)} \varphi_m$, where $\varphi_m(\epsilon) = e^{-im\epsilon}/\sqrt{2\pi}$, $m \in \mathbb{Z}$. Moreover, one can show that $\mathbf{T}[1] = T_0$ as a usual *strong* operator integral and also as a more general *weak* operator integral [18, Appendix A].

The covariance relation

$$e^{-itH} T e^{itH} = T - tI$$

is found to be satisfied for $T^{(0)}$ but not for any of its selfadjoint extensions since $e^{-itH} T^{(c)} e^{itH} = T^{(c')}$ with $c' = e^{i2\pi t} c$. In accordance with this, the canonical commutation relation between the Hamiltonian and the time operator is obtained only on the domain ³ of $T^{(0)}$, and therefore the preparation uncertainty relation

$$\Delta(\mathbf{T}_\varrho) \Delta(\mathbf{H}_\varrho) \geq \frac{\hbar}{2} \quad (17.5)$$

holds only on this dense subspace (here we have put $\hbar$ back by replacing H with $\hbar H$). Since the spectrum of H is a bounded interval of length $\lambda(H) = 2\pi$, there is an absolute bound to the temporal variance in any such state ϱ :

$$\Delta(\mathbf{T}_\varrho) \geq \frac{\hbar}{2\lambda(H)}. \quad (17.6)$$

To conclude this example, we compare briefly the time observable (17.3) with the general construction of Brunetti and Fredenhagen in [19]. Given a bounded positive operator A , consider the positive operator measure, defined first on intervals J via

$$\mathbf{B}(J) = \int_J e^{-itH} A e^{itH} dt.$$

³This is the set of states ϱ whose eigenvectors belong to the domain of $T^{(0)}$.

It can be shown [19] that in certain circumstances this can be turned into a normalised semispectral measure on a suitable closed subspace (provided this is not the null space). Indeed, it is easy to see that $\mathbf{B}$ is a normalised semispectral measure if A is a positive operator with the unit trace. In the present case of $\mathbf{T}$ as defined in (17.3), we see that the operator corresponding to A can be identified with the one-dimensional projection operator $|\varphi_0\rangle\langle\varphi_0|$. The observable $\mathbf{T}$ corresponds to a measurement of the time that the system spends (loosely speaking) in the state φ_0 . $\triangleleft$

Example 17.7 Perhaps the most important application up to now of arrival time observables is the analysis of [11] which resolves the question of a possible ‘superluminal signal transport through a tunnel’: transmission through a tunnel will always slow down the signal. The arrival time probabilities of a free particle, $H_0 = \frac{1}{2m}P^2$, are compared with those of the particle subjected to a tunnelling potential $V(Q)$ (a compactly supported bounded positive operator), $H = H_0 + V(Q)$, under the requirements that the initial vector states of the two systems are asymptotically (at $t \rightarrow -\infty$) the same and that the two arrival time observables are indistinguishable after the scattering event, that is, at large times $t \rightarrow \infty$. To formulate these requirements, let $\Omega_{\pm} = \text{s-lim}_{t \rightarrow \pm\infty} e^{itH} e^{-itH_0}$ be the (unitary) wave operators. Then, for any vector state ψ_0 of the free particle (with positive momenta), the corresponding state of the particle subjected to the tunnelling potential must be $\psi = \Omega_- \psi_0$. If $\mathbf{T}_0$ is the arrival time observable of the free particle, then the second requirement implies that the arrival time observable $\mathbf{T}$ of the other particle is to be unitarily equivalent to $\mathbf{T}_0$ under the outgoing wave operator Ω_+ , that is, $\mathbf{T}(\cdot) = \Omega_+ \mathbf{T}_0(\cdot) \Omega_+^*$. Clearly, such a $\mathbf{T}$ is covariant under the time shifts generated by H (or $-H$) if $\mathbf{T}_0$ is covariant under the time shifts generated by H_0 . In [11] it is then shown that for any (vector) state ψ_0 of the free particle (with positive momenta), the corresponding arrival time distributions $t \mapsto \langle \psi_0 | \mathbf{T}_0((-\infty, t]) \psi_0 \rangle$ and $t \mapsto \langle \psi | \mathbf{T}((-\infty, t]) \psi \rangle$, with $\psi = \Omega_- \psi_0$ and $\mathbf{T}(\cdot) = \Omega_+ \mathbf{T}_0(\cdot) \Omega_+^*$, satisfy the relation

$$\langle \psi | \mathbf{T}((-\infty, t]) \psi \rangle \leq \langle \psi_0 | \mathbf{T}_0((-\infty, t]) \psi_0 \rangle$$

for any t , showing that the probability of arrival by the time t is never larger for the tunnelled particle. $\triangleleft$

Periodic Systems

In the periodic case, there exists a minimal period $T_0 > 0$ such that $e^{iT_0 H} = I$, implying $e^{inT_0 H} = I$ for all $n \in \mathbb{Z}$. Hence, we choose T_0 to be the smallest positive number satisfying the above condition. Now the spectrum $\sigma(H)$ must be discrete. Indeed, $(e^{iT_0 H} \psi)(\epsilon) = e^{iT_0 \epsilon} \psi(\epsilon) = \psi(\epsilon)$ implies that $T_0 \epsilon = 2\pi n$, $n \in \mathbb{Z}$, or by denoting $\omega = 2\pi/T_0$, $\epsilon = \omega n$, $n \in \mathbb{Z}$. Hence, $\sigma(H) = \omega \mathbb{Z}_H$ where $\mathbb{Z}_H \subset \mathbb{Z}$. Denote $d_n = d(\epsilon = \omega n)$ and let $\psi_{n,k} \in \mathcal{H}$, $k < d_n + 1$, be the orthonormal eigenvectors associated with the eigenvalue ωn , $n \in \mathbb{Z}_H$. Hence, $\{\psi_{n,k} | n \in \mathbb{Z}_H, k < d_n + 1\}$ is an orthonormal basis of $\mathcal{H}$ and we may write $\mathcal{H} = \bigoplus_{n \in \mathbb{Z}_H} \mathcal{H}_n$ where $\mathcal{H}_n$ is the eigenspace associated with ωn and hence spanned by vectors $\psi_{n,k}$, $k < d_n + 1$. Thus, any $\psi \in \mathcal{H}$ can uniquely be written in the form $\psi = \sum_{n \in \mathbb{Z}_H} \psi_n$ where $\psi_n \in \mathcal{H}_n$.

We then define an angle parameter $\theta = \omega t$ and a generalised ‘number’ operator $N_H = \omega^{-1}H$ so that

$$N_H \psi_{n,k} = n \psi_{n,k} \text{ and } U_t = e^{itH} = e^{i\theta N_H}.$$

We say that an observable $\mathsf{T} : \mathcal{B}([0, T_0)) \rightarrow \mathcal{L}(\mathcal{H})$ is a *covariant (periodic) time observable* if it is time shift covariant:

$$U_t^* \mathsf{T}(\tau) U_t = \mathsf{T}(\tau \dot{+} t)$$

for all $t \in [0, T_0)$ and $\tau \in \mathcal{B}([0, T_0))$ (where $\dot{+}$ is addition modulo period). From the above discussion we see that, without restricting generality, we may always assume that $T_0 = 2\pi$ so that $\omega = 1$, $t = \theta$, $\sigma(H) = \mathbb{Z}_H$, and $H = N_H$. Thus, the characterisation of covariant periodic time observables is (essentially) the same as that of covariant phase observables associated with a generalised number operator (with possibly degenerate eigenvalues). The proof of the following theorem is left as an exercise:

Theorem 17.2 *An observable $\mathsf{T} : \mathcal{B}([0, T_0)) \rightarrow \mathcal{L}(\mathcal{H})$ is a covariant time observable if and only if there exist isometries $W_n : \mathcal{H}_n \rightarrow \mathcal{H}$ such that*

$$\langle \varphi | \mathsf{T}(\tau) \psi \rangle = \sum_{n,m \in \mathbb{Z}_H} \frac{\omega}{2\pi} \int_{\tau} e^{i(m-n)\omega t} dt \langle W_n \varphi_n | W_m \psi_m \rangle$$

for all $\varphi, \psi \in \mathcal{H}$ and $\tau \in \mathcal{B}(\mathbb{R})$. Moreover, T is a spectral measure if and only if $\mathbb{Z}_H = \mathbb{Z}$ and, for all $n \in \mathbb{Z}$, $d_n = c < \dim \mathcal{H} + 1$, and W_n is a unitary operator from $\mathcal{H}_c$ onto a c -dimensional subspace of $\mathcal{H}$ (which can be chosen to be $\mathcal{H}_c$).

Note that now relying on the Heisenberg commutation relation $[T, H] = iI$ (where $T = \int_0^{T_0} t d\mathsf{T}(t)$) to characterise time observables is even more problematic than in the nonperiodic case, since one may have cases where $\mathcal{H}$ is finite-dimensional so that $\langle \psi | [T, H] \psi \rangle = 0$ for any eigenvector ψ of H . However, the Weyl commutator relation can be defined as before.

Example 17.8 Consider the quantum harmonic oscillator (recall Example 17.3). One has

$$H = \frac{1}{2m} P^2 + \frac{k}{2} Q^2 = \hbar\omega(N + \frac{1}{2}I)$$

where $k > 0$, $\omega = \sqrt{k/m}$, and $N = \sum_{n=0}^{\infty} n |n\rangle \langle n|$ the usual number operator. Hence, the covariant time observables T are in a bijective correspondence with the covariant phase observables Φ via

$$\mathsf{T}(\tau) = \overline{\Phi(\omega\tau)}$$

where $\mathcal{L}(\mathcal{H}) \ni A \mapsto \overline{A} = \sum_{n,m=0}^{\infty} \overline{\langle n|A|m\rangle} |n\rangle\langle m| \in \mathcal{L}(\mathcal{H})$ is the conjugation map (with respect to the number basis). Hence, there exists a canonical time observable for the harmonic oscillator which is just the modified canonical phase:

$$T_{\text{can}}(\tau) = \overline{\Phi_{\text{can}}(\omega\tau)} = \sum_{n,m=0}^{\infty} \frac{\omega}{2\pi} \int_{\tau} e^{i(m-n)\omega t} dt |n\rangle\langle m|$$

for all $\tau \in \mathcal{B}([0, T_0])$.

◁

17.4 Exercises

1. Complete the proof of Theorem 17.1.
2. Work out the details of Example 17.4. For instance, by using the Baker-Campbell-Hausdorff formula, show that T_{can} is time-shift covariant. What are the other solutions of the covariance system according to Theorem 17.1? Show that if a covariant time observable is rank-1 then it is unitarily equivalent with T_{can} . If T is a convolution of T_{can} what can be said about the isometries $W(\epsilon)$? Examine the uncertainty relations of the pair (T_{can}, H) and compare the situation with the case of a free particle.
3. Compare Examples 17.1 and 17.4. What technical problems arise if you try to define a quantum arrival time operator as $T_{\text{arr}} = (mg)^{-1}(P + \sqrt{2mH})$?
4. In the context of Example 17.6, show that, for any $m \in \mathbb{Z}$, $\Delta(T_{\varrho}) = \infty$ where $\varrho = P[\varphi_m]$.
5. Prove Theorem 17.2.

References

1. Mackey, G.W.: Unitary Group Representations in Physics, Probability, and Number Theory. Advanced Book Classics, 2nd edn. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA (1989)
2. Varadarajan, V.S.: Geometry of Quantum Theory, 2nd edn. Springer, New York (1985)
3. Cassinelli, G., De Vito, E., Lahti, P., Levrero, A.: The Theory of Symmetry Actions in Quantum Mechanics - With an application to the Galilei Group. Lecture Notes in Physics, vol. 654. Springer, Berlin (2004)
4. Pauli, W.: General Principles of Quantum Mechanics. Springer, Berlin (1980). Translated from the German original of 1933 by P. Achuthan and K. Venkatesan
5. Muga, J.G., Mayato, R.S., Egusquiza, Í.L. (eds.): Time in Quantum Mechanics vol. 1. Volume 734 of Lecture Notes in Physics, 2nd edn. Springer, Berlin (2008)
6. Muga, J.G., Ruschhaupt, A., del Campo, A. (eds.): Time in Quantum Mechanics vol. 2. Volume 789 of Lecture Notes in Physics. Springer, Berlin (2009)
7. Thirring, W.: Quantum Mathematical Physics. Atoms, Molecules and Large Systems, 2nd edn. Springer, Berlin (2002). Translated from the 1979 and 1980 German originals by Evans M. Harrell II

8. Kiukas, J., Ruschhaupt, A., Schmidt, P.O., Werner, R.F.: Exact energy-time uncertainty relation for arrival time by absorption. *J. Phys. A* **45**(18), 185301 (2012). 13
9. Busch, P.: The time-energy uncertainty relation. In: Muga, J., Mayato, R., Egusquiza, A. (eds.) *Time in Quantum Mechanics. Lecture Notes in Physics*, vol. 734, pp. 73–105. Springer, Berlin Heidelberg (2007)
10. Werner, R.: Screen observables in relativistic and nonrelativistic quantum mechanics. *J. Math. Phys.* **27**(3), 793–803 (1986)
11. Kiukas, J., Ruschhaupt, A., Werner, R.F.: Tunneling times with covariant measurements. *Found. Phys.* **39**(7), 829–846 (2009)
12. Kholevo, A.S.: Generalized imprimitivity systems for abelian groups. *Soviet Math. (Iz. VUZ)* **27**(2), 53–80 (1983)
13. Dixmier, J.: *Von Neumann Algebras Volume 27 North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam (1981). Translated from the second French edition by F. Jellet
14. Cattaneo, U.: On Mackey's imprimitivity theorem. *Comment. Math. Helv.* **54**(4), 629–641 (1979)
15. Haapasalo, E., Pellonpää, J.-P.: Extreme covariant quantum observables in the case of an Abelian symmetry group and a transitive value space. *J. Math. Phys.* **52**(12), 122102, 23 (2011)
16. Holevo, A.S.: *Probabilistic and Statistical Aspects of Quantum Theory Volume 1 North-Holland Series in Statistics and Probability*. North-Holland Publishing Co., Amsterdam (1982). Translated from the Russian by the author
17. Akhiezer, N.I., Glazman, I.M.: *Theory of Linear Operators in Hilbert Space*. Dover Publications Inc., New York (1993). Translated from the Russian and with a preface by Merlynd Nestell, Reprint of the 1961 and 1963 translations, Two volumes bound as one
18. Dubin, D.A., Kiukas, J., Pellonpää, J.-P., Ylinen, K.: Operator integrals and sesquilinear forms. *J. Math. Anal. Appl.* **413**(1), 250–268 (2014)
19. Brunetti, R., Fredenhagen, K.: Time of occurrence observable in quantum mechanics. *Phys. Rev. A* (3) **66**(4), 044101 (2002)

Chapter 18

State Reconstruction

If a measurement of an observable is performed and a result is registered, then the system must have been in a state in which the result was possible. Since there is no effect operator $A \in \mathbf{E}(\mathcal{H})$ for which the map

$$S(\mathcal{H}) \ni \varrho \mapsto \text{tr}[\varrho A] \in [0, 1]$$

is injective, no single measurement result can ever—without prior knowledge on the state—determine the state of the system.

Consider then a subset $\mathcal{E} \subset \mathbf{E}(\mathcal{H})$ of effects, for instance the range of an observable $\mathbf{E}$, $\mathcal{E} = \text{ran}(\mathbf{E}) = \{\mathbf{E}(X) \mid X \in \mathcal{A}\}$, or a union of such ranges, and consider the assignment

$$S(\mathcal{H}) \ni \varrho \mapsto \{\text{tr}[\varrho A] \mid A \in \mathcal{E}\} \subset [0, 1].$$

If the set $\mathcal{E}$ is sufficiently large, then the associated measurement outcome probabilities separate states, that is, for any two states ϱ_1 and ϱ_2 , $\varrho_1 = \varrho_2$ if and only if $\text{tr}[\varrho_1 A] = \text{tr}[\varrho_2 A]$ for all $A \in \mathcal{E}$. Such a set $\mathcal{E}$ is called *informationally complete*. If $\mathcal{E}$ is informationally complete, one may try to *reconstruct the state* ϱ from the statistics $\{\text{tr}[\varrho A] \mid A \in \mathcal{E}\}$.

This chapter is devoted to a study of informational completeness and the related problem of state reconstruction. Special attention is given to the continuous variable case. The key concepts and the basic results will be introduced in Sect. 18.1, including a short discussion of the qubit case; the Pauli problem and the two basic ways of overcoming it are studied in Sect. 18.2. The final Sect. 18.3 is devoted to the problem of state reconstruction on the basis of informational completeness.

18.1 Informational Completeness

Let $\mathcal{E}$ be a subset of the set $\mathbf{E}(\mathcal{H})$ of effects. In the physical applications considered here the set $\mathcal{E}$ appears as the range $\text{ran}(\mathbf{E})$ of a single observable $\mathbf{E}$ or as the union of the ranges $\bigcup_{\mathbf{E} \in \mathcal{M}} \text{ran}(\mathbf{E})$ of a set $\mathcal{M}$ of observables. We say that two states ϱ_1 and ϱ_2 are $\mathcal{E}$ -equivalent if

$$\text{tr}[\varrho_1 A] = \text{tr}[\varrho_2 A] \quad \text{for all } A \in \mathcal{E}.$$

If ϱ_1 and ϱ_2 are $\mathcal{E}$ -equivalent states, we denote it as $\varrho_1 \sim_{\mathcal{E}} \varrho_2$. For any $\varrho \in \mathbf{S}(\mathcal{H})$ we let $[\varrho]^{\mathcal{E}}$ denote the equivalence class

$$[\varrho]^{\mathcal{E}} = \{\varrho' \in \mathbf{S}(\mathcal{H}) \mid \varrho' \sim_{\mathcal{E}} \varrho\}.$$

We use also the notions $[\varrho]^{\mathbf{E}}$, $[\varrho]^{(\mathbf{E}_1, \dots, \mathbf{E}_n)}$, or $[\varrho]^{\mathcal{M}}$ where $\mathcal{E}$ is the union of the ranges of the observables involved.

Definition 18.1 Two subsets of effects $\mathcal{E}_1, \mathcal{E}_2 \subset \mathbf{E}(\mathcal{H})$ are *informationally equivalent with respect to a set of states* $\mathcal{S} \subset \mathbf{S}(\mathcal{H})$ if $[\varrho]^{\mathcal{E}_1} = [\varrho]^{\mathcal{E}_2}$ for all $\varrho \in \mathcal{S}$. If $\mathcal{E}_1$ and $\mathcal{E}_2$ are informationally equivalent with respect to all states $\varrho \in \mathbf{S}(\mathcal{H})$, then $\mathcal{E}_1$ and $\mathcal{E}_2$ are *informationally equivalent*. Two sets $\mathcal{M}_1$ and $\mathcal{M}_2$ of observables are *informationally equivalent with respect to a set of states* $\mathcal{S} \subset \mathbf{S}(\mathcal{H})$ if $[\varrho]^{\mathcal{M}_1} = [\varrho]^{\mathcal{M}_2}$ for all $\varrho \in \mathcal{S}$, and they are *informationally equivalent* if they are informationally equivalent with respect to all states.

Definition 18.2 A subset $\mathcal{E}$ of effects is *informationally complete with respect to a set of states* $\mathcal{S} \subset \mathbf{S}(\mathcal{H})$ if $[\varrho]^{\mathcal{E}} = \{\varrho\}$ for all $\varrho \in \mathcal{S}$. The set $\mathcal{E}$ is *informationally complete* if it is informationally complete with respect to the whole set of states. A set $\mathcal{M}$ of observables is *informationally complete with respect to a set of states* $\mathcal{S}$ if $[\varrho]^{\mathcal{M}} = \{\varrho\}$ for all $\varrho \in \mathcal{S}$, and $\mathcal{M}$ is *informationally complete* if it is informationally complete with respect to the whole set of states.

The state distinction power of (sets of) informationally equivalent observables are the same. They separate exactly the same states. The informational completeness of a set $\mathcal{M}$ of observables means that the totality of the measurement outcome probabilities $\mathbf{E}_{\varrho}$, $\mathbf{E} \in \mathcal{M}$, determines the state ϱ of the system. Such (sets of) observables form the basis of quantum tomography.

The following proposition restates the notion of informational completeness in some alternative forms. We leave the proof as an exercise.

Proposition 18.1 *For any $\mathcal{E} \subset \mathbf{E}(\mathcal{H})$ the following conditions are equivalent:*

- (i) $\mathcal{E}$ is informationally complete;
- (ii) for any $T \in \mathcal{T}_s(\mathcal{H})$, if $\text{tr}[TA] = 0$ for all $A \in \mathcal{E}$, then $T = 0$;
- (iii) for any $T \in \mathcal{T}(\mathcal{H})$, if $\text{tr}[TA] = 0$ for all $A \in \mathcal{E}$, then $T = 0$;

- (iv) the ultraweak closure of the linear span of $\mathcal{E}$ is $\mathcal{L}(\mathcal{H})$;
- (v) the map

$$\mathcal{T}_s(\mathcal{H}) \ni T \mapsto \|T\|_{\mathcal{E}} = \sup \left\{ \frac{|\operatorname{tr}[TA]|}{\|A\|} \mid A \in \mathcal{E}, A \neq 0 \right\} \in [0, \infty)$$

is a norm (with the additional requirement that $I \in \mathcal{E}$).

Example 18.1 In the two dimensional Hilbert space $\mathbb{C}^2$ a state ϱ is uniquely determined by a vector $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{R}^3$ of length $\|\mathbf{n}\| \leq 1$,

$$\varrho = \varrho_{\mathbf{n}} = \frac{1}{2}(I + n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3) = \frac{1}{2}(I + \mathbf{n} \cdot \boldsymbol{\sigma}).$$

For any spin quantity $S_{\mathbf{a}} = \frac{1}{2} \mathbf{a} \cdot \boldsymbol{\sigma}$, $\mathbf{a} \in \mathbb{R}^3$, $\|\mathbf{a}\| = 1$, the measurement outcome probabilities are $E_{\varrho_{\mathbf{n}}}^{S_{\mathbf{a}}}(\pm) = \frac{1}{2}(1 \pm \mathbf{n} \cdot \mathbf{a})$. Choosing any three linearly independent directions $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ the associated probabilities determine the direction $\mathbf{n}$ and allow thus to reconstruct the state $\varrho_{\mathbf{n}}$. In other words, any three linearly independent spin quantities ($S_{\mathbf{a}}$, $S_{\mathbf{b}}$, $S_{\mathbf{c}}$) form an informationally complete set of observables and the state is directly obtained from the probabilities. Clearly, in this case informational completeness with respect to pure states already guarantees informational completeness. $\triangleleft$

Noncommutativity, even total noncommutativity is an essential feature of informational completeness. We note that $\mathcal{E}$ is totally noncommutative if its commutativity domain $\operatorname{com}(\mathcal{E}) = \{\varphi \in \mathcal{H} \mid AB\varphi = BA\varphi \text{ for all } A, B \in \mathcal{E}\}$ contains only the null vector. The commutativity domain of $\mathcal{E}$ equals the commutativity domain of the ultraweak closure of the linear span of $\mathcal{E}$. Hence, by Proposition 18.1 (iv), $\operatorname{com}(\mathcal{E}) = \operatorname{com}(\mathcal{L}(\mathcal{H})) = \{0\}$. We summarise this basic observation as a corollary.

Corollary 18.1 *If $\mathcal{E} \subset \mathbf{E}(\mathcal{H})$ is informationally complete, then $\mathcal{E}$ is totally noncommutative, that is, $\operatorname{com}(\mathcal{E}) = \{0\}$.*

The total noncommutativity of $\mathcal{E}$ is necessary but not sufficient for $\mathcal{E}$ to be informationally complete. For example, the canonical phase Φ_{can} is totally noncommutative but informationally incomplete (it gives the random phase distribution in all number states).

Observables given as spectral measures are of special importance. Since no single spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ is informationally complete it is useful to know what is the largest set of states $\mathcal{S} \subset \mathcal{S}(\mathcal{H})$ with respect to which $\mathbf{E}$ is informationally complete. The following simple proposition answers this question. Let $\mathcal{S}(\mathbf{E})$ denote the set of one-dimensional spectral projections of $\mathbf{E}$; we note that $\mathcal{S}(\mathbf{E})$ may be empty.

Proposition 18.2 *For any spectral measure $\mathbf{E} : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$, $[\varrho]^{\mathbf{E}} = \{\varrho\}$ if and only if $\varrho \in \mathcal{S}(\mathbf{E})$.*

18.2 The Pauli Problem

In a footnote on p. 17 to his 1933 *Wellenmechanik*, Wolfgang Pauli remarked that the question under which conditions the position and momentum distributions $|\varphi|^2$ and $|\hat{\varphi}|^2$ define the state function φ uniquely (modulo a phase factor) ‘has still not been investigated in all its generality’ [1]. This question is known as the *Pauli problem*. One of the first examples demonstrating the informational incompleteness of the pair (Q, P) , reported in [2], is due to V. Bargmann. The papers of Prugovečki [3] and Corbett and Hurst [4] contain further examples.

This subsection starts with a demonstration of the informational incompleteness of the position-momentum pair: there are vector states φ which are not determined by the distributions $|\varphi|^2$ and $|\hat{\varphi}|^2$. Following this, a completion of the pair (Q, P) in terms of the rotated quadrature observables is presented. The next subsection investigates another completion of this pair in terms of informationally equivalent jointly measurable pairs $(\mu * Q, \nu * P)$.

Consider the position and momentum Q and P of a quantum object with the concrete Hilbert space $\mathcal{H} = L^2(\mathbb{R})$. As shown in Sect. 15.1, position and momentum are totally noncommutative, $\text{com}(Q, P) = \{0\}$. However, in general, the distributions Q_ϱ and P_ϱ do not determine the state ϱ . As an example, consider the function $\varphi_{a,b}(q) = \left(\frac{2a}{\pi}\right)^{1/4} e^{-(a+ib)q^2}$, with $a, b \in \mathbb{R}, a > 0$. Computing its Fourier transform,

$$\hat{\varphi}_{a,b}(p) = \left[\frac{a}{2\pi(a^2 + b^2)} \right]^{1/4} \exp\left(-\frac{ap^2}{4(a^2 + b^2)}\right) \exp\left(\frac{ibp^2}{4(a^2 + b^2)} - \frac{i}{2} \arctan \frac{b}{a}\right),$$

one observes that the position and momentum distributions in this state are

$$|\varphi_{a,b}(q)|^2 = \left(\frac{2a}{\pi}\right)^{1/2} e^{-2aq^2}, \quad |\hat{\varphi}_{a,b}(p)|^2 = \left(\frac{a}{2\pi(a^2 + b^2)}\right)^{1/2} e^{-\frac{ap^2}{2(a^2 + b^2)}}.$$

For $b \neq 0$, the states $\varrho_1 = |\varphi_{a,b}\rangle\langle\varphi_{a,b}|$ and $\varrho_2 = |\varphi_{a,-b}\rangle\langle\varphi_{a,-b}|$ are different, but $Q_{\varrho_1} = Q_{\varrho_2}$ and $P_{\varrho_1} = P_{\varrho_2}$.

This result opens two lines of investigations: find a maximal set of states $\mathcal{S} \subset \mathcal{S}(\mathcal{H})$ with respect to which the pair (Q, P) is informationally complete; find a minimal set $\mathcal{M}$ of observables which completes the pair (Q, P) informationally. To the best of our knowledge, neither of these lines have been fully explored.

Rotated Quadratures

Let $N = a^*a$ be the selfadjoint number operator defined by the lowering and raising operators $a = \frac{1}{\sqrt{2}}(Q + iP)$ and $a^* = \frac{1}{\sqrt{2}}(Q - iP)$ associated with the Hermite functions $|n\rangle = h_n$. Let $U_\theta = e^{i\theta N}$, $\theta \in [0, 2\pi)$, and define the rotated quadrature operators $Q_\theta = U_\theta Q U_\theta^*$, with the spectral measures $Q_\theta = U_\theta Q U_\theta^*$. In particular, $Q_0 = Q$ and $Q_{\pi/2} = P$.

Theorem 18.1 *Let $J \subset [0, \pi)$. The set $\{\mathbf{Q}_\theta \mid \theta \in J\}$ of rotated quadrature observables is informationally complete if and only if J is dense in $[0, \pi)$.*

Proof The proof follows [5]. Assume that J is dense and $T \in \mathcal{T}(\mathcal{H})$ such that $\text{tr}[T\mathbf{Q}_\theta(X)] = 0$ for all $\theta \in J$ and $X \in \mathcal{B}(\mathbb{R})$. It follows that

$$0 = \frac{1}{\sqrt{2\pi}} \int e^{-ipq} \text{tr}[T\mathbf{Q}_\theta(dq)] = \frac{1}{\sqrt{2\pi}} \text{tr}[W(p \sin \theta, -p \cos \theta)T]$$

for all $p \in \mathbb{R}$ and $\theta \in J$. The continuity of $\mathbb{R}^2 \ni (x, y) \mapsto \text{tr}[W(x, y)T] \in \mathbb{C}$ implies that $\text{tr}[W(x, y)T] = 0$ for all $(x, y) \in \mathbb{R}^2$ and thus $T = 0$ by the injectivity of the Wigner transform (see, Remark 18.2 below). Thus, the set of quadratures is informationally complete. The ‘only if’-part is left to an exercise. $\square$

Remark 18.1 The set J in Theorem 18.1 can be chosen to be countable, e.g. $J = [0, 2\pi) \cap \mathbb{Q}$. Hence, in principle, it suffices to measure a sequence of quadrature observables in order to determine the state of the system. $\triangleleft$

Remark 18.2 The Wigner transform of a trace class operator $T \in \mathcal{T}(\mathcal{H})$ is the continuous square integrable phase space function

$$\mathbb{R}^2 \ni x \mapsto \mathcal{W}_T(q, p) = \frac{1}{\pi} \text{tr}[TW(q, p)\Pi W(-q, -p)] = \frac{1}{\pi} \text{tr}[W(2q, 2p)\Pi T] \in \mathbb{C},$$

where $W(q, p)$ are the Weyl operators and $\Pi = U(\pi)$ is the parity operator. For any two $T_1, T_2 \in \mathcal{T}(\mathcal{H})$,

$$\langle \mathcal{W}_{T_1} \mid \mathcal{W}_{T_2} \rangle_{L^2(\mathbb{R}^2)} = \int_{\mathbb{R}^2} \overline{\mathcal{W}_{T_1}(q, p)} \mathcal{W}_{T_2}(q, p) dq dp = \frac{1}{2\pi} \text{tr}[T_1 T_2],$$

showing, in particular, that the Wigner transform is injective—it separates states, that is, for any two states $\varrho_1, \varrho_2 \in \mathcal{S}(\mathcal{H})$, if $\mathcal{W}_{\varrho_1} = \mathcal{W}_{\varrho_2}$, then $\varrho_1 = \varrho_2$.

The Radon transform is a continuous linear map $\mathbf{R} : L^1(\mathbb{R}^2) \rightarrow L^1(\mathbb{T} \times \mathbb{R})$ given by

$$(\mathbf{R}f)(\theta, x) = \int_{\mathbb{R}} f(x \cos \theta - t \sin \theta, x \sin \theta + t \cos \theta) dt$$

for $d\theta dx$ -almost all $(\theta, x) \in [0, 2\pi) \times \mathbb{R}$, where we have identified θ with $e^{i\theta} \in \mathbb{T}$. Its restriction to smooth rapidly decreasing functions is injective [6, Theorem 2.4].

For any state $\varrho \in \mathcal{S}(\mathcal{H})$, let $\varrho^{\mathbf{Q}_\theta}$ denote the density of the probability measure $\mathcal{B}(\mathbb{R}) \ni X \mapsto \text{tr}[\varrho \mathbf{Q}_\theta(X)] \in [0, 1]$ defined by the pair $(\varrho, \mathbf{Q}_\theta)$. If ϱ is such that $\mathcal{W}_\varrho \in L^1(\mathbb{R}^2)$ then the Radon transform $\mathbf{R}(\mathcal{W}_\varrho)$ of $\mathcal{W}_\varrho$ can easily be computed, and one gets

$$\mathbf{R}(\mathcal{W}_\varrho)(\theta, x) = \int_{\mathbb{R}} \mathcal{W}_\varrho(x \cos \theta - t \sin \theta, x \sin \theta + t \cos \theta) dt = \varrho^{\mathbf{Q}_\theta}(x)$$

for (almost all) $(\theta, x) \in [0, 2\pi) \times \mathbb{R}$. Thus the Radon transform of the Wigner function of a state constitutes its rotated quadrature distributions. The Wigner transform separates states and the Radon transform separates smooth rapidly decreasing functions. Let $\mathcal{S}^{\text{smooth}} = \{\varrho \in \mathcal{S}(\mathcal{H}) \mid \mathcal{W}_\varrho \in C_{\downarrow}^\infty(\mathbb{R}^2)\}$. Then for any two states $\varrho_1, \varrho_2 \in \mathcal{S}^{\text{smooth}}$,

$$\varrho_1 \neq \varrho_2 \iff \mathbf{R}(\mathcal{W}_{\varrho_1}) \neq \mathbf{R}(\mathcal{W}_{\varrho_2}). \quad (18.1)$$

Although this result has formed the starting point of state tomography, it is not very useful since it requires the apriori knowledge on the smoothness of the state function. $\triangleleft$

Unsharp Pairs

In this subsection we consider informationally complete joint observables of unsharp approximations of position and momentum.

Let μ and ν be any two probability measures $\mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ and consider the unsharp position and momentum observables $\mu * \mathbf{Q}$ and $\nu * \mathbf{P}$ in $\mathcal{H} = L^2(\mathbb{R})$. By Theorem 15.1 they are jointly measurable if and only if they are the margins of a covariant phase space observable $\mathbf{G}^K$, in which case the measures μ and ν are Fourier related, that is, $\mu = \mathbf{Q}_{\Pi K \Pi}$ and $\nu = \mathbf{P}_{\Pi K \Pi}$. The following theorem characterises the informational completeness of a $\mathbf{G}^K$.

Theorem 18.2 *The phase space observable $\mathbf{G}^K$ is informationally complete if and only if the support of the function $(q, p) \mapsto \text{tr}[K W(q, p)]$ is the whole set $\mathbb{R}^2$.*

The proof of this theorem utilises the following two facts. First, for any $T \in \mathcal{T}_s(\mathcal{H})$, the density of the measure $\mathcal{B}(\mathbb{R}^2) \ni Z \mapsto \text{tr}[T \mathbf{G}^K(Z)] \in \mathbb{R}$ can be expressed as the Fourier transform (see e.g. [7, 8])

$$\int e^{-i(qp' - q'p)} \overline{\text{tr}[T W(q, p)]} \text{tr}[K W(q, p)] dq dp = 2\pi \text{tr}[T W(q', p') K W(q', p')^*]. \quad (18.2)$$

Second, a function $g : \mathbb{R}^2 \rightarrow \mathbb{C}$ is of the form

$$g(q, p) = \text{tr}[K W(q, p)], \quad (q, p) \in \mathbb{R}^2, \quad (18.3)$$

for some positive trace class operator K , if and only if g is continuous and positive definite [8, Proposition 3.4 (7)]. Recall that a function $g : \mathbb{R}^2 \rightarrow \mathbb{C}$ is positive definite if for all $n \in \mathbb{N}$ and $(q_1, p_1), \dots, (q_n, p_n) \in \mathbb{R}^2$

$$(i, j) \mapsto g(q_i - q_j, p_i - p_j) \quad (18.4)$$

is a positive definite matrix. The following condition [8] is a “twisted” version of this: for all $n \in \mathbb{N}$ and $(q_1, p_1), \dots, (q_n, p_n) \in \mathbb{R}^2$,

$$(i, j) \mapsto g(q_i - q_j, p_i - p_j) e^{\frac{1}{2}i(q_i p_j - q_j p_i)} \quad (18.5)$$

is a positive definite matrix. Bochner's theorem for the Weyl transform [8, Proposition 3.4(7)] states the following: A function $g : \mathbb{R}^2 \rightarrow \mathbb{C}$ is of the form $g(q, p) = \text{tr}[KW(q, p)]$, $(q, p) \in \mathbb{R}^2$, for some positive trace class operator K , if and only if g is continuous, and has the property (18.5).

Let $B((q, p), \epsilon)$ denote the open ϵ -ball centered at $(q, p) \in \mathbb{R}^2$. The result (18.3) can now be used to prove the following lemma.

Lemma 18.1 *For each $\epsilon > 0$, and $(q_0, p_0) \in \mathbb{R}^2$ there exists a trace class operator $T \neq 0$, such that the function $\mathbb{R}^2 \ni (q, p) \mapsto \text{tr}[TW(q, p)] \in \mathbb{C}$ is supported in $B((q_0, p_0), \epsilon)$.*

Proof Pick some positive operator ϱ_0 of trace one, and let χ denote the characteristic function of $B(0, \epsilon/4)$. Then the convolution $\chi * \chi : \mathbb{R}^2 \rightarrow \mathbb{C}$ is continuous, and since $\bar{\chi}(q, p) = \chi(-(q, p))$, it is positive definite. Since the elementwise product of two positive definite matrices is positive definite, it follows that the function $(q, p) \mapsto \text{tr}[\varrho_0 W(q, p)](\chi * \chi)(q, p)$ has the property (18.5). Since it is also continuous, the result (18.3) gives a positive trace class operator K such that $\text{tr}[KW(q, p)] = \text{tr}[\varrho_0 W(q, p)](\chi * \chi)(q, p)$, $(q, p) \in \mathbb{R}^2$. Now $K \neq 0$, since clearly $\text{tr}[KW(0)] = (\chi * \chi)(0) > 0$. Now the support of $\chi * \chi$ is contained in $B(0, \epsilon)$, so also the support of $(q, p) \mapsto \text{tr}[KW(q, p)]$ is included in this ball.

Since the trace class is an ideal in the algebra of bounded operators, then $T = KW(-(q_0, p_0))$ is a trace class operator as well. It is nonzero, because $W(-(q_0, p_0))$ is a unitary operator. Now

$$\text{tr}[TW(q, p)] = e^{i\frac{1}{2}(-q_0 p + q p_0)} \text{tr}[KW(-q_0 + q, -p_0 + p)],$$

for all (q, p) , so $\text{tr}[TW(q, p)] = 0$ if $(q, p) \notin B((q_0, p_0), \epsilon)$. $\square$

Proof (Theorem 18.2) If the support in question is full, then by (18.2), the condition

$$T \in \mathcal{T}(\mathcal{H}), T = T^*, \text{tr}[TG^K(q, p)] = 0 \text{ for all } X$$

implies that $\text{tr}[TW(q, p)] = 0$ for all (q, p) , because the Fourier transform is injective, and this function is continuous. But the Weyl transform is also injective, so $T = 0$, which shows that G^K is informationally complete.

Suppose now that $\text{tr}[W(q, p)K]$ vanishes in some open set, i.e. there is an $(q_0, p_0) \in \mathbb{R}^{2n}$, and $\epsilon > 0$, such that $\text{tr}[KW(q, p)] = 0$ for $(q, p) \in B((q_0, p_0), \epsilon)$. Since $W(q, p)^*$ equals $W(-(q, p))$ up to a phase, and $K^* = K$, it follows that $\text{tr}[KW(q, p)] = 0$ also for $(q, p) \in B(-(q_0, p_0), \epsilon)$. Choosing a trace class operator T as in the above lemma, we see that $\text{tr}[T^*W(q, p)] = 0$ if $(q, p) \notin B(-(q_0, p_0), \epsilon)$. As $T \neq 0$, at least one of the selfadjoint operators $T + T^*$ and $i(T - T^*)$ is nonzero. Let T' be this operator. Then the support of $(q, p) \mapsto \text{tr}[T'W(q, p)]$ is contained in $B((q_0, p_0), \epsilon) \cup B(-(q_0, p_0), \epsilon)$, and so $\text{tr}[KW(q, p)]\text{tr}[T'W(q, p)] = 0$ for all (q, p) . Hence, by (18.2), this T' violates the condition of informational completeness of G^K . Consequently, G^K is not informationally complete. $\square$

An important special case of informational completeness of a $\mathbf{G}^K$ is obtained whenever the function $(q, p) \mapsto \text{tr}[KW(q, p)]$ has no zeros, that is, the generating operator K is *regular*. According to Wiener's approximation theorem [8, Proposition 3.5], K is regular if and only if for any trace class operator T , the condition $\text{tr}[TW(q, p)KW(q, p)^*] = 0$ for all (q, p) , implies $T = 0$, or, equivalently, the functions $(q, p) \mapsto \text{tr}[KW(q, p)TW(q, p)^*]$, T in trace class, are dense in $L^1(\mathbb{R}^2)$.

The margins $X \mapsto \mathbf{G}^K(X \times \mathbb{R})$ and $Y \mapsto \mathbf{G}^K(\mathbb{R} \times Y)$ of a phase space observable $\mathbf{G}^K$ are the approximate position and momentum observables $\mu * \mathbf{Q}$ and $\nu * \mathbf{P}$. Consistently with the notation $\text{tr}[\varrho \mathbf{Q}(q, p)] = \int_X \mathbf{Q}_\varrho(q, p) dq dp$, the densities of the probability measures μ and ν can be written as $(\Pi K \Pi^*)^{\mathbf{Q}}$ and $(\Pi K \Pi^*)^{\mathbf{P}}$, respectively, where Π is the parity operator.

Similarly, one may compute the “ θ -marginal observable” of $\mathbf{G}^K$ in any direction θ , that is, in any θ -rotated coordinate system. It is an easy exercise to show that the resulted marginal observable $\mu^\theta * \mathbf{Q}_\theta$ has the density

$$K^{\mathbf{Q}(\theta-\pi)} * \varrho^{\mathbf{Q}_\theta}, \quad (18.6)$$

for any state ϱ . Since the parity operator $\Pi = U(\pi)$ and the Fourier–Plancherel operator $F = U(\frac{\pi}{2})$, we note that (18.6) contains, in particular, the Cartesian marginal distributions

$$K^{\mathbf{Q}(-\pi)} * \varrho^{\mathbf{Q}}, \quad \text{and} \quad K^{\mathbf{Q}(-\frac{\pi}{2})} * \varrho^{\mathbf{P}}.$$

The marginal observable $\mu^\theta * \mathbf{Q}_\theta$ is a smeared version of the sharp observables $\mathbf{Q}_\theta$, so that its state distinction power is a priori weaker than that of $\mathbf{Q}_\theta$. Taking the Fourier transform of the convolution $K^{\mathbf{Q}(\theta-\pi)} * \varrho^{\mathbf{Q}_\theta}$, and observing that, for any $p \in \mathbb{R}$,

$$\begin{aligned} \int_{\mathbb{R}} e^{-ipx} K^{\mathbf{Q}(\theta-\pi)}(q, p) dx &= \text{tr}[U(\theta - \pi) K U(\theta - \pi)^* W(0, p)] \\ &= \text{tr}[K W(-p \sin \theta, -p \cos \theta)], \end{aligned}$$

we note that if the support of the function $p \mapsto \text{tr}[K W(-p \sin \theta, -p \cos \theta)]$ is the whole space $\mathbb{R}$, then the unsharp distribution $K^{\mathbf{Q}(\theta-\pi)} * \varrho^{\mathbf{Q}_\theta}$ determines the sharp distribution $\varrho^{\mathbf{Q}_\theta}$, meaning that the unsharp observable $\mu^\theta * \mathbf{Q}_\theta$ is informationally equivalent to the sharp observables $\mathbf{Q}_\theta$. Clearly, if K is regular, then all the unsharp quadratures $\mu^\theta * \mathbf{Q}_\theta$ are informationally equivalent to the sharp quadratures $\mathbf{Q}_\theta$. Thus, in this case, the informational completeness of the phase space observable $\mathbf{G}^K$ is directly related to the informational completeness of the quadrature observables $\mathbf{Q}_\theta$, $\theta \in [0, \pi)$.

For a regular phase space observable $\mathbf{G}^K$ one thus has:

$$\{\varrho\} = [\varrho]^{\mathbf{G}^K} \subset [\varrho]^{\{\mu^\theta * \mathbf{Q}_\theta \mid \theta \in [0, \pi)\}} = [\varrho]^{\{\mathbf{Q}_\theta \mid \theta \in [0, \pi)\}} = \{\varrho\} \quad (18.7)$$

for any state $\varrho \in \mathcal{S}(\mathcal{H})$, showing, in particular, that the inclusion in (18.7) is an equality.

Remark 18.3 For a phase space observable $\mathbf{G}^K$, let

$$Z(K) = \{(q, p) \in \mathbb{R}^2 \mid \text{tr}[KW(q, p)] = 0\}$$

denote the zeros of the function $(q, p) \mapsto \text{tr}[KW(q, p)]$. By Theorem 18.2, $\mathbf{G}^K$ is informationally complete if and only if the set $Z(K)$ does not contain an open set, that is, $\overline{Z(K)}^c = \mathbb{R}^2$. Clearly, if $Z(K)$ is of (Lebesgue) measure zero, that is, $\lambda(Z(K)) = 0$, then it cannot contain an open set, and thus $\mathbf{G}^K$ is informationally complete. An even more stringent condition is that $Z(K)$ is empty, that is, $\mathbf{G}^K$ is regular. The proof of Theorem 18.2 is adopted from [9]. There it is also shown that neither of the obvious implications $Z(K) = \emptyset \implies \lambda(Z(K)) = 0$ and $\lambda(Z(K)) = 0 \implies \overline{Z(K)}^c = \mathbb{R}^2$ can be reversed. These results correct some of the earlier statements claiming the necessity of the condition $\lambda(Z(K)) = 0$ [7, 10, 11] or even the necessity of $Z(K) = \emptyset$ [12]. Since $\overline{Z(K)}^c = \mathbb{R}^2$ is indeed a necessary condition, this also shows, contrary to some formal state reconstruction formulas [13], that not all covariant phase space observables can be used in quantum tomography. For instance, a phase space observable generated by a slit state $K = |\varphi\rangle\langle\varphi|$, with a compactly supported φ , is not informationally complete. $\triangleleft$

18.3 State Reconstruction

A state $\varrho \in \mathcal{S}(\mathcal{H})$ of a quantum system can be identified with the totality of the measurement outcome probabilities $\mathbf{E}_\varrho$, with $\mathbf{E}$ running through all the observables of the system. The notion of informational completeness singles out classes of observables $\mathcal{M}$ which are such that their measurement statistics $\mathbf{E}_\varrho$, $\mathbf{E} \in \mathcal{M}$, already suffice to separate the states. If no a priori knowledge of a state ϱ is given, then informational completeness of $\mathcal{M}$ is necessary for the identification of the state ϱ . The problem of state reconstruction is the question of determining the state ϱ from an informationally complete statistics $\mathbf{E}_\varrho$, $\mathbf{E} \in \mathcal{M}$.

Any state ϱ is uniquely determined by its matrix elements $\varrho_{nm} = \langle n | \varrho | m \rangle$ with respect to an orthonormal basis $\{|n\rangle\}$. In the strategies considered here such matrix elements ϱ_{nm} will be constructed from the statistics $\mathbf{E}_\varrho$, $\mathbf{E} \in \mathcal{M}$, using appropriate integration or differentiation techniques and infinite matrix inversion [14–16].

Remark 18.4 The early methods to reconstruct the state from the rotated quadrature statistics $\varrho^{\mathbf{Q}_\theta}$, $\theta \in [0, 2\pi)$, were based on the fact that this statistics constitutes the Radon transform $\mathbf{R}(\mathcal{W}_\varrho)$ of the Wigner transform $\mathcal{W}_\varrho$ of the state ϱ , that is,

$$\mathbf{R}(\mathcal{W}_\varrho)(\theta, x) = \int_{\mathbb{R}} \mathcal{W}_\varrho(x \cos \theta - t \sin \theta, x \sin \theta + t \cos \theta) dt = \varrho^{\mathbf{Q}_\theta}(x)$$

for (almost all) $(\theta, x) \in [0, 2\pi) \times \mathbb{R}$ (see Remark 18.2). Inverting this integral transform would give $\mathcal{W}_\varrho$. While the Wigner transform provides a unique description of a state and it can be used e.g. to compute expectation values of several observables, it does not provide direct access to the state ϱ . Moreover, this inversion method is only known to work for smooth functions $\mathcal{W}_\varrho$.

To our knowledge, the reconstruction formula (18.12) was first given by Leonhardt and D'Ariano [17], who wanted to replace the traditional and mathematically troublesome inverse Radon transform scheme with a direct reconstruction of the density matrix in terms of the quadrature data. The same formula was later obtained as a special case of more general group theoretical results by Cassinelli et al. [18]. The derivation of (18.12) followed in this text is due to [14] $\triangleleft$

State Reconstruction Via Rotated Quadratures

For any rotated quadrature $\mathbf{Q}_\theta$, $\theta \in [0, 2\pi]$, and for any state $\varrho \in \mathcal{S}(\mathcal{H})$, the probability measure $X \mapsto \text{tr}[\varrho \mathbf{Q}_\theta(X)]$ has a density $\varrho^{\mathbf{Q}_\theta}$. Let D_ϱ^{qd} denote the function $\mathbb{R} \times [0, 2\pi) \ni (x, \theta) \mapsto \varrho^{\mathbf{Q}_\theta}(x) \in \mathbb{R}^+$. This function is both measurable and integrable. It is easy to show that, for any bounded measurable function $g : \mathbb{R} \rightarrow \mathbb{C}$,

$$\int_{\mathbb{R}} g(x) D_{\varrho,k}^{\text{qd}}(x) dx = \sum_{n=0}^{\infty} \varrho_{n+k,n} \int_{\mathbb{R}} g(x) h_n(x) h_{n+k}(x) dx, \quad (18.8)$$

where we define

$$D_{\varrho,k}^{\text{qd}}(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} D_\varrho^{\text{qd}}(x, \theta) d\theta.$$

Equation (18.8) gives a reconstruction formula for the state ϱ if one is able to choose functions g so that the matrix elements $\varrho_{n+k,n}$ are obtained from this equation.

Remark 18.5 Dawson's integral is the function on $\mathbb{R}$ given by the integral

$$\text{daw}(x) = e^{-x^2} \int_0^x e^{t^2} dt$$

(see e.g. [19, pp. 298–299] or [20, Chap. 42]). It is an analytic (bounded) function whose derivatives $\text{daw}^{(k)}$, $k \in \mathbb{N}_0$, all vanish at infinity, that is,

$$\lim_{x \rightarrow \pm\infty} \text{daw}^{(k)}(x) = 0.$$

Denote $Y = 2 \text{daw}^{(1)}$ and let $p \in \mathbb{N}_0$. One can write the following Hermite series representations:

$$Y^{(2p)}(x) = (-1)^p 2^p \sum_{k=0}^{\infty} \frac{(-1)^k (k+p)!}{2^k (2k)!} H_{2k}(x),$$

$$Y^{(2p+1)}(x) = (-1)^{p+1} 2^p \sum_{k=0}^{\infty} \frac{(-1)^k (k+p+1)!}{2^k (2k+1)!} H_{2k+1}(x)$$

for all $x \in \mathbb{R}$. Using the above series one can calculate the matrix elements of the *selfadjoint* bounded operators $Y^{(k)}(Q) = \int_{\mathbb{R}} Y^{(k)}(x) dQ(x)$ on $L^2(\mathbb{R})$ [14]: For all $m, n, p, u \in \mathbb{N}$,

$$\begin{aligned} \langle m | Y^{(2p)}(Q) | m + 2u + 1 \rangle &= 0, \\ \langle m | Y^{(2p)}(Q) | m + 2u \rangle &= (-1)^{p+u+m} 2^p (p+u)! \sqrt{\frac{m!}{(m+2u)!}} \binom{p-u}{m} \\ &= 0, \quad \text{if and only if } p \geq u \text{ and } m > p - u, \\ \langle m | Y^{(2p+1)}(Q) | m + 2u \rangle &= 0, \\ \langle m | Y^{(2p+1)}(Q) | m + 2u + 1 \rangle &= (-1)^{p+u+m+1} 2^{p+1/2} (p+u+1)! \times \\ &\quad \times \sqrt{\frac{m!}{(m+2u+1)!}} \binom{p-u}{m} \\ &= 0, \quad \text{if and only if } p \geq u \text{ and } m > p - u. \end{aligned}$$

◁

Consider the function $Y = 2 \operatorname{daw}^{(1)}$ and put $g = Y^{(k+2l)}$ in (18.8): one gets

$$D_{\varrho, k, l}^{\text{qd}} = \int_{\mathbb{R}} Y^{(k+2l)}(x) D_{\varrho, k}^{\text{qd}}(x) dx = \sum_{n=0}^l c_{ln}(k) \varrho_{n+k, n}, \quad (18.9)$$

where, by the above remark,

$$\begin{aligned} c_{ln}(k) &= \int_{\mathbb{R}} Y^{(k+2l)}(x) h_n(x) h_{n+k}(x) dx = \langle n | Y^{(k+2l)}(Q) | n+k \rangle \\ &= (-1)^{l+n+k} 2^{k/2+l} (k+l)! \sqrt{\frac{n!}{(n+k)!}} \binom{l}{n}. \end{aligned}$$

It is now clear that the elements $\varrho_{n+k, n}$ can be solved recursively from (18.9). In order to get an explicit formula, observe that for any fixed $k \in \mathbb{N}$, the matrix relation (18.9) now assumes the form

$$y_l = \sum_{n=0}^l (-1)^n \binom{l}{n} x_n, \quad (18.10)$$

where

$$y_l = (-1)^l \frac{D_{\varrho,k,l}^{\text{qd}}}{2^l (k+l)!}, \quad x_n = (-1)^k \sqrt{\frac{2^k n!}{(k+n)!}} \varrho_{n+k,n}. \quad (18.11)$$

According to [21, p. 43], the Eq. (18.10) can be solved for x_n to give

$$x_n = \sum_{l=0}^n (-1)^l \binom{n}{l} y_l,$$

which immediately yields an explicit reconstruction formula for the matrix elements of the state ϱ , in terms of the quantities $D_{\varrho,k,l}^{\text{qd}}$ determined by the measurement statistics $\varrho^{\mathbf{Q}_\theta}$:

$$\varrho_{n+k,n} = (-1)^k \sqrt{\frac{(n+k)!}{2^k n!}} \sum_{l=0}^n \binom{n}{l} \frac{D_{\varrho,k,l}^{\text{qd}}}{2^l (k+l)!}, \quad n, k \in \mathbb{N}. \quad (18.12)$$

where

$$D_{\varrho,k,l}^{\text{qd}} = \int_{\mathbb{R}} Y^{(k+2l)}(x) \int_0^{2\pi} e^{ik\theta} \varrho^{\mathbf{Q}_\theta}(x) \frac{d\theta}{2\pi} dx.$$

State Reconstruction Via the Phase Space Observable $\mathbf{G}^{(0)}$

A phase space observable $\mathbf{G}^K$ is informationally complete if and only if the generating operator K is such that the support of the function $(q, p) \mapsto \text{tr}[K W(q, p)]$ is the whole $\mathbb{R}^2$. In particular, if K is the vacuum state $K = |0\rangle\langle 0|$, then the density

$$(q, p) \mapsto g_\varrho^{(0)}(q, p) = \text{tr}[\varrho W(q, p)|0\rangle\langle 0|W(q, p)^*] = \langle 0 | W(q, p)^* \varrho W(q, p) | 0 \rangle$$

of the probability measure $\mathbf{G}_\varrho^{(0)}$ is the Q -function of the state ϱ for which a reconstruction formula can quite easily be obtained. Indeed, writing the number $\langle 0 | W(q, p)^* \varrho W(q, p) | 0 \rangle$ in the form

$$g_\varrho^{(0)}(q, p) = \sum_{n,m=0}^{\infty} \langle 0 | W(q, p)^* | n \rangle \langle n | \varrho | m \rangle \langle m | W(q, p) | 0 \rangle,$$

and noting that

$$\langle 0 | W(q, p)^* | n \rangle = e^{-|z|^2/2} \frac{\bar{z}^n}{\sqrt{n!}} = \frac{1}{\sqrt{n!}} e^{-in\theta} e^{-r^2/2} r^n,$$

where $z = \frac{1}{\sqrt{2}}(q + ip) = r e^{i\theta}$, one gets

$$g_{\varrho}^{(0)}(r e^{i\theta}) = \sum_{n,m=0}^{\infty} \langle n | \varrho | m \rangle e^{i(m-n)\theta} \frac{1}{\sqrt{n!m!}} e^{-r^2} r^{n+m}.$$

Since $|g_{\varrho}^{(0)}(r e^{i\theta})| \leq e^{-r^2} \left(\sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} r^n \right)^2 < \infty$ one gets

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} g_{\varrho}^{(0)}(r e^{i\theta}) d\theta = \sum_{n=0}^{\infty} \langle n | \varrho | n+k \rangle \frac{1}{\sqrt{n!(n+k)!}} e^{-r^2} r^{2n+k}, \quad (18.13)$$

and

$$\langle n | \varrho | n+k \rangle = \frac{\sqrt{n!(n+k)!}}{(2n+k)!} \frac{\partial^{2n+k}}{\partial r^{2n+k}} \left[e^{r^2} \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} g_{\varrho}^{(0)}(r e^{i\theta}) d\theta \right]_{r=0}. \quad (18.14)$$

The next subsection generalises the results (18.13) and (18.14). Actually, the proof covers all the so-called Cahill–Glauber s -parametrised distributions which include, in particular, phase space observable generated by the operators $K^\lambda = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k |k\rangle \langle k|$, $0 \leq \lambda < 1$.

Cahill–Glauber s -parametrised Distributions

The s -parametrised quasiprobability distributions were introduced in quantum optics as mathematical tools [22, 23], but they have since become accessible also to direct measurements. A scheme for indirect determination of these distributions was suggested by Vogel and Risken [24], and the pioneering experimental work was done by Smith et al. [25]. The s -distributions also arise when considering realistic measurements, where the detectors are not assumed to be ideal [26, 27]. In that case, the measurement outcome statistics correspond to certain s -distributions, where the parameter s is related to the efficiency η of the detectors, by $s = 1 - 2/\eta$. For our purposes it is convenient to define a parameter λ as $\lambda = \frac{s+1}{s-1}$. This gives a bijective mapping on $\mathbb{C} \setminus \{1\}$. We will use λ as the parameter for these distributions, and for this reason we will call the s -parametrised quasiprobability distributions λ -distributions.

Let $\lambda \in \mathbb{C}$, $|\lambda| \leq 1$. Define a bounded operator $K^\lambda = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k |k\rangle \langle k|$; it has norm $\|K^\lambda\| = |1 - \lambda|$. If $|\lambda| < 1$ then K^λ is a trace class operator with

$$\text{tr}[K^\lambda] = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k = 1.$$

For each $\lambda \in \mathbb{C}$, $|\lambda| \leq 1$, define a (weakly continuous bounded) function $D^\lambda : [0, \infty) \times [0, 2\pi) \rightarrow \mathcal{L}(\mathcal{H})$ by $D^\lambda(r, \theta) = W(re^{i\theta})K^\lambda W(re^{i\theta})^*$ where $W(z) = W(q, p)$, $z = \frac{1}{\sqrt{2}}(q + ip) = re^{i\theta}$. For each state ϱ define the λ -distribution as the phase space distribution $D_\varrho^\lambda : [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{C}$,

$$D_\varrho^\lambda(r, \theta) = \text{tr}[\varrho D^\lambda(r, \theta)] = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k \langle k | D(re^{i\theta})^* \varrho D(re^{i\theta}) | k \rangle.$$

Note that $\lambda = 0$ gives us the Q -function, and $\lambda = -1$ gives us the Wigner function (up to a constant scaling factor) of the state.

If $\lambda \geq 0$, then the operator density D^λ defines a covariant phase space observable

$$\mathcal{B}(\mathbb{C}) \ni Z \mapsto \mathbf{G}^{K^\lambda}(Z) = \frac{1}{\pi} \int_Z D^\lambda(r, \theta) r dr d\theta \in \mathcal{L}(\mathcal{H}),$$

in which case the λ -distribution can be measured via eight-port homodyne detection. For other values of λ , the λ -distribution is not a measurable quantity in the sense of the theory of measurement. However, as shown in [28], it can be constructed from the rotated quadrature distributions, which are obtainable via balanced homodyne detection.

Let us consider now the operator valued function D^λ . Using the fact that $W(re^{i\theta}) = R(\theta)W(r)R(\theta)^*$, a direct calculation gives us the matrix elements of $D^\lambda(r, \theta)$ with respect to the number basis:

$$\langle n | D^\lambda(r, \theta) | m \rangle = e^{i(n-m)\theta} K_{nm}^\lambda(r, \theta),$$

where

$$K_{nm}^\lambda(r) = \langle n | D(r) K^\lambda D(r)^* | m \rangle.$$

By the formula of Cahill and Glauber [14, 22] we get

$$K_{nm}^\lambda(r) = \sqrt{\frac{n!}{m!}} (1 - \lambda)^{m-n+1} e^{-(1-\lambda)r^2} r^{m-n} \lambda^n L_n^{m-n}((2 - \lambda - \lambda^{-1})r^2) \quad (18.15)$$

where

$$L_n^\alpha(x) = \sum_{u=0}^n \frac{(-1)^u}{u!} \binom{n+\alpha}{n-u} x^u$$

is the associated Laguerre polynomial.

Since $\|D^\lambda(r, \theta)\| \leq \|K^\lambda\| = |1 - \lambda|$ for all $r \in [0, \infty)$, $\theta \in [0, 2\pi)$, the operator

$$D_k^\lambda(r) = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} D^\lambda(r, \theta) d\theta$$

is well defined as a weak integral for each $k \in \mathbb{N}_0$. In addition, we have

$$D_{\varrho,k}^{\lambda}(r) = \text{tr}[\varrho D_k^{\lambda}(r)] = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} D_{\varrho}^{\lambda}(r, \theta) d\theta,$$

for each state ϱ . Since

$$\begin{aligned} D_{\varrho}^{\lambda}(r, \theta) &= \text{tr}[\varrho D^{\lambda}(r, \theta)] \\ &= \sum_{m,n=0}^{\infty} \varrho_{mn} \langle n | D^{\lambda}(r, \theta) | m \rangle = \sum_{m,n=0}^{\infty} \varrho_{mn} e^{i(n-m)\theta} K_{nm}^{\lambda}(r), \end{aligned}$$

we have

$$D_{\varrho,k}^{\lambda}(r) = \frac{1}{2\pi} \sum_{m,n=0}^{\infty} \varrho_{mn} \int_0^{2\pi} e^{i(k+n-m)\theta} K_{nm}^{\lambda}(r) d\theta = \sum_{n=0}^{\infty} \varrho_{n+k,n} K_{n,n+k}^{\lambda}(r).$$

Suppose then that ϱ is such that its matrix elements $\varrho_{mn} \neq 0$ only for finitely many pairs $(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0$. Then

$$D_{\varrho,k,l}^{\lambda} = \left. \frac{\partial^l [e^{(1-\lambda)r^2} D_{\varrho,k}^{\lambda}(r)]}{\partial r^l} \right|_{r=0} = \sum_{n=0}^{\infty} \varrho_{n+k,n} \left. \frac{\partial^l [e^{(1-\lambda)r^2} K_{n,n+k}^{\lambda}(r)]}{\partial r^l} \right|_{r=0}$$

where the sum is now finite. By (18.15) and

$$\left. \frac{\partial^l r^{2u+k}}{\partial r^l} \right|_{r=0} = l! \delta_{2u+k,l}$$

one gets two (nonzero) cases for $D_{\varrho,k,l}^{\lambda}$: (i) when k is even, then l must be even, and (ii) when k is odd, then l must be odd.

Let $h, j \in \mathbb{N}$ and consider case (i) when $k = 2h$ and $l = 2j$. Now

$$\begin{aligned} D_{\varrho,2h,2j}^{\lambda} &= \sum_{n=0}^{\infty} \varrho_{n+2h,n} \sqrt{\frac{n!}{(n+2h)!}} (1-\lambda)^{2h+1} \times \\ &\quad \times \sum_{u=0}^n \frac{\lambda^{n-u}}{u!} \binom{n+2h}{n-u} (1-\lambda)^{2u} (2j)! \delta_{u,j-h} \\ &= \sum_{n=\max\{0,j-h\}}^{\infty} \varrho_{n+2h,n} \sqrt{\frac{n!}{(n+2h)!}} \frac{(1-\lambda)^{2j+1} \lambda^{n-j+h} (n+2h)! (2j)!}{(j-h)! (n-j+h)! (j+h)!} \end{aligned}$$

$$= \binom{2j}{j-h} (1-\lambda)^{2j+1} \lambda^{h-j} \sum_{n=\max\{0, j-h\}}^{\infty} \frac{\varrho_{n+2h,n} \lambda^n \sqrt{(n+2h)!n!}}{(n-j+h)!}. \quad (18.16)$$

Consider then case (ii) and put $k = 2h + 1$ and $l = 2j + 1$. Then

$$\begin{aligned} D_{\varrho, 2h+1, 2j+1}^{\lambda} &= \binom{2j+1}{j-h} (1-\lambda)^{2j+2} \lambda^{h-j} \times \\ &\times \sum_{n=\max\{0, j-h\}}^{\infty} \frac{\varrho_{n+2h+1,n} \lambda^n \sqrt{(n+2h+1)!n!}}{(n-j+h)!}. \end{aligned} \quad (18.17)$$

Define, for all $n, m \in \mathbb{N}_0$,

$$\tilde{\varrho}_{mn}^{\lambda} = \varrho_{mn} \lambda^{\min\{m,n\}} \sqrt{m!n!}.$$

Then both Eqs. (18.16) and (18.17) reduce to the single one:

$$\sum_{n=\max\{0, (l-k)/2\}}^{\infty} \frac{\tilde{\varrho}_{n+k,n}^{\lambda}}{(n - (l-k)/2)!} = D_{\varrho, k, l}^{\lambda} \binom{l}{(l-k)/2}^{-1} (1-\lambda)^{-(l+1)} \lambda^{(l-k)/2} \quad (18.18)$$

where either $k = 2h$ and $l = 2j$ or $k = 2h + 1$ and $l = 2j + 1$. In order to invert Eq. (18.18), we need the following lemma which is easy to prove.

Lemma 18.2 *Let $(x_n)_{n \in \mathbb{N}_0}$ be a sequence of complex numbers such that $x_n \neq 0$ only for finitely many n and*

$$\sum_{n=p}^{\infty} \frac{x_n}{(n-p)!} = y_p \in \mathbb{C}, \quad p \in \mathbb{N}_0.$$

Then

$$\sum_{p=n}^{\infty} \frac{(-1)^{p-n} y_p}{(p-n)!} = x_n.$$

Now Lemma 18.2 may be applied to Eq. (18.18) to give

$$\tilde{\varrho}_{n+k,n}^{\lambda} = \sum_{p=n}^{\infty} \frac{(-1)^{p-n}}{(p-n)!} D_{\varrho, k, 2p+k}^{\lambda} \binom{2p+k}{p}^{-1} (1-\lambda)^{-(2p+k+1)} \lambda^p.$$

Thus, we get the reconstruction formula

$$\begin{aligned}\varrho_{n+k,n} &= \frac{1}{\sqrt{(n+k)!n!}} \sum_{p=n}^{\infty} \binom{2p+k}{p}^{-1} \frac{(-1)^{p-n} \lambda^{p-n}}{(1-\lambda)^{2p+k+1} (p-n)!} D_{\varrho,k,2p+k}^{\lambda} \\ &= \sqrt{\frac{n!}{(n+k)!}} \sum_{p=n}^{\infty} \binom{p}{n} \frac{(p+k)!}{(2p+k)!} \frac{(-\lambda)^{p-n}}{(1-\lambda)^{2p+k+1}} \left. \frac{\partial^{2p+k} [e^{(1-\lambda)r^2} D_{\varrho,k}^{\lambda}(r)]}{\partial r^{2p+k}} \right|_{r=0}\end{aligned}\quad (18.19)$$

which holds for all $n, k \in \mathbb{N}_0$ and $\lambda \in [-1, 1)$ when the matrix (ϱ_{mn}) of a state ϱ is finite. It is shown in [14] that if $|\lambda| < \frac{1}{2}$ then the reconstruction formula (18.19) works for all states ϱ . A special case of this reconstruction formula is the formula (18.14) which is obtained from (18.19) with $\lambda = 0$.

18.4 Exercises

1. Prove Proposition 18.1.
2. Let $\mathcal{E} \subset \mathbf{E}(\mathcal{H})$ and assume that there is a $B \in \mathcal{E}, 0 \neq B \neq I$, such that $BA = AB$ for all $A \in \mathcal{E}$. Show that $\mathcal{E}$ is not informationally complete. Hint: assume that there is such a B , and let ψ be any unit vector which is not an eigenvector of B . Then the states $\varrho = P[\psi]$ and $\varrho_1 = P[e^{iB}\psi]$ are different but $\text{tr}[\varrho A] = \text{tr}[\varrho_1 A]$ for all $A \in \mathcal{E}$.
3. Prove Proposition 18.2. Hint: adopt a proof from [10].
4. The set of pure states $\mathbf{P}_1(\mathcal{H})$ exhaust the set of states $\mathbf{S}(\mathcal{H})$ in the sense that $\sigma - \text{conv}(\mathbf{P}_1(\mathcal{H})) = \mathbf{S}(\mathcal{H})$. However, informational completeness of a set $\mathcal{M}$ of observables with respect to $\mathbf{P}_1(\mathcal{H})$ does not necessarily imply that $\mathcal{M}$ is informationally complete [29]. Give an example. Hint: consider the case $\dim(\mathcal{H}) = 3$.
5. Let $a > 0$ and $b \in \mathbb{R}$. Show that the Fourier transform $\hat{\varphi}_{a,b}$ of the function $\varphi_{a,b}$,

$$\varphi_{a,b}(q) = \left(\frac{2a}{\pi}\right)^{1/4} \exp(-(a+ib)q^2)$$

is

$$\hat{\varphi}_{a,b}(p) = \left[\frac{a}{2\pi(a^2+b^2)}\right]^{1/4} \exp\left(-\frac{ap^2}{4(a^2+b^2)} + \frac{ibp^2}{4(a^2+b^2)} - \frac{i}{2} \arctan \frac{b}{a}\right).$$

6. Let $\varphi \in C_{\downarrow}^{\infty}(\mathbb{R})$ be any unit vector and define the set $\mathcal{S}(\varphi)$ of vector states $W(q, p)\varphi, (q, p) \in \mathbb{R}^2$. Show that $\mathbf{Q}$ and $\mathbf{P}$ are informationally complete with respect to $\mathcal{S}(\varphi)$. Note that it suffices to show that the expectation values $\langle W(q, p)\varphi | QW(q, p)\varphi \rangle$ and $\langle W(q, p)\varphi | PW(q, p)\varphi \rangle$ already separate these vector states.

7. Show that the pair $(\mathbf{Q}, \mathbf{P})$ is informationally complete with respect to the set of mixed states of the form $\varrho_z = e^{-|z|^2} \sum_{n=0}^{\infty} \frac{|z|^{2n}}{n!} |n\rangle\langle n|$, $z \in \mathbb{C}$ (where $|n\rangle$ is identified with the n th Hermite function h_n).
8. Show that for any spectral measure $\mathbf{E}$, if $\mathcal{S}(\mathbf{E}) \neq \emptyset$, then $[\varrho]^{(\mathbf{Q}, \mathbf{P})} \cap \mathcal{S}(\mathbf{E}) = \{\varrho\}$ for any $\varrho \in \mathcal{S}(\mathbf{E})$.
9. Show that the quadrature observables $\mathbf{Q}_\theta$ are not informationally complete in the sense of statistical expectation, that is, the numbers $\text{tr}[\varrho \mathbf{Q}_\theta]$, $\theta \in [0, 2\pi)$, do not, in general, determine the state ϱ . Hint: consider number states $|n\rangle = h_n$.
10. Assume that $J \subset [0, \pi)$ is not dense. Show that the set $\{\mathbf{Q}_\theta \mid \theta \in J\}$ of rotated quadrature observables is not informationally complete. Hint: pick an open interval $J' \subset [0, \pi) \setminus J$ and a nonzero $T \in \mathcal{T}(\mathcal{H})$ such that $\text{tr}[W(p \sin \theta, -p \cos \theta) T] = 0$ for all $\theta \in [0, \pi) \setminus J'$ and $p \in \mathbb{R}$. Show then that $\text{tr}[T \mathbf{Q}_\theta(X)] = 0$ for all $\theta \in J \subset [0, \pi) \setminus J'$ and $X \in \mathcal{B}(\mathbb{R})$.
11. Show that

$$\mathbf{R}(\mathcal{W}_\varrho)(\theta, x) = \int_{\mathbb{R}} \mathcal{W}_\varrho(x \cos \theta - t \sin \theta, x \sin \theta + t \cos \theta) dt = \varrho^{\mathbf{Q}_\theta}(x).$$

(for the notations, see Remark 18.2).

12. Show that a phase space observable $\mathbf{G}^K$ with K being the one-dimensional projection defined by a number state, or a coherent state, or a squeezed state, is informationally complete (one can even show that $\mathbf{G}^K$ is informationally complete if $\langle n|K|m\rangle \neq 0$ only for a finite number of n 's and m 's). Take then $K = |\psi\rangle\langle\psi|$, with a compactly supported ψ (in L^2 realisation), and show that $\mathbf{G}^K$ is not informationally complete.
13. For any $0 < \epsilon < 1$, let μ_ϵ be the Gaussian probability measure with the density

$$f_\epsilon(x, y) = \frac{\epsilon}{\pi(1-\epsilon)} \exp\left(-\frac{\epsilon}{1-\epsilon}(x^2 + y^2)\right).$$

Show that the convolution of μ_ϵ with the operator $|0\rangle\langle 0|$ is the operator $K^{1-\epsilon} = \epsilon \sum_k (1-\epsilon)^k |k\rangle\langle k|$, that is, $\mu_\epsilon * |0\rangle\langle 0| = K^{1-\epsilon}$, where $\mu_\epsilon * |0\rangle\langle 0|$ is defined as the weak integral

$$\mu_\epsilon * |0\rangle\langle 0| = \int W(q, p) |0\rangle\langle 0| W(q, p)^* dq dp.$$

14. Prove Lemma 18.2.

References

1. Pauli, W.: General Principles of Quantum Mechanics. Springer, Berlin (1980). Translated from the German by P. Achuthan and K. Venkatesan, With an introduction by Charles P. Enz
2. Reichenbach, H.: Philosophic Foundations of Quantum Mechanics. University of California Press, California (1944)
3. Prugovečki, E.: Information-theoretical aspects of quantum measurement. *Int. J. Theor. Phys.* **16**(5), 321–331 (1977)
4. Corbett, J.V., Hurst, C.A.: Are wave functions uniquely determined by their position and momentum distributions? *J. Austral. Math. Soc. Ser. B* **20**(2), 182–201 (1977/78)
5. Kiukas, J., Schultz, J.: Informationally complete sets of gaussian measurements. *J. Phys. A* **46**(48), 485303 (2013)
6. Helgason, S.: The Radon Transform. Progress in Mathematics, vol. 5, 2nd edn. Birkhäuser, Boston (1999)
7. Ali, S.T., Prugovečki, E.: Classical and quantum statistical mechanics in a common Liouville space. *Phys. A* **89**(3), 501–521 (1977)
8. Werner, R.: Quantum harmonic analysis on phase space. *J. Math. Phys.* **25**(5), 1404–1411 (1984)
9. Kiukas, J., Lahti, P., Schultz, J., Werner, R.F.: Characterization of informational completeness for covariant phase space observables. *J. Math. Phys.* **53**(10), 102103 (2012)
10. Busch, P., Lahti, P.J.: The determination of the past and the future of a physical system in quantum mechanics. *Found. Phys.* **19**(6), 633–678 (1989)
11. Healy Jr., D.M., Schroeck Jr., F.E.: On informational completeness of covariant localization observables and Wigner coefficients. *J. Math. Phys.* **36**(1), 453–507 (1995)
12. D’Ariano, G.M., Perinotti, P., Sacchi, M.F.: Informationally complete measurements and group representation. *J. Opt. B* **6**(6), S487–S491 (2004)
13. Wünsche, A., Bužek, V.: Reconstruction of quantum states from propensities. *Quantum Semi-class. Opt.* **9**(4), 631–653 (1997)
14. Kiukas, J., Pellonpää, J.-P., Schultz, J.: State reconstruction formulae for the S -distributions and quadratures. *Rep. Math. Phys.* **66**(1), 55–84 (2010)
15. Kiukas, J., Pellonpää, J.-P., Schultz, J.: Density matrix reconstruction from displaced photon number distributions. *J. Phys. A* **43**(9), 095303, 18 (2010)
16. Schultz, J.: State reconstruction via infinite matrix inversion. *Phys. Scripta* **T140**(1), 014060 (2010)
17. Leonhardt, U., Paul, H., D’Ariano, G.M.: Tomographic reconstruction of the density matrix via pattern functions. *Phys. Rev. A* **52**, 4899–4907 (1995)
18. Cassinelli, G., D’Ariano, G.M., De Vito, E., Levrero, A.: Group theoretical quantum tomography. *J. Math. Phys.* **41**(12), 7940–7951 (2000)
19. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions. Dover, New York (1972)
20. Spanier, J., Oldham, K.B.: An Atlas of Functions. Hemisphere Publishing Company, New York (1987)
21. Riordan, J.: Combinatorial Identities. Wiley, New York (1968)
22. Cahill, K.E., Glauber, R.J.: Ordered expansions in boson amplitude operators. *Phys. Rev.* **177**, 1857–1881 (1969)
23. Cahill, K.E., Glauber, R.J.: Density operators and quasiprobability distributions. *Phys. Rev.* **177**, 1882–1902 (1969)
24. Vogel, K., Risken, H.: Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **40**, 2847–2849 (1989)
25. Smithey, D.T., Beck, M., Raymer, M.G., Faridani, A.: Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**, 1244–1247 (1993)
26. Leonhardt, U., Paul, H.: Realistic optical homodyne measurements and quasiprobability distributions. *Phys. Rev. A* **48**, 4598–4604 (1993)

27. D'Ariano, G.M., Macchiavello, C., Paris, M.G.: Optimized phase detection. *Phys. Lett. A* **198**(4), 286–294 (1995)
28. Pellonpää, J.-P.: Quantum tomography, phase-space observables and generalized Markov kernels. *J. Phys. A* **42**(46), 465303, 18 (2009)
29. Heinosaari, T., Mazzarella, L., Wolf, M.M.: Quantum tomography under prior information. *Comm. Math. Phys.* **318**(2), 355–374 (2013)

Chapter 19

Measurement Implementations

Following the introduction and study of a range of concrete operator measures representing quantum mechanical observables in earlier chapters (notably Chaps. 8, 14–17), we now apply the tools of measurement theory developed in Chap. 10 to illustrate the implementation of more or less realistic measurement schemes for typical observables.

In a measurement, a probe interacts with the object system after which the pointer observable is measured on the probe. In addition, if the object system is still available after the measurement interaction has ceased, one may also perform another measurement on it, taking into account its change of state due to the first measurement.

Our focus will be on the realisation of joint approximate measurements of non-commuting pairs of observables. We start in Sect. 19.1 with discussing a model of a joint position-momentum measurement due to Arthurs and Kelly. It combines approximate position and momentum measurements of the standard model into a single measurement. This allows us to investigate also in detail the influence that a measurement necessarily has on the object system and the mutual disturbance that two measurements exert on each other if applied simultaneously or in succession.

Most of the actual implementations of approximate joint measurements of incompatible quantities are measurements on light using beam splitters, phase shifters, and photodetection. Therefore, Sect. 19.2 gives a brief introduction of the relevant aspects of quantum optics. This allows us to use a balanced homodyne detection scheme to realise approximate quadrature measurements a single mode optical field, Sect. 19.3, and a double homodyne detection scheme to realise their approximate joint measurements, Sect. 19.4, as well as measurements of covariant phase observables, Sect. 19.5. Finally, we use (Sect. 19.6) Mach–Zehner interferometry to investigate aspects of the phenomenon of wave-particle duality.

19.1 Arthurs–Kelly Model

An approximate position measurement, described by the standard (von Neumann) model, followed by a sharp momentum measurement, results in a sequential joint measurement which realises a covariant phase space observable $\mathbf{G}^T : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ (see Chap. 10). The margins of $\mathbf{G}^T$ are

$$\mathbf{G}_1^T = \mu * \mathbf{Q}, \quad \mathbf{G}_2^T = \nu * \mathbf{P},$$

where the probability measures μ and ν have the densities $e_0^\lambda(q) = \lambda |\phi(-\lambda q)|^2$ and $f_0^\lambda(p) = \frac{1}{\lambda} |\hat{\phi}(-\frac{p}{\lambda})|^2$, which depend on the initial probe state $\phi \in L^2(\mathbb{R})$, $\|\phi\| = 1$, and the coupling parameter $\lambda > 0$, while the generating density operator $T = |\Pi \phi^{(\lambda)}\rangle \langle \Pi \phi^{(\lambda)}|$ with $\phi^{(\lambda)} = \sqrt{\lambda} \phi(\lambda q)$. This model constitutes an approximate joint measurement of position and momentum with measurement error product (see (15.28))

$$\Delta_2(\mathbf{Q}, \mathbf{G}_1^T) \Delta_2(\mathbf{P}, \mathbf{G}_2^T) = \sqrt{\mu[2]} \sqrt{\nu[2]} \geq \Delta(\mu) \Delta(\nu) \geq \frac{1}{2}.$$

More realistic approximate joint measurement models arise when both, the position and the momentum measurement, are approximate (standard) measurements, and the two measurements are performed either sequentially or simultaneously on the object system. Though both models define covariant phase space observables, it is instructive to study their details and compare the approximation errors in the two cases.

Thus, let us consider an object and two probe particles with associated Hilbert spaces $\mathcal{H}$, $\mathcal{H}_1$, $\mathcal{H}_2$, respectively, all given by $L^2(\mathbb{R})$ in the position representation. The fixed initial probe states are denoted ϕ_1 , ϕ_2 , and the two measurement couplings are given by the unitary maps $U_1 = \exp(-i\lambda Q \otimes P_1 \otimes I_2)$ and $U_2 = \exp(i\kappa P \otimes I_1 \otimes Q_2)$, here already extended so as to act on $\mathcal{H} \otimes \mathcal{H}_1 \otimes \mathcal{H}_2$. A natural choice of pointer observables¹ is $\mathbf{Z}_1 = \mathbf{Q}_1$ and $\mathbf{Z}_2 = \mathbf{P}_2$, and for the pointer functions we take $x \mapsto \lambda^{-1}x$ and $y \mapsto \kappa^{-1}y$. (We shall put generally $\hbar = 1$ in the following considerations but will allow the symbol $\hbar$ to appear in some formulas. Note that the coupling strength parameters κ , λ are dimensionless.)

If the two couplings are applied simultaneously, the coupling is given by

$$U = \exp(-i\lambda Q \otimes P_1 \otimes I_2 + i\kappa P \otimes I_1 \otimes Q_2). \quad (19.1)$$

This is the Arthurs–Kelly model [1] (see also [2–4]) for an approximate joint measurement of position and momentum.² In contrast, combining the two measurements sequentially (Sect. 10.3), with (say) the position measurement performed first, the

¹The coupling operator used by Arthurs and Kelly has P_2 in the place of Q_2 , so that in this case the appropriate pointer observable for the second probe is Q_2 .

²For the usual technical precautions concerning equations such as (19.1), see Footnote 1 in Chap. 15.

total coupling is

$$U' = U_2 U_1 = \exp(i\kappa P \otimes I_1 \otimes Q_2) \exp(-i\lambda Q \otimes P_1 \otimes I_2). \quad (19.2)$$

The two interactions are related with each other through

$$U = \exp\left(\frac{i\lambda\kappa}{2} I \otimes P_1 \otimes Q_2\right) U', \quad (19.3)$$

which is easily confirmed by a formal application of the Baker–Campbell–Hausdorff formula.³ This also shows that in the Arthurs–Kelly model the position and momentum measurement probes become intertwined via their interactions with the object, a feature which is not present in the sequential scheme.

The measured observable $\mathbf{G}$ is obtained, in both models, by application of the probability reproducibility condition, written here for product sets:

$$\langle \varphi | \mathbf{G}(X \times Y) \varphi \rangle = \langle U^{(m)} \Psi_0 | I \otimes \mathbf{Q}_1(\lambda X) \otimes \mathbf{P}_2(\kappa Y) U^{(m)} \Psi_0 \rangle, \quad (19.4)$$

where $X, Y \in \mathcal{B}(\mathbb{R})$ and the coupling $U^{(m)}$ is either U or U' . In both cases $\mathbf{G}$ is found to be a covariant phase space observable and the marginal observables, unsharp position and unsharp momentum, are easily computed.

In the sequential application of the two standard models, the first marginal observable $\mathbf{G}'_1$ is the observable measured first; in this case, the unsharp position $\mathbf{G}'_1 = \mu_1 * \mathbf{Q}$, with $\mu_1 = \mathbf{Q}_{1, \phi_1^{(\lambda)}}$. The second marginal observable $\mathbf{G}'_2$ is a ‘distorted’ version of the unsharp momentum $\nu_2 * \mathbf{P}$, that is, $\mathbf{G}'_2$ has the effects $\mathbf{G}'_2(Y) = \mathfrak{I}_1(\mathbb{R})^*(\nu_2 * \mathbf{P})(Y)$ with $\nu_2 = \mathbf{P}_{2, \phi_2^{(1/\kappa)}}$.

For any state ϱ , one now has

$$\begin{aligned} \text{tr}[\varrho \mathbf{G}'_2(Y)] &= \text{tr}[\varrho \mathfrak{I}_1(\mathbb{R})^*(\nu_2 * \mathbf{P})(Y)] \\ &= \text{tr}[\mathfrak{I}_1(\mathbb{R})(\varrho)(\nu_2 * \mathbf{P})(Y)] \\ &= (\nu_2 * \mathbf{P}_{\mathfrak{I}_1(\mathbb{R})(\varrho)})(Y) = (\nu_2 * (\nu_1 * \mathbf{P}_\varrho))(Y), \end{aligned}$$

which gives

$$\mathbf{G}'_1 = \mu_1 * \mathbf{Q}, \quad \mathbf{G}'_2 = \nu_1 * \nu_2 * \mathbf{P}, \quad (19.5)$$

where $\nu_1 = \mathbf{P}_{1, \phi_1^{(\lambda)}}$. According to Corollary 15.1, there is again a density operator $K(\lambda, \kappa)$ such that $\mathbf{G}'_i = \mathbf{G}_i^{K(\lambda, \kappa)}$ for $i = 1, 2$.

The measurement errors in this model are

$$\Delta_2(\mathbf{Q}, \mathbf{G}'_1) = \sqrt{\mu_1[2]} \geq \Delta(\mu_1), \quad \Delta_2(\mathbf{P}, \mathbf{G}'_2) = \sqrt{(\nu_1 * \nu_2)[2]} \geq \sqrt{\Delta(\nu_1)^2 + \Delta(\nu_2)^2}.$$

³A rigorous proof of this equation is obtained by adaptation of arguments from [5] along the lines of Exercises 6, 7 and 8 of Chap. 15. This also shows that (19.1) defines a unitary operator, without giving explicitly the spectral decomposition of the unique selfadjoint extension of the essentially selfadjoint operator $-\lambda Q \otimes P_1 \otimes I_2 + \kappa P \otimes I_1 \otimes Q_2$.

Though the errors in the two standard models can be chosen independently to be arbitrarily small, the relation $\Delta(\mu_1)\Delta(\nu_1) \geq \frac{1}{2}$ ensures that the sequential joint measurement fulfills the necessary trade-off relation.

In the Arthurs–Kelly model, the appearance of the additional coupling term $\exp\left(\frac{i\lambda\kappa}{2}I \otimes P_1 \otimes Q_2\right)$ in Eq. (19.3) causes further blurring in the approximate position and momentum distributions. It is again a straightforward calculation to determine the Cartesian marginal observables G_1 , G_2 of the observable G defined by this measurement scheme.

To determine the margin G_1 by evaluation of the probabilities $\langle \varphi | G_1(X) \varphi \rangle$, one needs to spell out the definitions of the operators, such as, for example,

$$\exp\left(\frac{i\lambda\kappa}{2}I \otimes P_1 \otimes Q_2\right) = I \otimes \int \exp\left(\frac{i\lambda\kappa}{2}zP_1\right) \otimes Q_2(dz),$$

and use the multiplicativity of spectral measures and the translation covariance of position to get

$$\begin{aligned} \langle \varphi | G_1(X) \varphi \rangle &= \int \int Q_\varphi(dx) \left\langle \phi_1 \left| Q_1\left(\lambda\left(X + \frac{\kappa}{2}z - x\right)\right) \phi_1 \right\rangle Q_{2,\phi_2}(dz) \right. \\ &= \int \left(Q_\varphi * Q_{1,\phi_1^{(\lambda)}} \right) \left(X + \frac{\kappa}{2}z \right) Q_{2,\phi_2}(dz) \\ &= \left(\left(Q_\varphi * Q_{1,\phi_1^{(\lambda)}} \right) * Q_{2,\Pi\phi_2^{(2/\kappa)}} \right) (X), \end{aligned}$$

so that

$$G_1 = (\tilde{\mu}_2 * \mu_1) * Q, \quad (19.6)$$

where $\tilde{\mu}_2 = Q_{2,\Pi\phi_2^{(2/\kappa)}}$. A similar calculation for G_2 gives

$$\langle \varphi | G_2(Y) \varphi \rangle = \left(P_\varphi * P_{2,\phi_2^{(1/\kappa)}} * P_{1,\phi_1^{(\lambda/2)}} \right) (Y),$$

and so

$$G_2 = (\tilde{\nu}_1 * \nu_2) * P, \quad (19.7)$$

where $\tilde{\nu}_1 = P_{1,\phi_1^{(\lambda/2)}}$.

The measurement errors

$$\Delta_2(Q, G_1) = \sqrt{(\tilde{\mu}_2 * \mu_1)[2]}, \quad \Delta_2(P, G_2) = \sqrt{(\tilde{\nu}_1 * \nu_2)[2]} \quad (19.8)$$

are again readily available. Below we shall give a more detailed analysis of the structure of the uncertainty product.

By virtue of Corollary 15.1 there is a state K such that $G_i = G_i^K$ for $i = 1, 2$. We determine next the explicit form of this covariant phase space observable, that is, the operator K , as well as the associated instrument using the Arthurs–Kelly interaction.

By (19.3), the final state $\Psi = U\Psi_0$ is given by

$$\Psi(q, \xi_1, \xi_2) = \varphi(q + \kappa\xi_2) \phi_1(\xi_1 - \lambda q - \frac{\lambda\kappa}{2}\xi_2) \phi_2(\xi_2). \quad (19.9)$$

Since the second pointer observable is momentum, it is useful to Fourier transform this wave function with respect to the last variable:

$$\widehat{\Psi}(q, \xi_1, \pi_2) = \frac{1}{\sqrt{2\pi}} \int e^{-i\xi_2\pi_2} \Psi(q, \xi_1, \xi_2) d\xi_2. \quad (19.10)$$

We determine the measured observable $\mathbf{G}$ using (19.4) and the instrument $\mathfrak{I}$ via

$$\text{tr}[\mathfrak{I}(X \times Y)(P[\varphi])A] = \langle \Psi | A \otimes \mathbf{Q}_1(\lambda X) \otimes \mathbf{P}_2(\mu Y) \Psi \rangle \quad (19.11)$$

which is to hold for all initial object states φ , for all outcome sets X, Y , and for all bounded selfadjoint operators A acting in $\mathcal{H}$. One obtains:

$$\mathbf{G}(X \times Y) = \int_{X \times Y} K_{qp}^* K_{qp} dq dp = \frac{1}{2\pi} \int_{X \times Y} T_{qp} dq dp \quad (19.12)$$

$$\mathfrak{I}(X \times Y)(P[\varphi]) = \int_{X \times Y} K_{qp} P[\varphi] K_{qp}^* dq dp \quad (19.13)$$

$$K_{qp}(x, x') = \frac{1}{\sqrt{2\pi}} e^{ip(x-x')} \phi_1^{(\lambda)}(q - \frac{1}{2}(x+x')) \phi_2^{(\kappa)}(x-x') \quad (19.14)$$

It thus follows that K_{qp} is an integral operator with kernel given by Eq. (19.14).

At this stage we formulate some conditions to ensure that (19.12), (19.14) and the subsequent operator relations are well-defined. We shall assume that the functions $\phi_1(\xi_1), \phi_2(\xi_2)$, as well as their Fourier transforms, are continuous and bounded, and that they all have vanishing first and finite second moments. If seen in the position representation, the operators K_{qp} are integral operators with a kernel which is an element of $L^2(\mathbb{R}^2)$. This is to say that the K_{qp} are Hilbert-Schmidt operators and consequently, that the operators T_{qp} in (19.12) are positive trace class operators. Since the normalisation of the observable $\mathbf{G}$ is guaranteed by virtue of (19.4), one immediately has $\text{tr}[T_{qp}] = 1$. A little further analysis of the K_{qp} shows that they form a phase-space translation covariant family:

$$K_{qp} = W(q, p) K_{00} W(q, p)^{-1} \quad (19.15)$$

It is evident that $\mathbf{G}$ inherits this covariance and therefore is a phase space observable; one obtains

$$\begin{aligned} K_{00}\varphi(x) &= \langle \psi_x | \varphi \rangle, & \overline{\psi_x(x')} &= K_{00}(x', x) \\ T_{00} &= 2\pi \int |\psi_x\rangle \langle \psi_x| dx. \end{aligned} \quad (19.16)$$

The operator T_{00} can be a rank one projection only when $K_{00}(x, x')$ factorises, which is possible only for certain Gaussian functions ϕ_1, ϕ_2 [2].

Trade-Off Relation for Unsharpness and Inaccuracies

The margins of the phase space observable (19.12) are unsharp position and momentum observables (19.6) and (19.7),

$$\mathbf{G}_1 = \mathbf{Q}_{\Pi T_{00} \Pi} * \mathbf{Q} = \tilde{\mu}_2 * \mu_1 * \mathbf{Q}, \quad \mathbf{G}_2 = \mathbf{P}_{\Pi T_{00} \Pi} * \mathbf{P} = \tilde{\nu}_1 * \nu_2 * \mathbf{P}.$$

We recall that the probability measures $\tilde{\mu}_2 * \mu_1$ and $\tilde{\nu}_1 * \nu_2$ reflect both, the *intrinsic unsharpness* of the observables $\mathbf{G}_1$ and $\mathbf{G}_2$ as well as the (*in*-)accuracy with which they approximate the sharp position and momentum. We will refer to them as inaccuracy measures.

The explicit forms of the inaccuracy measures are

$$\mu = \mathbf{Q}_{\Pi T_{00} \Pi} = \tilde{\mu}_2 * \mu_1 = \mathbf{Q}_{2, \Pi \phi_2^{(2/\kappa)}} * \mathbf{Q}_{1, \phi_1^{(\lambda)}} \quad (19.17)$$

$$\nu = \mathbf{P}_{\Pi T_{00} \Pi} = \tilde{\nu}_1 * \nu_2 = \mathbf{P}_{1, \phi_1^{(\lambda/2)}} * \mathbf{P}_{2, \phi_2^{(1/\kappa)}} \quad (19.18)$$

and the measurement errors (19.8) satisfy the inequality

$$\begin{aligned} \Delta_2(\mathbf{Q}, \mathbf{G}_1) \Delta_2(\mathbf{P}, \mathbf{G}_2) &= \sqrt{(\tilde{\mu}_2 * \mu_1)[2]} \sqrt{(\tilde{\nu}_1 * \nu_2)[2]} \\ &\geq \Delta(\tilde{\mu}_2 * \mu_1) \Delta(\tilde{\nu}_1 * \nu_2) \geq \frac{1}{2} \hbar, \end{aligned} \quad (19.19)$$

where the first inequality is an equality exactly when the initial probe states ϕ_1 and ϕ_2 are such that $\mathbf{Q}_{\phi_i} = \mathbf{P}_{\phi_i} = 0$ for $i = 1, 2$. The second inequality will be saturated if, in addition, the probe states are Gaussians. On the other hand, by Theorem 15.3, the lower bound is reached exactly when T_{00} is the one-dimensional projection operator defined by the ground state of a harmonic oscillator.

The measures μ_1 and ν_2 are the inaccuracy measures of the original single measurements, which can be recovered from the present joint measurement model by switching off one ($\kappa = 0$) or the other ($\lambda = 0$) coupling. One thus obtains a full specification of all features of a phase space observable in operational terms. As indicated by the convolution structure, the original undisturbed inaccuracies are each changed due to the presence of the other device. In other words the simultaneous application of the measuring devices for $\mu_1 * \mathbf{Q}$ and $\nu_2 * \mathbf{P}$ is a joint measurement of smeared versions $\mathbf{G}_1$ and $\mathbf{G}_2$ of these observables.

It is instructive to analyse further the explicit form of the variances of μ and ν . As we shall see, this gives a more refined account, in the context of the Arthurs–Kelly model, of the way in which quantum mechanics prohibits accurate joint measurements of position and momentum. There are two factors at play: the mutual disturbance of the two measurements when performed simultaneously, and the quantum nature of the probes.

The mutual influence of the two measurements being carried out simultaneously becomes strikingly manifest in the variances of μ and ν :

$$\begin{aligned}\Delta_2(\mathbf{Q}, \mathbf{G}_1)^2 &\geq \Delta(\mu)^2 = \Delta_2(\tilde{\mu}_2 * \mu_1)^2 = \Delta_2(\tilde{\mu}_2)^2 + \Delta_2(\mu_1)^2 \\ &= \frac{1}{\lambda^2} \text{Var}(\mathbf{Q}_1, \phi_1) + \frac{\kappa^2}{4} \text{Var}(\mathbf{Q}_2, \phi_2),\end{aligned}\quad (19.20)$$

$$\begin{aligned}\Delta_2(\mathbf{P}, \mathbf{G}_2)^2 &\geq \Delta(\nu)^2 = \Delta_2(\tilde{\nu}_1 * \nu_2)^2 = \Delta_2(\tilde{\nu}_1)^2 + \Delta_2(\nu_2)^2 \\ &= \frac{1}{\kappa^2} \text{Var}(\mathbf{P}_2, \phi_2) + \frac{\lambda^2}{4} \text{Var}(\mathbf{P}_1, \phi_1).\end{aligned}\quad (19.21)$$

There are two ways to make the ‘undisturbed’ variances (the first terms) small: either by choosing large coupling constants or by preparing ‘pointer’ states having sharply peaked distributions $|\phi_1|^2$, $|\hat{\phi}_2|^2$. Both options have the same consequence: they produce large contributions to the other quantity’s unsharpness, by increasing the second terms. Thus there is no way of getting both quantities $\text{Var}(\mu)$ and $\text{Var}(\nu)$ small in one and the same experiment. In fact it is apriori clear that they fulfil the uncertainty relation $\text{Var}(\mu)\text{Var}(\nu) \geq \hbar^2/4$ since the probability measures μ and ν are defined by a common state. But we are now in a position to see the dynamical mechanism at work that ensures the phase space measurement inaccuracies to be in accord with this relation.

We evaluate the product of the variances,

$$\text{Var}(\mu) \cdot \text{Var}(\nu) = \mathcal{Q} + \mathcal{D}, \quad (19.22)$$

$$\mathcal{Q} = \frac{1}{4} \text{Var}(\mathbf{Q}_1, \phi_1) \text{Var}(\mathbf{P}_1, \phi_1) + \frac{1}{4} \text{Var}(\mathbf{Q}_2, \phi_2) \text{Var}(\mathbf{P}_2, \phi_2), \quad (19.23)$$

$$\mathcal{D} = \frac{1}{\lambda^2 \kappa^2} \text{Var}(\mathbf{Q}_1, \phi_1) \text{Var}(\mathbf{P}_2, \phi_2) + \frac{\lambda^2 \kappa^2}{16} \text{Var}(\mathbf{Q}_2, \phi_2) \text{Var}(\mathbf{P}_1, \phi_1). \quad (19.24)$$

Making use of the uncertainty relations $\text{Var}(\mathbf{Q}_k, \phi_k) \text{Var}(\mathbf{P}_k, \phi_k) \geq \hbar^2/4$ for the two probe systems, we find that both terms $\mathcal{Q}, \mathcal{D}$ can be estimated from below. Putting $\xi = 16 \text{Var}(\mathbf{Q}_1, \phi_1) \text{Var}(\mathbf{P}_2, \phi_2)/(\lambda \kappa \hbar)^2$, we obtain:

$$\mathcal{Q} \geq \frac{1}{4} \left(\frac{\hbar^2}{4} + \frac{\hbar^2}{4} \right) = \frac{\hbar^2}{8}, \quad (19.25)$$

$$\mathcal{D} \geq \frac{\hbar^2}{16} \left(\xi + \frac{1}{\xi} \right) \geq \frac{\hbar^2}{8}. \quad (19.26)$$

This gives

$$\Delta_2(\mathbf{Q}, \mathbf{G}_1) \cdot \Delta_2(\mathbf{P}, \mathbf{G}_2) \geq \text{Var}(\mu) \cdot \text{Var}(\nu) = \mathcal{Q} + \mathcal{D} \geq \frac{\hbar^2}{4}, \quad (19.27)$$

confirming the error bound appropriate to a covariant phase space observable.

It is remarkable that either one of the terms $\mathcal{Q}$ and $\mathcal{D}$ suffices to provide an absolute lower bound for the error product. Hence there are two sources of inaccuracy that give rise to an uncertainty relation.

First, neglecting $\mathcal{D}$ it would be simply the validity of the preparation uncertainty relations for the two parts of the apparatus which forbids making the term $\mathcal{Q}$ arbitrarily small. This is in the spirit of Bohr's argument according to which it is the quantum nature of part of the measuring device that makes it impossible to escape the uncertainty relation. Note that the two terms occurring in $\mathcal{Q}$ each refer to one of the probe systems, and they contribute independently to the lower bound for $\mathcal{Q}$; furthermore no coupling parameters appear in $\mathcal{Q}$. There is no trace of a mutual influence between the two measurements being carried out simultaneously.

Second, neglecting the term $\mathcal{Q}$, one would still be left with the two contributions collected in $\mathcal{D}$, the combination of which has again a lower bound. The terms in $\mathcal{D}$ are products of variances and coupling terms associated with the two probe systems, showing that $\mathcal{D}$ reflects the mutual disturbance of the two measurements. This is in accord with Heisenberg's illustrations of the uncertainty relation. For example if a particle is measured so as to have a rather well-defined momentum, then a subsequent measurement of position by means of a slit influences the effect of the preceding momentum measurement to the extent required by the uncertainty relation.

Finally we see that the nature of the measurement 'inaccuracy', or unsharpness, is determined by the preparations of the apparatus. Insofar as the pointer observables are indeterminate and not merely subjectively unknown, this interpretation applies to the measurement uncertainties as well: each individual measurement outcome is intrinsically unsharp, reflecting thereby a genuine quantum noise inherent in the measurement process, so that the inequality (19.27) demonstrates quantum indeterminacy as the source of the necessary inaccuracy in joint approximate measurements of position and momentum. This interpretation will be further substantiated by analysing the way the measurement affects the object's states. To this end we investigate the instrument (19.13).

Repeatability Features

The minimal measurement uncertainty in this model is achieved exactly when $T_{00} = |\psi\rangle\langle\psi|$ where $\psi(x) = h_0(x) = \pi^{-1/4}e^{-x^2/2}$ is the ground state of the oscillator. In this case the instrument (19.13) assumes a particularly simple form

$$\mathcal{I}(X \times Y)(\varrho) = \frac{1}{2\pi} \int_{X \times Y} \langle \psi_{qp} | \varrho \psi_{qp} \rangle P[\psi_{qp}] dq dp. \quad (19.28)$$

This is a mixture of coherent states which depends on the initial state ϱ only via the weights. Consider 'point-like' readings, that is, intervals $X = [q_o - \delta/2, q_o + \delta/2]$, $Y = [p_o - \gamma/2, p_o + \gamma/2]$ of lengths small compared to the widths of the position and momentum distributions of $\psi_{qp} = W(q, p)\psi$. Within sufficiently small interval lengths, the continuous function $(q, p) \mapsto \langle \psi_{qp} | \varrho \psi_{qp} \rangle$ will be approximately constant. For such intervals one can approximate the integration to obtain

$$\mathfrak{I}(X \times Y)(\varrho) \simeq P[\psi_{q_o, p_o}] \left\{ \psi_{q_o, p_o} \mid \varrho \psi_{q_o, p_o} \right\} \frac{\delta\gamma}{2\pi}. \quad (19.29)$$

After normalisation no mark is left of the original object state. On the contrary, the characteristics of the measurement scheme are ideally imprinted into the system: its final state is localised at the ‘point’ (q_o, p_o) indicated by the reading, and the variances are exactly given by the measurement inaccuracies (19.20) and (19.21). This means that the measurement is quasi-preparatory, in the sense that the properties of the final object state correlate well to the value of the outcome obtained.

To conclude, our analysis of the Arthurs–Kelly model shows how the measurement uncertainty relation for position and momentum is enforced by two factors: the quantum nature of the probe systems as it manifests itself in their preparation uncertainty relations, and the unavoidable mutual disturbance of the probe system resulting from their simultaneous or sequential couplings to the object system. Furthermore, in the case of a quasi-preparatory phase space measurement, the preparation uncertainty of position and momentum in the object’s final state is directly related to the measurement uncertainty relation, reflecting the close connection between these two facets of the uncertainty principle that was already exhibited in the general proof of the position-momentum error relation of Chap. 15. This result corroborates Heisenberg’s informal model derivation of the measurement uncertainty relation in his 1927 paper [6], which was based on an intuitive notion of a quasi-preparatory measurement and an identification of measurement errors with the respective preparation uncertainties in his model.

19.2 Photon Detection, Phase Shifters and Beam Splitters

In this section we introduce classical and quantum descriptions of a single-mode optical field. We use the coarsest nontrivial approximation for a propagating electromagnetic field, namely, a monochromatic plane wave. This approximation is sufficient for the applications of this chapter since it allows us to present all essential features of ideal laser light.

Let us consider a classical electromagnetic field in a (e.g. cubic) cavity of finite volume $\mathcal{V}$. The Maxwell equations with the Coulomb gauge condition lead to the wave equation

$$\nabla^2 \mathbf{A} = \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} \quad (19.30)$$

for the vector potential $\mathbf{A}$ which depends on position $\mathbf{r} = (x, y, z) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ and time t , c being the speed of light. The electric and magnetic field vectors are $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$ and $\mathbf{H} = \mu_0^{-1} \nabla \times \mathbf{A}$ where μ_0 is the magnetic permeability of free space. The monochromatic plane wave is a solution of (19.30),

$$\mathbf{A}(\mathbf{r}, t) = \sqrt{\frac{2\mathcal{E}}{\epsilon_0 \mathcal{V} \omega^2}} \cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \phi) \mathbf{u}. \quad (19.31)$$

Here $\mathbf{k}$ is the wavevector, $\omega = c|\mathbf{k}|$, $\mathbf{u}$ is the unit polarisation vector ($\mathbf{u} \perp \mathbf{k}$), $\mathcal{E}$ is the cycle-averaged energy content of the plane wave in the volume $\mathcal{V}$, ϵ_0 is the free space permittivity, and ϕ is an (absolute) phase of the plane wave. Assuming the periodic boundary conditions all other solutions of (19.30) can be expressed as linear combinations or series of these single mode plane waves.

Suppose that the plane wave is moving in the positive direction of the z -axis so that $\mathbf{k} = (\omega/c)\hat{\mathbf{k}}$ implying $\mathbf{k} \cdot \mathbf{r} = (\omega/c)z$ and we may choose $\mathbf{u} = \hat{\mathbf{i}}$. The energy density is then

$$W(\mathbf{r}, t) = \frac{\epsilon_0}{2} |\mathbf{E}(\mathbf{r}, t)|^2 + \frac{\mu_0}{2} |\mathbf{H}(\mathbf{r}, t)|^2 = \frac{2\mathcal{E}}{\mathcal{V}} \sin^2((\omega/c)z - \omega t + \phi),$$

the Poynting vector

$$\mathbf{E}(\mathbf{r}, t) \times \mathbf{H}(\mathbf{r}, t) = c W(\mathbf{r}, t) \hat{\mathbf{k}},$$

and the cycle-averaged energy content is

$$\int_{\mathcal{V}} \left[\frac{\omega}{2\pi} \int_0^{2\pi/\omega} W(\mathbf{r}, t) dt \right] dx dy dz = \mathcal{E}.$$

Suppose that there is a photodetector at the origin and the active region of the detector is perpendicular to the z -axis. The intensity of light streaming into the detector at time t is

$$c W(0, t) = \frac{2c\mathcal{E}}{\mathcal{V}} \sin^2(-\omega t + \phi).$$

In all experiments measuring light intensity the time of measurement is long compared to the period of a single oscillation. This means that one does not measure $c W(0, t)$ but rather its temporal average over a time $2\pi/\omega$. Thus, the optimally measured intensity is

$$\mathcal{I} = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} c W(0, t) dt = \frac{c\mathcal{E}}{\mathcal{V}}.$$

If the efficiency coefficient of the detector is $\epsilon \in [0, 1]$, then $\epsilon\mathcal{I}$ is the actually measured intensity. The measured energy during a time interval $[0, T]$ is

$$\epsilon \mathcal{A} \int_0^T c W(0, t) dt = \epsilon \mathcal{A} \mathcal{I} \left[T + \frac{\sin(2\phi - 2\omega T) - \sin(2\phi)}{2\omega} \right] \sim \epsilon \mathcal{A} T \mathcal{I}$$

when $T \gg 2\pi/\omega$ and $\mathcal{A}$ is the area of the active region of the detector. This justifies the above approximation for intensity when a measurement period is long compared to $2\pi/\omega$. Next we consider the quantum plane wave.

First we recall the quantum mechanical treatment of the harmonic oscillator (of unit mass, $m = 1$). Let $\mathcal{H} = L^2(\mathbb{R})$ be the position representation space and

$(Q\psi)(x) = x\psi(x)$ and $(P\psi)(x) = -i\hbar\psi'(x)$ with their spectral measures $\mathbf{Q}$ and $\mathbf{P}$, respectively. Denote the n th Hermite function h_n by $|n\rangle$ so that the energy operator $H = \frac{1}{2}P^2 + \frac{\omega^2}{2}Q^2$ can be written in the form

$$H = \hbar\omega \left(N + \frac{1}{2}I \right)$$

where the number operator $N = a^*a = \sum_{n=0}^{\infty} n |n\rangle\langle n|$ and the lowering operator $a = \sum_{n=0}^{\infty} \sqrt{n+1} |n\rangle\langle n+1|$ is the (closure of the) operator $(\sqrt{\omega}Q + i\sqrt{\omega}^{-1}P)/\sqrt{2\hbar}$.

The single-mode electromagnetic vector potential operator [7, Eq. (6.105)] in a cavity at $(\mathbf{r}, t)$ is a selfadjoint (rotated) quadrature operator

$$\begin{aligned} \hat{\mathbf{A}}(\mathbf{r}, t) &= \sqrt{\frac{\hbar}{2\epsilon_0\mathcal{V}\omega}} (ae^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} + a^*e^{-i\mathbf{k}\cdot\mathbf{r}+i\omega t}) \mathbf{u} \\ &= \frac{\mathbf{u}}{\sqrt{\epsilon_0\mathcal{V}}} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)N} Q e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)N} \end{aligned} \quad (19.32)$$

where we have used the equation $e^{-i\theta N} a e^{i\theta N} = e^{i\theta} a$ for all $\theta \in \mathbb{R}$. Also, it follows that the energy operator generates time shifts,

$$e^{-iHt'/\hbar} \hat{\mathbf{A}}(\mathbf{r}, t) e^{iHt'/\hbar} = \hat{\mathbf{A}}(\mathbf{r}, t + t').$$

(see, Sect. 17.3).

Since laser light is quasimonochromatic, it suffices to treat a laser as a single-mode optical system in this context. *Coherent states*

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad \alpha \in \mathbb{C},$$

are known to describe ideal laser light in quantum optics [8, p. 232]. We note that $a|\alpha\rangle = \alpha|\alpha\rangle$ and $\langle\alpha|N|\alpha\rangle = |\alpha|^2$ for all $\alpha \in \mathbb{C}$. The mean value of the operator $\hat{\mathbf{A}}(\mathbf{r}, t)$ in state $|\alpha\rangle$ is

$$\begin{aligned} \langle\alpha|\hat{\mathbf{A}}(\mathbf{r}, t)|\alpha\rangle &= \sqrt{\frac{\hbar}{2\epsilon_0\mathcal{V}\omega}} (\alpha e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} + \bar{\alpha} e^{-i\mathbf{k}\cdot\mathbf{r}+i\omega t}) \mathbf{u} \\ &= \sqrt{\frac{2\tilde{\mathcal{E}}}{\epsilon_0\mathcal{V}\omega^2}} \cos(\mathbf{k}\cdot\mathbf{r} - \omega t + \arg \alpha) \mathbf{u} \end{aligned} \quad (19.33)$$

where

$$\tilde{\mathcal{E}} = \langle\alpha|H|\alpha\rangle - \frac{1}{2}\hbar\omega = \hbar\omega\langle\alpha|N|\alpha\rangle = \hbar\omega|\alpha|^2$$

is the mean field energy (total energy minus vacuum energy). Hence,

$$\langle \alpha | N | \alpha \rangle = |\alpha|^2 = \frac{\tilde{\mathcal{E}}}{\hbar\omega}. \quad (19.34)$$

Comparing Eqs. (19.31) and (19.33) one sees that the classical monochromatic plane wave can be used to describe a quantum laser light in the coarsest nontrivial semiclassical approximation. Note that the variance of $\hat{A}(\mathbf{r}, t)$ in $|\alpha\rangle$ tends to 0 in the high amplitude limit $|\alpha| \rightarrow \infty$ (i.e. $\tilde{\mathcal{E}} \rightarrow \infty$).

As one knows from the theory of photodetection, a photon detector counts the number of photons in some detection time interval $[0, T]$. Hence, in a semiclassical theory the energy of light streaming into a photodetector, $\epsilon AT\mathcal{I}$, is proportional to the measured number of photons n in the detection time interval $[0, T]$. This means that one can write $\mathcal{I} = \gamma n/T$ where $\gamma > 0$ is some constant depending on the structure of the photodetector. Since the classical intensity $\mathcal{I} = c\mathcal{E}/\mathcal{V}$ it follows that $\mathcal{E} = (\mathcal{V}/c)\mathcal{I} = (\mathcal{V}/c)(\gamma/T)n$ and, comparing this to Eq. (19.34), one can define the semiclassical photon ‘number’ $\mathcal{N} = \mathcal{E}/(\hbar\omega)$ which is proportional to n/T . This suggests that an *ideal photon detector* serves as a measurement of the photon number observable $\mathbf{N} : n \mapsto |n\rangle\langle n|$. If the counting efficiency of the detector is ϵ , then the detector measures, instead, the *unsharp photon number*

$$n \mapsto E_n^\epsilon = \sum_{m=n}^{\infty} \binom{m}{n} \epsilon^n (1-\epsilon)^{m-n} |m\rangle\langle m|, \quad j = 1, 2, \quad (19.35)$$

whose first moment operator is ϵN , corresponding to the classical value $\epsilon\mathcal{N}$, see, e.g. [9, pp. 79–83] or [10, pp. 177–180]. Note that, for any (initial) number state $|m\rangle$ and $n \leq m$,

$$\langle m | E_n^\epsilon | m \rangle = \binom{m}{n} \epsilon^n (1-\epsilon)^{m-n}$$

is the probability of detecting n photons when ϵ is the probability for one photon. This probability distribution is called *binomial*.

Now Eq. (19.31) gets the form

$$\begin{aligned} A(\mathbf{r}, t) &= \sqrt{\frac{2\hbar}{\epsilon_0 \mathcal{V} \omega}} \operatorname{Re} \left(e^{i\phi} \sqrt{\mathcal{N}} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} \right) \mathbf{u} \\ &= \sqrt{\frac{\hbar}{2\epsilon_0 \mathcal{V} \omega}} \left(e^{i\phi} \sqrt{\mathcal{N}} e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t} + e^{-i\phi} \sqrt{\mathcal{N}} e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega t} \right) \mathbf{u}. \end{aligned} \quad (19.36)$$

Comparing this equation to the operator $\hat{A}(\mathbf{r}, t)$ of Eq. (19.32) one sees that in the quantisation of the electromagnetic single-mode field one substitutes the lowering operator a for $e^{i\phi} \sqrt{\mathcal{N}}$. Moreover, since $\langle \alpha | a | \alpha \rangle = \alpha = e^{i \arg \alpha} |\alpha|$, it is convenient

to call $|\alpha|$ and $\arg \alpha$ the *intensity* or *energy parameter* and *phase parameter* of a coherent state $|\alpha\rangle$, respectively.

The classical (δ -)phase shifter shifts the phase of $\mathbf{A}(\mathbf{r}, t)$ by the amount $\delta \in [0, 2\pi)$, that is, after the phase shifter one gets the plane wave

$$\sqrt{\frac{2\hbar}{\epsilon_0 \mathcal{V} \omega}} \operatorname{Re} \left(e^{i(\phi+\delta)} \sqrt{\mathcal{N}} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} \right) \mathbf{u}.$$

Hence, in quantum theory, a (δ -)phase shifter is described by the unitary operator $e^{i\delta N}$ for which

$$e^{-i\delta N} a e^{i\delta N} = e^{i\delta} a, \quad e^{i\delta N} |\alpha\rangle = |\alpha e^{i\delta}\rangle.$$

Remark 19.1 Writing the operator a in the form $a = V_\psi \sqrt{N}$, where

$$V_\psi = \sum_{n=0}^{\infty} |n\rangle \langle n+1| + |\psi\rangle \langle 0|, \quad \psi \in \mathcal{H},$$

Dirac [11] assumed that the operator V_ψ could be expressed as $e^{i\Phi}$, where Φ is a selfadjoint phase operator. (Actually, Dirac used a decomposition $a = \sqrt{N} e^{-i\Phi}$, but our notation does not change Dirac's basic idea.) This would mean that V_ψ is a unitary operator, which is not true since $\langle 0 | V_\psi^* V_\psi | 0 \rangle = \|\psi\|^2 = 1$ only when ψ is a unit vector, whereas $\langle k | V_\psi V_\psi^* | k \rangle = 1 + |\langle k | \psi \rangle|^2 = 1$ for all k only when ψ is the null vector. Hence, there is no quantum phase operator in Dirac's sense and one has to proceed in another way. As we have seen in Chap. 16, to overcome this difficulty, it is convenient to define phase observables as phase shift covariant observables. None of them is projection valued. $\triangleleft$

In photodetection one measures the light intensity $\epsilon \mathcal{I}$ (or a smeared number observable in quantum theory) which does not depend on the absolute phase ϕ . Therefore, using a single photodetector one cannot measure the absolute phase of an incoming plane wave. To measure ϕ , one needs a reference field which is combined with the signal in a beam splitter. The resulting plane waves are then measured with two photodetectors.

A beam splitter is a device that splits a beam of light in two parts. It can also be used to combine two light beams. Suppose that two light beams 1 and 2 (here, plane waves as before) of the same ω but possibly different semiclassical photon numbers $\mathcal{N}_1, \mathcal{N}_2$ and absolute phases ϕ_1, ϕ_2 are mixed by a lossless 50 : 50 beam splitter at the origin $\mathbf{r} = 0$. A beam splitter that consists of a glass plate with a reflective dielectric coating on one side gives a phase shift of 0 or π , depending on the side from which it is incident. Transmitted waves have no phase shift. Reflected waves entering from the reflective side are phase-shifted by π , whereas reflected waves entering from the glass side have no phase shift.

Assume that the first resultant wave 3 after the shifter and beam splitter has a polarisation vector $\mathbf{u}$, a wavevector $\mathbf{k}$ with $|\mathbf{k}| = \omega/c$, a semiclassical photon number

$\mathcal{N}_3$, and a phase ϕ_3 . The reflected wave 2 entering from the reflective side is phase-shifted by π . By the superposition principle and Eq. (19.36), the first resultant wave is of the form

$$\begin{aligned} & \sqrt{\frac{2\hbar}{\epsilon_0 \mathcal{V} \omega}} \operatorname{Re} \left(\sqrt{\mathcal{N}_3} e^{i\phi_3} e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t} \right) \mathbf{u} \\ &= \sqrt{\frac{2\hbar}{\epsilon_0 \mathcal{V} \omega}} \operatorname{Re} \left[\left(2^{-1/2} \sqrt{\mathcal{N}_1} e^{i\phi_1} + 2^{-1/2} \sqrt{\mathcal{N}_2} e^{i(\phi_2 + \pi)} \right) e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t} \right] \mathbf{u}, \end{aligned}$$

that is,

$$\sqrt{\mathcal{N}_3} e^{i\phi_3} = 2^{-1/2} \sqrt{\mathcal{N}_1} e^{i\phi_1} - 2^{-1/2} \sqrt{\mathcal{N}_2} e^{i\phi_2}. \quad (19.37)$$

and

$$\begin{aligned} \mathcal{N}_3 &= \left| 2^{-1/2} \sqrt{\mathcal{N}_1} e^{i\phi_1} - 2^{-1/2} \sqrt{\mathcal{N}_2} e^{i\phi_2} \right|^2 \\ &= \frac{\mathcal{N}_1}{2} + \frac{\mathcal{N}_2}{2} - \sqrt{\mathcal{N}_1 \mathcal{N}_2} \cos(\phi_1 - \phi_2). \end{aligned} \quad (19.38)$$

Similarly, for the second resultant wave 4, one has

$$\sqrt{\mathcal{N}_4} e^{i\phi_4} = 2^{-1/2} \sqrt{\mathcal{N}_1} e^{i\phi_1} + 2^{-1/2} \sqrt{\mathcal{N}_2} e^{i\phi_2}$$

so that

$$\mathcal{N}_4 = \frac{\mathcal{N}_1}{2} + \frac{\mathcal{N}_2}{2} + \sqrt{\mathcal{N}_1 \mathcal{N}_2} \cos(\phi_1 - \phi_2). \quad (19.39)$$

In quantum theory, the above beam splitter system is described as follows. Associate the Hilbert space $\mathcal{H}_i = L^2(\mathbb{R})$, $i = 1, 2, 3, 4$, with both of the incoming plane waves (input modes 1 and 2) and with both of the resultant beams (output modes 3 and 4). Hence, the Hilbert space for incoming waves is $\mathcal{H}_{\text{in}} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and for outgoing waves $\mathcal{H}_{\text{out}} = \mathcal{H}_3 \otimes \mathcal{H}_4$. For input modes, we denote $a_1 = a \otimes I$ and $a_2 = I \otimes a$, and for output modes, $a_3 = a \otimes I$ and $a_4 = I \otimes a$. From the classical treatment above (put $\alpha_i = \sqrt{\mathcal{N}_i} e^{i\phi_i}$, $i = 1, 2, 3, 4$, in Eqs. (19.37) and (19.39)), one sees that our quantum *beam splitter* with two phase shifters is described by a unitary operator $U_{\text{BS}} : \mathcal{H}_{\text{in}} \rightarrow \mathcal{H}_{\text{out}}$ defined by its action on coherent states:

$$U_{\text{BS}}(|\alpha_1, \alpha_2\rangle) = \left| \frac{1}{\sqrt{2}}(\alpha_1 - \alpha_2), \frac{1}{\sqrt{2}}(\alpha_1 + \alpha_2) \right\rangle = |\alpha_3, \alpha_4\rangle \quad (19.40)$$

where $\alpha_1, \alpha_2 \in \mathbb{C}$ and, for instance, $|\alpha_1, \alpha_2\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle$. Since the linear span of the two-mode coherent states $|\alpha, \beta\rangle$, $\alpha, \beta \in \mathbb{C}$, is dense in $\mathcal{H} \otimes \mathcal{H}$, the linear extension of (19.40) indeed defines a unitary operator $U_{\text{BS}} : \mathcal{H}_{\text{in}} \rightarrow \mathcal{H}_{\text{out}}$. In this chapter, we will model a quantum beam splitter by the above unitary operator U_{BS} .

It is an easy exercise to show that, for the two-mode number states $|10\rangle = |1\rangle \otimes |0\rangle$ and $|01\rangle = |0\rangle \otimes |1\rangle$, one gets $U_{\text{BS}}(|10\rangle) = 2^{-1/2} |10\rangle + 2^{-1/2} |01\rangle$ and $U_{\text{BS}}(|01\rangle) = -2^{-1/2} |10\rangle + 2^{-1/2} |01\rangle$. Hence, the beam splitter acts as a unitary matrix

$$\tilde{U}_{\text{BS}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (19.41)$$

in the two dimensional Hilbert space with the orthonormal basis $\{|10\rangle, |01\rangle\}$.

Remark 19.2 The mapping U_{BS} of Eq. (19.40) can easily be adjusted to more realistic beam splitters which are neither lossless nor 50:50, or may have phase shifts of magnitudes other than π ; see, for instance, [9, 10, 12]. The possible losses in a beam splitter can be taken into account by a non-ideal photon detector. Ignoring losses, we consider here only beam splitter interactions of a form similar to (19.40). $\triangleleft$

19.3 Balanced Homodyne Detection and Quadrature Observables

The balanced homodyne detection is a basic measurement scheme in many quantum optical applications, including continuous variable quantum tomography as well as continuous variable quantum teleportation. In this section we demonstrate that such a detection scheme can be used to measure (smeared) rotated quadrature observables. The results of this section are largely from [13].

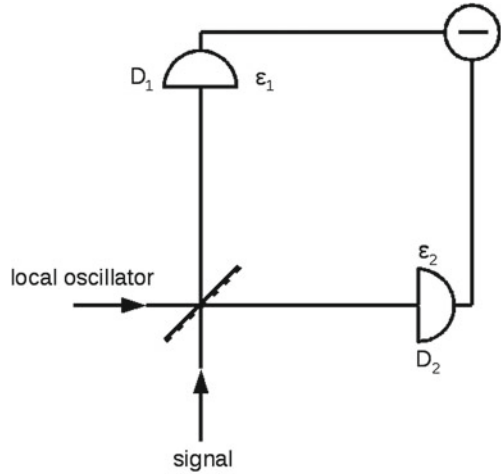
A balanced homodyne detector consists of a beam splitter (possibly with the additional phase shifters), two photodetectors D_1 and D_2 with the efficiencies ϵ_1 and ϵ_2 , and one registers a scaled difference of counts as the result. Classically, this (unscaled) difference is simply $\epsilon_1 \mathcal{N}_3 - \epsilon_2 \mathcal{N}_4$.

In quantum case, the scheme involves the two modes, signal field 1 and an auxiliary field 2, called the local oscillator. Let $\varrho \in \mathcal{S}(\mathcal{H}_1)$ be the initial state of the signal field and assume the auxiliary field is in a coherent state $|z\rangle$, $z \in \mathbb{C}$. These fields are coupled via a beam splitter, described by the unitary operator of Eq. (19.40), and the detectors D_1 and D_2 , associated with output modes 3 and 4 (Fig. 19.1) measure the smeared photon numbers $n \mapsto E_n^{ej}$, $j = 1, 2$. To register a scaled photon number differences, the set of possible measurement outcomes is taken to be

$$\Omega = \left\{ \frac{1}{\sqrt{2}|z|} \left(\frac{n}{\epsilon_2} - \frac{m}{\epsilon_1} \right) \mid m, n \in \mathbb{N} \right\}.$$

This specific choice for the scaling is motivated by the fact that it assures that for a coherent signal state the first moment of the probability measure remains finite in the limit $|z| \rightarrow \infty$. The detection statistics is thus represented by the observable $E_{\epsilon_1, \epsilon_2} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_4)$,

Fig. 19.1 Balanced homodyne detector



$$E_{\epsilon_1, \epsilon_2}(X) = \sum_X E_m^{\epsilon_1} \otimes E_n^{\epsilon_2}$$

where the summation is over those $m, n \in \mathbb{N}_0$ for which $\frac{1}{\sqrt{2}|z|} \left(\frac{n}{\epsilon_2} - \frac{m}{\epsilon_1} \right) \in X$. The signal observable $E_{\epsilon_1, \epsilon_2}^z : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H}_1)$ measured with this setup is determined by the relation

$$\text{tr}[\varrho E_{\epsilon_1, \epsilon_2}^z(X)] = \text{tr}[U_{\text{BS}}(\varrho \otimes |z\rangle\langle z|)U_{\text{BS}}^* E_{\epsilon_1, \epsilon_2}(X)], \quad \varrho \in \mathcal{S}(\mathcal{H}_1), \quad X \in \mathcal{B}(\mathbb{R}),$$

and is thus of the form

$$E_{\epsilon_1, \epsilon_2}^z(X) = V_z^* U_{\text{BS}}^* E_{\epsilon_1, \epsilon_2}(X) U_{\text{BS}} V_z, \quad X \in \mathcal{B}(\mathbb{R})$$

where $V_z : \mathcal{H}_1 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ is the linear isometry $\varphi \mapsto \varphi \otimes |z\rangle$. We are interested in the high-amplitude limit $|z| \rightarrow \infty$ of this observable. The proposition below shows that this limit is a smeared rotated quadrature observable, namely the convolution of $\mathbf{Q}_\theta$, $\theta = \arg z$, by a probability measure $\mu_{\epsilon_1, \epsilon_2}$ depending on the detection efficiencies ϵ_1 and ϵ_2 .

We fix the phase $\theta \in [0, 2\pi)$ of the local oscillator and take an arbitrary sequence $(r_k)_{k \in \mathbb{N}}$ of positive numbers such that $\lim_{k \rightarrow \infty} r_k = \infty$. Let $z_k = r_k e^{i\theta}$, so that we obtain a sequence $(E_{\epsilon_1, \epsilon_2}^{z_k})_{k \in \mathbb{N}}$ of observables $\mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H}_1)$. Suppose that $\epsilon_1 < 1$ or $\epsilon_2 < 1$, and define the probability density $f_{\epsilon_1, \epsilon_2} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f_{\epsilon_1, \epsilon_2}(x) = \sqrt{\frac{2\epsilon_1\epsilon_2}{\pi(\epsilon_1 - 2\epsilon_1\epsilon_2 + \epsilon_2)}} \exp\left(-\frac{2\epsilon_1\epsilon_2}{\epsilon_1 - 2\epsilon_1\epsilon_2 + \epsilon_2}x^2\right). \quad (19.42)$$

Let $\mu_{\epsilon_1, \epsilon_2} : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be the probability measure determined by $f_{\epsilon_1, \epsilon_2}$, that is, $\mu_{\epsilon_1, \epsilon_2}(X) = \int_X f_{\epsilon_1, \epsilon_2}(x) dx$ for all $X \in \mathcal{B}(\mathbb{R})$. We extend the definition of $\mu_{\epsilon_1, \epsilon_2}$ to include also the case of ideal detectors, and thus we define $\mu_{1,1}$ as the Dirac measure concentrated at the origin.

Proposition 19.1 *For all $\epsilon_1, \epsilon_2 \in (0, 1]$ the sequence $(E_{\epsilon_1, \epsilon_2}^{z_k})_{k \in \mathbb{N}}$ converges to $\mu_{\epsilon_1, \epsilon_2} * Q_\theta$ in the following sense:*

$$\lim_{k \rightarrow \infty} \text{tr}[\varrho E_{\epsilon_1, \epsilon_2}^{z_k}(X)] = \text{tr}[\varrho(\mu_{\epsilon_1, \epsilon_2} * Q_\theta)(X)] \quad (19.43)$$

for all $\varrho \in \mathcal{T}(\mathcal{H}_1)$ and $X \in \mathcal{B}(\mathbb{R})$ such that $(\mu_{\epsilon_1, \epsilon_2} * Q_\theta)(\partial X) = 0$ where ∂X is the boundary of X .

Proof By polarisation and Theorem 2.1 of [14], one sees that (19.43) is equivalent with the condition

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} f(x) \text{tr}[\varrho E_{\epsilon_1, \epsilon_2}^{z_k}(dx)] = \int_{\mathbb{R}} f(x) \text{tr}[\varrho(\mu_{\epsilon_1, \epsilon_2} * Q_\theta)(dx)]$$

for all $\varrho \in \mathcal{T}(\mathcal{H}_1)$ and for all bounded continuous functions $f : \mathbb{R} \rightarrow \mathbb{R}$, or with the condition

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} e^{itx} \text{tr}[\varrho E_{\epsilon_1, \epsilon_2}^{z_k}(dx)] = \int_{\mathbb{R}} e^{itx} \text{tr}[\varrho(\mu_{\epsilon_1, \epsilon_2} * Q_\theta)(dx)] \quad (19.44)$$

for all $\varrho \in \mathcal{T}(\mathcal{H}_1)$ and $t \in \mathbb{R}$ according to the continuity theorem [14, Theorem 7.6]. Since the sequences $k \mapsto \int e^{itx} E_{\epsilon_1, \epsilon_2}^{z_k}(dx) \in \mathcal{L}(\mathcal{H})$ are norm bounded and the linear combinations of coherent states are dense in $\mathcal{H}$ it is enough to show that (19.44) holds for any $\varrho = |\beta\rangle\langle\alpha|$ where $\alpha, \beta \in \mathbb{C}$ and $t \in \mathbb{R}$, that is, we need to show that

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} e^{itx} \langle\alpha| E_{\epsilon_1, \epsilon_2}^{z_k}(dx) |\beta\rangle = \int_{\mathbb{R}} e^{itx} \langle\alpha| (\mu_{\epsilon_1, \epsilon_2} * Q_\theta)(dx) |\beta\rangle. \quad (19.45)$$

Since $\int_{\mathbb{R}} e^{itx} \langle\alpha| E_{\epsilon_1, \epsilon_2}^{z_k}(dx) |\beta\rangle = \sum_{n,m=0}^{\infty} \exp\left[\frac{it}{\sqrt{2}|z_k|} \left(\frac{n}{\epsilon_2} - \frac{m}{\epsilon_1}\right)\right] c_{nm}$ where

$$\begin{aligned} c_{nm} &= \langle\alpha| E_{\epsilon_1, \epsilon_2}^{z_k} (\{2^{-1/2}|z_k|^{-1} (\epsilon_2^{-1}n - \epsilon_1^{-1}m)\}) |\beta\rangle \\ &= \langle\alpha| V_{z_k}^* U_{\text{BS}}^* (E_m^{\epsilon_1} \otimes E_n^{\epsilon_2}) U_{\text{BS}} V_{z_k} |\beta\rangle \\ &= \langle 2^{-1/2}(\alpha - z_k) | E_m^{\epsilon_1} | 2^{-1/2}(\beta - z_k) \rangle \langle 2^{-1/2}(\alpha + z_k) | E_n^{\epsilon_2} | 2^{-1/2}(\beta + z_k) \rangle \\ &= \frac{1}{m!n!} \left(\frac{\epsilon_1}{2}(\bar{\alpha} - \bar{z}_k)(\beta - z_k)\right)^m \left(\frac{\epsilon_2}{2}(\bar{\alpha} + \bar{z}_k)(\beta + z_k)\right)^n \times \\ &\quad \times e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 - |z_k|^2} e^{\frac{1}{2}(1-\epsilon_1)(\bar{\alpha}-\bar{z}_k)(\beta-z_k) + \frac{1}{2}(1-\epsilon_2)(\bar{\alpha}+\bar{z}_k)(\beta+z_k)}, \end{aligned}$$

the left-hand side of (19.45) reduces to

$$e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \bar{\alpha}\beta} e^{\frac{i\theta}{\sqrt{2}}(\bar{\alpha}e^{i\theta} + \beta e^{-i\theta})} e^{-\frac{t^2}{8}\left(\frac{1}{\epsilon_1} + \frac{1}{\epsilon_2}\right)}$$

by straightforward calculation. Next we show that this is the right-hand side of (19.45).

For all $X \in \mathcal{B}(\mathbb{R})$ one has

$$\langle \alpha | \mathbf{Q}_\theta(X) | \beta \rangle = \langle \alpha | e^{i\theta N} \mathbf{Q}(X) e^{-i\theta N} | \beta \rangle = \langle e^{-i\theta} \alpha | \mathbf{Q}(X) | e^{-i\theta} \beta \rangle$$

so that, by denoting $\alpha' = e^{-i\theta} \alpha$ and $\beta' = e^{-i\theta} \beta$, we may express the density of the measure $X \mapsto \langle \alpha | (\mu_{\epsilon_1, \epsilon_2} * \mathbf{Q}_\theta)(X) | \beta \rangle$ as

$$x \mapsto \int_{\mathbb{R}} f_{\epsilon_1, \epsilon_2}(x - y) \langle \alpha' | \mathbf{Q}(dy) | \beta' \rangle$$

if $\epsilon_1 \epsilon_2 < 1$. Putting $\alpha' = \frac{1}{\sqrt{2}}(q + ip)$ and $\beta' = \frac{1}{\sqrt{2}}(u + iv)$ we find, by using the position representation, that the right-hand side of (19.45) is

$$\begin{aligned} \int_{\mathbb{R}} e^{itx} \langle \alpha | (\mu_{\epsilon_1, \epsilon_2} * \mathbf{Q}_\theta)(dx) | \beta \rangle &= \int_{\mathbb{R}} e^{itx} \left(\int_{\mathbb{R}} f_{\epsilon_1, \epsilon_2}(x - y) \langle \alpha' | \mathbf{Q}(dy) | \beta' \rangle \right) dx \\ &= \frac{1}{\pi} \sqrt{\frac{2\epsilon_1 \epsilon_2}{\epsilon_1 - 2\epsilon_1 \epsilon_2 + \epsilon_2}} \times \\ &\times \int_{\mathbb{R}} e^{itx} \left(\int_{\mathbb{R}} e^{-\frac{2\epsilon_1 \epsilon_2}{\epsilon_1 - 2\epsilon_1 \epsilon_2 + \epsilon_2}(x-y)^2} e^{\frac{i}{2}(qp - uv)} e^{iy(v-p)} e^{-\frac{1}{2}(y-q)^2 - \frac{1}{2}(y-u)^2} dy \right) dx \\ &= e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \bar{\alpha}\beta} e^{\frac{i\theta}{\sqrt{2}}(\bar{\alpha}e^{i\theta} + \beta e^{-i\theta})} e^{-\frac{t^2}{8}\left(\frac{1}{\epsilon_1} + \frac{1}{\epsilon_2}\right)} \end{aligned}$$

when $\epsilon_1 \epsilon_2 < 1$. In the case $\epsilon_1 = \epsilon_2 = 1$, $\mu_{1,1}$ is the Dirac measure at 0, so that $\mu_{1,1} * \mathbf{Q}_\theta = \mathbf{Q}_\theta$ and

$$\begin{aligned} \int_{\mathbb{R}} e^{itx} \langle \alpha | (\mu_{1,1} * \mathbf{Q}_\theta)(dx) | \beta \rangle &= \int_{\mathbb{R}} e^{itx} \langle \alpha' | \mathbf{Q}(dx) | \beta' \rangle \\ &= \int_{\mathbb{R}} e^{itx} e^{\frac{i}{2}(qp - uv)} e^{ix(v-p)} e^{-\frac{1}{2}(x-q)^2 - \frac{1}{2}(x-u)^2} dx \\ &= e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \bar{\alpha}\beta} e^{\frac{i\theta}{\sqrt{2}}(\bar{\alpha}e^{i\theta} + \beta e^{-i\theta})} e^{-\frac{t^2}{4}}. \end{aligned}$$

Hence, Eq. (19.45) holds and the proposition is proved. $\square$

19.4 Eight-Port Homodyne Detection and Phase Space Observables

Eight-port homodyne detection is another important quantum optical tool for the reconstruction of a quantum state of light. It also provides an optical implementation of covariant phase space observables which serve as approximate joint measurements of position and momentum observables.

The eight-port homodyne detector involves four input modes, four 50:50 beam splitters, a phase shifter, and four photon detectors, Fig. 19.2.⁴ If $\mathcal{H}_j = L^2(\mathbb{R})$, $j = 1, 2, 3, 4$, is the Hilbert space of the j th input mode, then the Hilbert space of the entire four mode field is $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \mathcal{H}_4$. We denote by ϱ the state of the signal field and by S the state of the parameter field. If the coherent local oscillator is in the state $|\sqrt{2}z\rangle$, $z \in \mathbb{C}$, the initial state of the four-mode field is

$$\varrho \otimes S \otimes |0\rangle\langle 0| \otimes |\sqrt{2}z\rangle\langle \sqrt{2}z|.$$

In this case we use the notation U_{ij} for the unitary transform U_{BS} representing the 50:50 beam splitter. Here the subscripts refer to the primary and secondary input modes, that is, the first and second components of the tensor product in Eq. (19.40). The dashed lines in Fig. 19.2 represent the primary input modes. The phase shifter with phase shift δ is modelled with the unitary operator $e^{i\delta N}$.

We assign to each detector D_j a quantum efficiency $\epsilon_j \in (0, 1]$, so that each detector measures the observable defined in Eq. (19.35). The detection is represented by the biobservable

$$(X, Y) \mapsto E_{\epsilon_1, \epsilon_3}(X) \otimes E_{\epsilon_2, \epsilon_4}(Y) = \sum_{X, Y} E_k^{\epsilon_1} \otimes E_l^{\epsilon_2} \otimes E_m^{\epsilon_3} \otimes E_n^{\epsilon_4},$$

where the summation is now taken over those $k, l, m, n \in \mathbb{N}_0$ for which

$$\frac{1}{\sqrt{2}|z|} \left(\frac{m}{\epsilon_3} - \frac{k}{\epsilon_1} \right) \in X \quad \text{and} \quad \frac{1}{\sqrt{2}|z|} \left(\frac{n}{\epsilon_4} - \frac{l}{\epsilon_2} \right) \in Y.$$

The state of the entire four-mode field before detection is

$$\sigma_{\varrho, S, z, \delta} = U_{13} \otimes U_{24} (U_{12}(\varrho \otimes S) U_{12}^* \otimes |z\rangle\langle z| \otimes |ze^{i\delta}\rangle\langle ze^{i\delta}|) U_{13}^* \otimes U_{24}^*,$$

so that the detection statistics is given by the probability bimeasure

$$(X, Y) \mapsto \text{tr}[\sigma_{\varrho, S, z, \delta} E_{\epsilon_1, \epsilon_3}(X) \otimes E_{\epsilon_2, \epsilon_4}(Y)].$$

There is a unique signal observable $E^{S, z, \delta} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H}_1)$ such that for all ϱ, X, Y ,

⁴This presentation follows closely [13].

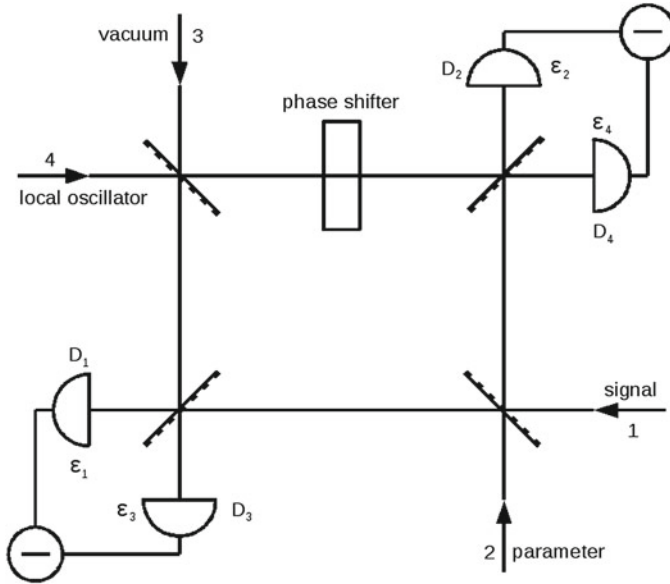


Fig. 19.2 Eight-port homodyne detector

$$\mathrm{tr}[\varrho E^{S,z,\delta}(X \times Y)] = \mathrm{tr}[\sigma_{\varrho,S,z,\delta} E_{\epsilon_1,\epsilon_3}^z(2^{-1/2}X) \otimes E_{\epsilon_2,\epsilon_4}^z(2^{-1/2}Y)],$$

where the scaling has been chosen for later convenience. In order to calculate the high-amplitude limit we express $E^{S,z,\delta}$ in terms of the unsharp homodyne detection observables $E_{\epsilon_1,\epsilon_3}^z$ and $E_{\epsilon_2,\epsilon_4}^{ze^{i\delta}}$. In fact, after simple calculations we find that

$$\mathrm{tr}[\varrho E^{S,z,\delta}(X \times Y)] = \mathrm{tr}[U_{12}(\varrho \otimes S) U_{12}^* E_{\epsilon_1,\epsilon_3}^z(2^{-1/2}X) \otimes E_{\epsilon_2,\epsilon_4}^{ze^{i\delta}}(2^{-1/2}Y)]$$

for all $X, Y \in \mathcal{B}(\mathbb{R})$. Denote again $z_k = r_k e^{i\theta}$, where $\theta \in [0, 2\pi)$ is fixed and $(r_k)_{k \in \mathbb{N}}$ is an arbitrary sequence of positive numbers such that $\lim_{k \rightarrow \infty} r_k = \infty$. It follows from Proposition 19.1 and the boundedness of the associated operators that for all $X, Y \in \mathcal{B}(\mathbb{R})$ such that the boundaries ∂X and ∂Y are of zero Lebesgue measure, we have the convergence

$$\begin{aligned} & \lim_{k \rightarrow \infty} \mathrm{tr}[\varrho E^{S,z_k,\delta}(X \times Y)] \\ &= \mathrm{tr}[U_{12}(\varrho \otimes S) U_{12}^* (\mu_{\epsilon_1,\epsilon_3} * \mathbf{Q}_\theta)(2^{-1/2}X) \otimes (\mu_{\epsilon_2,\epsilon_4} * \mathbf{Q}_{\theta+\delta})(2^{-1/2}Y)]. \end{aligned}$$

Note that the condition of zero Lebesgue measure follows from the fact that each $\mathbf{Q}_\theta$ is unitarily equivalent to $\mathbf{Q}$ which is absolutely continuous with respect to the Lebesgue measure. In particular, we may choose $\theta = 0$ and $\delta = \frac{\pi}{2}$ to obtain the limit

$$\begin{aligned} & \lim_{k \rightarrow \infty} \operatorname{tr} [\varrho \mathbf{E}^{S, r_k, \frac{\pi}{2}} (X \times Y)] \\ &= \operatorname{tr} [U_{12} (\varrho \otimes S) U_{12}^* (\mu_{\epsilon_1, \epsilon_3} * \mathbf{Q}) (2^{-1/2} X) \otimes (\mu_{\epsilon_2, \epsilon_4} * \mathbf{P}) (2^{-1/2} Y)]. \end{aligned}$$

To find the explicit form of the high-amplitude limit observable, denote $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ and let $\mu_{\bar{\epsilon}} : \mathcal{B}(\mathbb{R}^2) \rightarrow [0, 1]$ be the unique probability measure satisfying

$$\mu_{\bar{\epsilon}}(X \times Y) = \mu_{\epsilon_1, \epsilon_3}(2^{-1/2} X) \mu_{\epsilon_2, \epsilon_4}(2^{-1/2} Y) \quad (19.46)$$

for all $X, Y \in \mathcal{B}(\mathbb{R})$. Let C denote the conjugation map $\psi \mapsto (x \mapsto \overline{\psi(x)})$ and let $(r_k)_{k \in \mathbb{N}}$ be as before. The high-amplitude limit observable, given by the proposition below, is a smeared phase space observable $\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}$. Recall that $\mu_{1,1,1,1} = \mu_{1,1} \times \mu_{1,1}$ is the point measure at the origin, so that the limit observable associated with the ideal detectors is the phase space observable $\mathbf{G}^{\text{CSC}^{-1}}$.

Proposition 19.2 *The sequence $(\mathbf{E}^{S, r_k, \frac{\pi}{2}})_{k \in \mathbb{N}}$ converges to $\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}$ in the following sense:*

$$\lim_{k \rightarrow \infty} \operatorname{tr} [\varrho \mathbf{E}^{S, r_k, \frac{\pi}{2}} (X \times Y)] = \operatorname{tr} [\varrho (\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}) (X \times Y)]$$

for any $\varrho \in \mathcal{T}(\mathcal{H}_1)$ and $X, Y \in \mathcal{B}(\mathbb{R})$ such that ∂X and ∂Y are of zero Lebesgue measure.

Proof According to the preceding discussion, the proposition follows from

$$\begin{aligned} & \operatorname{tr} [U_{12} (\varrho \otimes S) U_{12}^* (\mu_{\epsilon_1, \epsilon_3} * \mathbf{Q}) (2^{-1/2} X) \otimes (\mu_{\epsilon_2, \epsilon_4} * \mathbf{P}) (2^{-1/2} Y)] \\ &= \operatorname{tr} [\varrho (\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}) (X \times Y)], \quad X, Y \in \mathcal{B}(\mathbb{R}^2), \end{aligned} \quad (19.47)$$

which we will prove next. Let $\varphi \in \mathcal{H}_1$ and $\psi \in \mathcal{H}_2$ be unit vectors. We use the fact that $\mathbf{P}(\cdot) = F^{-1} \mathbf{Q}(\cdot) F$, where F is the Fourier–Plancherel operator. Furthermore, the relation

$$(I \otimes F) U_{12} (\varphi \otimes \psi) (x, y) = \frac{1}{\sqrt{\pi}} \langle W_{\sqrt{2}x, \sqrt{2}y} C \psi | \varphi \rangle$$

holds for all $y \in \mathbb{R}$ and almost all $x \in \mathbb{R}$ (see, e.g., the proof of [15, Lemma 2]). Now a direct calculation shows that

$$\begin{aligned} & \operatorname{tr} [U_{12} (P[\varphi] \otimes P[\psi]) U_{12}^* (\mu_{\epsilon_1, \epsilon_3} * \mathbf{Q}) (2^{-1/2} X) \otimes (\mu_{\epsilon_2, \epsilon_4} * \mathbf{P}) (2^{-1/2} Y)] \\ &= \langle (I \otimes F) U_{12} (\varphi \otimes \psi) | \\ & \quad (\mu_{\epsilon_1, \epsilon_3} * \mathbf{Q}) (2^{-1/2} X) \otimes (\mu_{\epsilon_2, \epsilon_4} * \mathbf{Q}) (2^{-1/2} Y) (I \otimes F) U_{12} (\varphi \otimes \psi) \rangle \\ &= \iint \mu_{\epsilon_1, \epsilon_3} (2^{-1/2} X - x) \mu_{\epsilon_2, \epsilon_4} (2^{-1/2} Y - y) |(I \otimes F) U_{12} \varphi \otimes \psi (x, y)|^2 dx dy \\ &= \frac{1}{\pi} \iint \mu_{\epsilon_1, \epsilon_3} (2^{-1/2} X - x) \mu_{\epsilon_2, \epsilon_4} (2^{-1/2} Y - y) |\langle W_{\sqrt{2}x, \sqrt{2}y} C \psi | \varphi \rangle|^2 dx dy \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \iint \mu_{\epsilon_1, \epsilon_3}(2^{-1/2}(X - x')) \mu_{\epsilon_2, \epsilon_4}(2^{-1/2}(Y - y')) |\langle W_{x', y'} C \psi | \varphi \rangle|^2 dx' dy' \\
&= \frac{1}{2\pi} \iint \mu_{\bar{\epsilon}}(X \times Y - (x', y')) |\langle W_{x', y'} C \psi | \varphi \rangle|^2 dx' dy' \\
&= \langle \varphi | (\mu_{\bar{\epsilon}} * \mathbf{G}^{C\psi})(X \times Y) \varphi \rangle
\end{aligned}$$

for all $X, Y \in \mathcal{B}(\mathbb{R})$, so that Eq. (19.47) holds for $\varrho = P[\varphi]$ and $S = P[\psi]$. Since both sides of (19.47) depend linearly and continuously on ϱ and S , the validity of the equation in the general case follows by using the spectral representations for ϱ and S . This completes our proof. $\square$

It can be shown that the covariance is not lost in the process of smearing. Indeed, let

$$\mu_{\bar{\epsilon}} * CSC^{-1} = \int W_{qp} CSC^{-1} W_{qp}^* d\mu_{\bar{\epsilon}}(q, p), \quad (19.48)$$

be a state defined as the weak integral [16]. Then the high-amplitude limit observable $\mu_{\bar{\epsilon}} * \mathbf{G}^{CSC^{-1}}$ is a covariant phase space observable with the generating operator (19.48), that is,

$$\mu_{\bar{\epsilon}} * \mathbf{G}^{CSC^{-1}} = \mathbf{G}^{\mu_{\bar{\epsilon}} * CSC^{-1}}.$$

Since there is a one-to-one correspondence between the covariant phase space observables and the generating operators, many questions concerning the properties of a given observable can be answered by studying only the properties of the generating operator. As an example, consider the extremality of an observable $\mathbf{G}^S$. The set of all covariant phase space observables is a convex set, and the convex combination of two observables $\mathbf{G}^{S_1}$ and $\mathbf{G}^{S_2}$ is simply $t\mathbf{G}^{S_1} + (1-t)\mathbf{G}^{S_2} = \mathbf{G}^{tS_1 + (1-t)S_2}$. Hence, an observable $\mathbf{G}^S$ is an extreme point of the convex set of covariant phase space observables if and only if S is an extreme point of the set of positive trace class operators with unit trace. Furthermore, the extreme points of this set are the one-dimensional projections $P[\varphi]$, $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$, that is, the pure states. In the case of our specific observable, we obtain the following result.

Proposition 19.3 *The generating operator $\mu_{\bar{\epsilon}} * CSC^{-1}$ is a pure state if and only if S is a pure state and the detectors are ideal.*

The consequence of Proposition 19.3 is that whenever detector inefficiencies are present, the measured observable can be written as a nontrivial convex combination $\mathbf{G}^{\mu_{\bar{\epsilon}} * CSC^{-1}} = t\mathbf{G}^{S_1} + (1-t)\mathbf{G}^{S_2}$ for some generating operators $S_1 \neq S_2$, and for some weight factor $t \in (0, 1)$. This is usually taken to correspond to classical randomisation between the two observables $\mathbf{G}^{S_1}$ and $\mathbf{G}^{S_2}$. In particular, Proposition 19.3 thus verifies the perhaps intuitive fact that ideal detectors are necessary for the measurement to be a pure quantum measurement.

In the process of smearing, the state distinction power of the observable is usually diminished. But this need not always be the case, particularly when the convolving measure is Gaussian. In fact, the following result holds true.

Proposition 19.4 *The observables $\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}$ and $\mathbf{G}^{\text{CSC}^{-1}}$ are informationally equivalent. In particular, $\mu_{\bar{\epsilon}} * \mathbf{G}^{\text{CSC}^{-1}}$ is informationally complete if and only if $\mathbf{G}^{\text{CSC}^{-1}}$ is informationally complete.*

19.5 Eight-Port Homodyne Detection and Phase Observables

An important class of covariant phase observables (see Chap. 16) arise as the angle margins of certain covariant phase space observables. Their physical significance is further emphasised by the fact that any phase space observable can in principle be measured via eight-port homodyne detection as we have seen. It should be stressed that even though covariant phase observables arise naturally as the margins of covariant phase space observables, not all of them are obtained in this way. In particular, the canonical phase observable Φ_{can} is not the angle margin of any phase space observable [17] but this does not mean that Φ_{can} cannot be measured by using an eight-port homodyne detector as we will see in the next subsection.

Any covariant phase-space observable $\mathbf{G}^\sigma$ is generated by a unique positive unit trace operator σ . Now let us denote by $\Phi^\sigma : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$ the angle margin of $\mathbf{G}^\sigma$, that is,

$$\Phi^\sigma(X) = \mathbf{G}^\sigma(X \times [0, \infty)), \quad X \in \mathcal{B}([0, 2\pi)),$$

where the relation between the polar and Cartesian coordinates is given by $re^{i\theta} = \frac{1}{\sqrt{2}}(q + ip)$. The key result needed in our study is [17, Theorem 4.1] which states that Φ^σ is a phase observable if and only if σ is diagonal in the number basis. The simplest and, from the experimental point of view, most useful example is the case $\sigma = |0\rangle\langle 0|$, that is, when the observable is generated by the vacuum state. In this case the phase distribution is just the angle margin of the Husimi Q -function of the field.

As we saw in the preceding section, the high-amplitude limit observable of double homodyne detection is a covariant phase space observable with the generating operator $\sigma = \mu_{\bar{\epsilon}} * \text{CSC}^{-1}$ where $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$. By denoting $\epsilon_{ij} = \frac{2\epsilon_i\epsilon_j}{\epsilon_i + \epsilon_j}$ we have

Proposition 19.5 *If S is diagonal, then $\mu_{\bar{\epsilon}} * \text{CSC}^{-1}$ is diagonal if and only if $\epsilon_{13} = \epsilon_{24}$. Conversely, if $\epsilon_{13} = \epsilon_{24}$, then $\mu_{\bar{\epsilon}} * \text{CSC}^{-1}$ is diagonal if and only if S is diagonal.*

Note that neither of the conditions in Proposition 19.5 is necessary for $\mu_{\bar{\epsilon}} * \text{CSC}^{-1}$ to be diagonal in the number basis [18].

Let $\sigma = \mu_{\bar{\epsilon}} * \text{CSC}^{-1}$. Since the observable measured in the double homodyne detection scheme is $\mathbf{G}^\sigma$ it is a slight misuse of terminology to call this a *direct* measurement of the angle margin Φ^σ . However, the analysis below shows that this scheme can be used to directly measure Φ^σ . For a local oscillator with a finite intensity

$r = |z|$ this scheme defines an observable $\mathbf{E}^{S,r,\frac{\pi}{2}} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(\mathcal{H})$ (which converges to $\mathbf{G}^\sigma$ when $r \rightarrow \infty$). Now $\mathbf{E}^{S,r,\frac{\pi}{2}}$ is a discrete observable and the measurement outcomes consist of pairs $(q, p) \in \mathbb{R}^2$. Let $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow [0, 2\pi)$ be the pointer function which assigns to each pair the corresponding argument, that is, $f(q, p) = \alpha_{qp}$ defined by

$$\cos \alpha_{qp} = \frac{q}{\sqrt{q^2 + p^2}}, \quad \sin \alpha_{qp} = \frac{p}{\sqrt{q^2 + p^2}}$$

and denote $\Phi_r^\sigma : \mathcal{B}([0, 2\pi)) \rightarrow \mathcal{L}(\mathcal{H})$,

$$\Phi_r^\sigma(X) = \mathbf{E}^{S,r,\frac{\pi}{2}}(f^{-1}(X) \cup \{(0, 0)\}).$$

Then it can be shown that

$$\lim_{r \rightarrow \infty} \Phi_r^\sigma = \Phi^\sigma.$$

In this sense, by choosing to record only the values α_{qp} we see that eight-port homodyne detection in the high-amplitude limit can be used as a direct measurement of Φ^σ .

Finally, we study the classical limit of phase space phase observables arising from double homodyne detection. Suppose that σ is a diagonal state and the signal field is in a coherent state $|\alpha\rangle$. Now the phase probability measure $X \mapsto \langle \alpha | \Phi^\sigma(X) | \alpha \rangle$ tends to a (2π) -periodic Dirac measure concentrated on $\arg \alpha - \theta$ when $|\alpha| \rightarrow \infty$ [19]. Here θ is the fixed phase of the high amplitude local oscillator state $|z\rangle$. This holds for *any* state S of the parameter field. Hence, when $|\alpha| \gg 0$ and $|z| \gg 0$, one can measure the classical phase difference of the phase parameters of the signal state $|\alpha\rangle$ and the local oscillator state $|z\rangle$.

Direct Measurement of the Canonical Phase Via Eight-Port Homodyne Detection

In this subsection, we show that, in principle, the canonical phase of the signal state can be measured via eight-port homodyne detection by adding an extra two-mode unitary coupling in front of the eight-port detector. This subsection is based on [20].

First note that in the usual eight-port homodyne detector, the overall unitary coupling consisting of the beam splitter and the phase shifter is given by $U_{12} : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_3 \otimes \mathcal{H}_4$,

$$(U_{12}|m\rangle \otimes |n\rangle)(x, y) = \frac{1}{\sqrt{\pi}} \langle n | W_{\sqrt{2}x, \sqrt{2}y}^* | m \rangle,$$

where $\mathcal{H}_1 \otimes \mathcal{H}_2$ consists of the input signal and parametric field modes and $\mathcal{H}_3 \otimes \mathcal{H}_4 \simeq L^2(\mathbb{R}^2)$ is the output space (see, e.g., the proof of [15, Lemma 2]). By using the identification $\mathbb{R}^2 \ni (x, y) \mapsto (x + iy)/\sqrt{2} \in \mathbb{C}$ and polar coordinates $(x + iy = re^{i\theta})$, the observable measured with the eight-port detector is then the canonical spectral measure $\mathbf{Q} : \mathcal{B}(\mathbb{C}) \rightarrow \mathcal{L}(\mathcal{H}_3 \otimes \mathcal{H}_4)$. As we have seen, if ϱ and S are the states of the signal and parameter fields, respectively, then

$$\mathrm{tr}[U_{12}(\varrho \otimes S)U_{12}^* \mathbf{Q}(Z)] = \mathrm{tr}[\varrho \mathbf{G}^{\mathrm{CSC}^{-1}}(Z)]$$

for all $Z \in \mathcal{B}(\mathbb{C})$.

Now, adding a unitary coupling $W : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ in front of the detector amounts to applying an extra optical component to the measurement setup prior to the beam splitter entangling the input and parameter fields. The total coupling $U_{12}W$ will then be denoted by V . We define W by requiring that V is of the form

$$(V|m\rangle \otimes |n\rangle)(r, \theta) = \frac{1}{\sqrt{\pi}} e^{i\theta(n-m)} L_{\min\{m,n\}}(r^2) e^{-\frac{r^2}{2}},$$

where $L_k(x)$ is the k th Laguerre polynomial. Since V is clearly a unitary operator, the unitarity of $W = U_{12}^* V$ is also ensured. If the parameter field is in the vacuum state then the measured observable $\mathbf{E} : \mathcal{B}(\mathbb{C}) \rightarrow \mathcal{L}(\mathcal{H}_1)$ of the signal field is determined by

$$\mathrm{tr}[V(\varrho \otimes |0\rangle\langle 0|)V^* \mathbf{Q}(Z)] = \mathrm{tr}[\varrho \mathbf{E}(Z)]$$

for all $Z \in \mathcal{B}(\mathbb{C})$. It is easy to see that

$$\mathbf{E}(Z) = \sum_{m,n=0}^{\infty} \frac{1}{2\pi} \int_Z e^{i\theta(m-n)} e^{-r^2} dr^2 d\theta |m\rangle\langle n|.$$

Especially, the angle margin of $\mathbf{E}$ is Φ_{can} .

The action of the operator W is now given by

$$W(|m\rangle \otimes |n\rangle) = \sum_{k,l=0}^{\infty} \langle U_{12}|k\rangle \otimes |l\rangle |V|m\rangle \otimes |n\rangle \rangle |k\rangle \otimes |l\rangle.$$

Let us denote $\alpha_{kl,mn} = \langle U|k\rangle \otimes |l\rangle |V|m\rangle \otimes |n\rangle \rangle$ so that

$$\begin{aligned} \alpha_{kl,mn} &= (-1)^{\max\{0,l-k\}} \sqrt{\frac{\min\{k,l\}!}{\max\{k,l\}!}} \frac{1}{\pi} \int_0^\infty \int_0^{2\pi} e^{i\theta(k-l+n-m)} r^{|k-l|} \\ &\quad \cdot L_{\min\{k,l\}}^{[k-l]}(r^2) L_{\min\{m,n\}}(r^2) e^{-r^2} r dr d\theta \\ &= \delta_{l,k+n-m} (-1)^{\max\{0,n-m\}} \sqrt{\frac{\min\{k,k+n-m\}!}{\max\{k,k+n-m\}!}} \int_0^\infty x^{\frac{1}{2}|n-m|} \\ &\quad \cdot L_{\min\{k,k+n-m\}}^{[n-m]}(x) L_{\min\{m,n\}}(x) e^{-x} dx. \end{aligned}$$

The above integral can be manipulated further, but it seems that no simple closed expression exists. Since the canonical phase is obtained as the special case $n = 0$ of the above consideration, we will next focus on that. Now the above integrals are

easily evaluated and we obtain the expressions $W(|0\rangle \otimes |0\rangle) = |0\rangle \otimes |0\rangle$ and, for $m > 0$,

$$\begin{aligned} W(|m\rangle \otimes |0\rangle) &= \sum_{k=0}^{\infty} \frac{\Gamma(k + \frac{m}{2} + 1)}{\sqrt{k!(k+m)!}} \frac{m}{2k+m} |k+m\rangle \otimes |k\rangle \\ &= \frac{m}{2\sqrt{m!}} \sum_{k=0}^{\infty} \Gamma(k + \frac{m}{2}) \binom{m+k}{m}^{-1/2} |k+m\rangle \otimes |k\rangle. \end{aligned}$$

To summarise, the canonical phase distribution $\Theta \mapsto \text{tr}[\varrho \Phi_{\text{can}}(\Theta)]$ of any state ϱ can be measured directly by first producing an entangled two-mode input state $S(\varrho) = W(\varrho \otimes |0\rangle\langle 0|)W^*$ and then measuring $S(\varrho)$ via eight-port homodyne detection.

Quantum Homodyne Tomography and Generalised Markov Kernels

As we have seen, in the high amplitude limit of the eight-port homodyne detection, one can measure a phase space observable $\mathbf{G}^\sigma$ which can be used to reconstruct the possibly unknown signal state ϱ (see, Chap. 18). Similarly, in the ideal balanced homodyne detection, one can measure the rotated quadratures $\mathbf{Q}_\theta$, $\theta \in [0, 2\pi)$, by varying the phase $\theta = \arg z$ of the high amplitude local oscillator state $|z\rangle$, $|z| \gg 0$ (see, Sect. 19.3). Measuring the whole collection of rotated quadratures in a signal state ϱ enables one to reconstruct ϱ (see, Chap. 18). We call this combination of $\mathbf{Q}_\theta$ -measurements the *quantum (balanced) homodyne tomography*. Now the set $\{\mathbf{Q}_\theta \mid \theta \in [0, 2\pi)\}$ of all rotated quadratures can be combined to a single observable $\mathbf{Q}_{\text{ht}} : \mathcal{B}([0, 2\pi) \times \mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$ where

$$\mathbf{Q}_{\text{ht}}(\Theta \times X) = \frac{1}{2\pi} \int_{\Theta} \mathbf{Q}_\theta(X) d\theta, \quad \Theta \in \mathcal{B}([0, 2\pi)), \quad X \in \mathcal{B}(\mathbb{R}).$$

(Cf. [21–23]). Next we find a connection between $\mathbf{Q}_{\text{ht}}$ and $\mathbf{G}^\sigma$ when σ is sufficiently regular. Indeed, we will show that $\mathbf{G}^\sigma$ is related to $\mathbf{Q}_{\text{ht}}$ via a generalised Markov kernel which we define next.

Let $\mathbf{E}_1 : \mathcal{A}_1 \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{E}_2 : \mathcal{A}_2 \rightarrow \mathcal{L}(\mathcal{H})$ be observables (where $\mathcal{A}_1$ and $\mathcal{A}_2$ are σ -algebras of the outcome sets Ω_1 and Ω_2 , respectively). Let $\mathcal{A}_1^0 \subset \mathcal{A}_1$ be a generating set of $\mathcal{A}_1$, that is, $\mathcal{A}_1$ is the smallest σ -algebra which contains $\mathcal{A}_1^0$. Since $\mathbf{E}_1$ is fully determined by its values on $\mathcal{A}_1^0$ the following definitions are reasonable. We say that $p : \mathcal{A}_1^0 \times \Omega_2 \rightarrow \mathbb{C}$ is a *generalised Markov kernel* if, for all $X \in \mathcal{A}_1^0$, the mapping $\Omega_2 \ni \omega \mapsto p(X, \omega) \in \mathbb{C}$ is measurable, and $\mathbf{E}_1$ is *subordinate* to $\mathbf{E}_2$ when there exists a generalised Markov kernel $p : \mathcal{A}_1^0 \times \Omega_2 \rightarrow \mathbb{C}$ such that

$$\mathbf{E}_1(X) = \int_{\Omega_2} p(X, \omega) d\mathbf{E}_2(\omega), \quad X \in \mathcal{A}_1^0.$$

Thus, if $\mathbf{E}_1$ is subordinate to $\mathbf{E}_2$ one can measure $\mathbf{E}_2$ in a state ϱ and then *construct* the probability measure $X \mapsto \text{tr}[\varrho \mathbf{E}_1(X)]$ by integrating as above.

Let $\mathcal{C}(\mathbb{R}^2)$ be the set of compact subsets of $\mathbb{R}^2$ which is a generating set of $\mathcal{B}(\mathbb{R}^2)$, and let $\mathcal{S}_\downarrow(\mathcal{H})$ be the set consisting of states $\sigma \in \mathcal{S}(\mathcal{H})$ such that the double sequences $(\sigma_{nm})_{n,m \in \mathbb{N}_0}$ of their matrix elements $\sigma_{nm} = \langle n|\sigma|m \rangle$ are rapidly decreasing, that is, the series $\sum_{n,m=0}^{\infty} n^{2j} m^{2k} |\sigma_{nm}|^2$ converge for all $j, k \in \mathbb{N}$. We have the following theorem proved in [23].

Theorem 19.1 *For all $\sigma \in \mathcal{S}_\downarrow(\mathcal{H})$ there exists a generalised Markov kernel*

$$p^\sigma : \mathcal{C}(\mathbb{R}^2) \times ([0, 2\pi) \times \mathbb{R}) \rightarrow \mathbb{C}$$

such that, for all $C \in \mathcal{C}(\mathbb{R}^2)$,

$$\begin{aligned} \mathbf{G}^\sigma(C) &= \int_0^{2\pi} \int_{\mathbb{R}} p^\sigma(C, (\theta, x)) d\mathbf{Q}_{\text{ht}}(\theta, x) \\ &= \sum_{n,m=0}^{\infty} \sigma_{nm} \int_0^{2\pi} \int_{\mathbb{R}} \left[\int_C e^{i(n-m)\theta} f_{nm}(x - q \cos \theta - p \sin \theta) dq dp \right] d\mathbf{Q}_{\text{ht}}(\theta, x) \end{aligned}$$

where the double series converges weakly, and for all $r \in \mathbb{R}$,

$$\begin{aligned} f_{nm}(r) &= \frac{(-1)^{n+m}}{\pi} \sqrt{\frac{n!m!}{2^{n+m}}} \cdot \\ &\quad \cdot \sum_{v=0}^{\min\{n,m\}} \frac{2^v}{v!(n-v)!(m-v)!} \frac{d^{n+m-2v+1}}{dr^{n+m-2v+1}} \left[e^{-r^2} \int_0^r e^{t^2} dt \right]. \end{aligned}$$

For example, when $n = m = 0$ one gets

$$f_{00}(r) = \frac{1}{\pi} \frac{d}{dr} \left[e^{-r^2} \int_0^r e^{t^2} dt \right] = \frac{1}{\pi} (1 - 2r) e^{-r^2} \int_0^r e^{t^2} dt$$

and, for $C = [1, 2] \times [0, 1]$,

$$\begin{aligned} p^{[0]\langle 0|}(C, (0, 0)) &= \int_C f_{00}(-q) dq dp = \int_{-2}^{-1} f_{00}(r) dr \\ &= \frac{1}{\pi} \left[e^{-4} \int_0^2 e^{t^2} dt - e^{-1} \int_0^1 e^{t^2} dt \right] \approx -0.2 < 0 \end{aligned}$$

which shows that $p^{[0]\langle 0|}$ is not a Markov kernel in the usual sense and it cannot be interpreted as a transition probability from $\mathbf{Q}_{\text{ht}}$ to $\mathbf{G}^{[0]}$. Thus, generalised Markov kernels do not seem to have an independent operational meaning.

19.6 Mach–Zehnder Interferometer

We begin with a presentation of variants of the classic split-beam interference experiment using a Mach–Zehnder setup.⁵ The purpose of the original experiment is to store or extract *which way* information and to investigate how this affects the observed interference pattern.

We apply a simplified description, restricting the quantum-optical description of splitters (given in Sect. 19.2) to two “path” modes and single photon state. The two possible paths through the interferometer 1 and 2 are represented by two orthogonal states, denoted $|10\rangle$, $|01\rangle$, associated with the two field modes corresponding to these paths—see Fig. 19.3. Hence the photon’s path degree of freedom is described in terms of a two-dimensional Hilbert space spanned by the single photon states of the two path modes involved.

To define and identify relevant observables operationally, we will allow arbitrary pure and mixed states entering the interferometer. Thus, a photon enters the interferometer on one of the two input paths 1 or 2. It passes a beam-splitter BS_1 , propagates into the two available paths which cross at a second beam-splitter, BS_2 , and propagates to one of two detectors, D_1 and D_2 where it will be recorded. One of the paths contains a phase shifter PS_δ .

Optionally, immediately before entry into BS_1 the photon may be made to interact with a path-detecting probe and become entangled with the path-marking states $|p_1\rangle$, $|p_2\rangle$, which span the probe Hilbert space. The probe could be taken to be another photon coupled through a Kerr medium placed in one of the input paths (see, e.g., [10, Sect. VII.4.1.]).

The experiment concludes with one of the detectors firing and simultaneously the probe being measured if it was activated.

The purpose of this joint measurement is to obtain information about the photon state immediately prior to the interaction with the probe and subsequent passage through the interferometer. Such information is available in the form of the output probabilities, which can be expressed in terms of the photon’s input state and thus define a unique observable of the photon.

State Evolution in the Interferometer

We first consider the evolution of the state of the photon in the interferometer with or without the involvement of the problem effecting the path marking before the beam splitter BS_1 (Fig. 19.3).

A general input from I_1 , I_2 without the path marking interaction switched on (object in input state ψ and probe remaining in a neutral state $|p_0\rangle$) can be represented by the unit vector

$$\psi \otimes |p_0\rangle = (\alpha|10\rangle + \beta|01\rangle) \otimes |p_0\rangle.$$

⁵The development of the examples in this section follows closely the review [24].

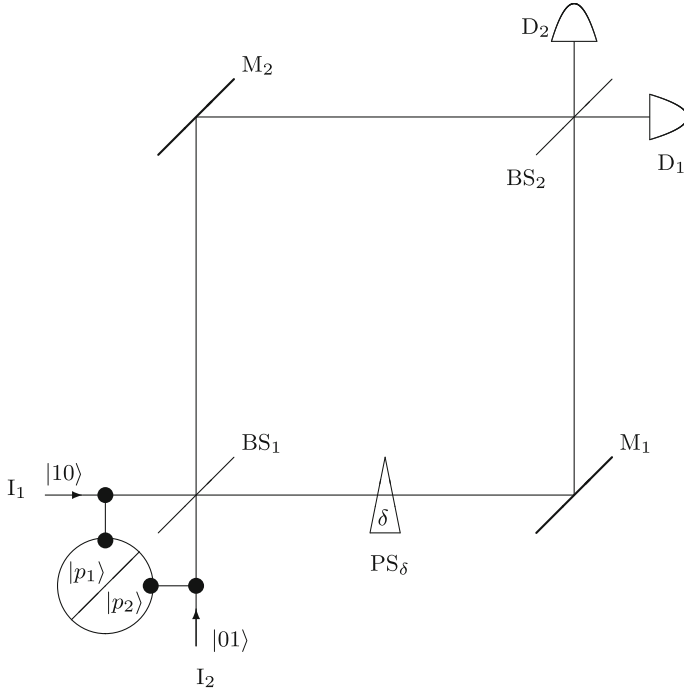


Fig. 19.3 Scheme of a Mach-Zehnder interferometer with path marking and phase shifter

Introducing now the phase shifter in path 1, the photon input states $|10\rangle$, $|01\rangle$ undergo an evolution upon passage through the interferometer and before entering one of the detectors D_1 , D_2 , which we model by the unitary map

$$U'_{BS} \tilde{U}_\delta U'_{BS} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{i\delta} + 1 & i(-e^{i\delta} + 1) \\ i(e^{i\delta} - 1) & e^{i\delta} + 1 \end{pmatrix}.$$

This gives the transitions

$$\begin{aligned} |10\rangle &\rightarrow \frac{1}{2}(e^{i\delta} + 1)|10\rangle + \frac{i}{2}(e^{i\delta} - 1)|01\rangle, \\ |01\rangle &\rightarrow \frac{i}{2}(-e^{i\delta} + 1)|10\rangle + \frac{1}{2}(e^{i\delta} + 1)|01\rangle. \end{aligned} \quad (19.49)$$

Note that in the absence of a phase shifter ($\delta = 0$) this becomes $|10\rangle \rightarrow |10\rangle$ and $|01\rangle \rightarrow |01\rangle$, so that a photon entering through path 1 (2) is found with certainty in detector D_1 (D_2).

The path marking interaction is defined through the map

$$\psi \otimes |p_0\rangle = (\alpha|10\rangle + \beta|01\rangle) \otimes |p_0\rangle \mapsto \alpha|10\rangle \otimes |p_1\rangle + \beta|01\rangle \otimes |p_2\rangle = \Psi_e$$

Since the photon states $|10\rangle, |01\rangle$ evolve according to (19.49), the total final (output) state after the photon passes through beam splitter BS_2 is

$$\begin{aligned}\Psi_f^\delta = & \frac{1}{2}\alpha[(e^{i\delta} + 1)|10\rangle + i(e^{i\delta} - 1)|01\rangle] \otimes |p_1\rangle \\ & + \frac{1}{2}\beta[-i(e^{i\delta} - 1)|10\rangle + (e^{i\delta} + 1)|01\rangle] \otimes |p_2\rangle.\end{aligned}\quad (19.50)$$

In this final photon-probe state the sharp output observable Z , with projections

$$Z_{k\ell} = |\varphi_k\rangle\langle\varphi_k| \otimes |\phi_\ell\rangle\langle\phi_\ell|, \quad (19.51)$$

is measured. Here $|\varphi_k\rangle, k = 1, 2$, are the photon's "path" states, $\varphi_1 = |10\rangle$ and $\varphi_2 = |01\rangle$, interpreted as eigenstates of the detector observable corresponding to D_1, D_2 , and $|\phi_\ell\rangle, \ell = 1, 2$, are eigenstates of a pointer observable of the probe, for which different choices will be made in each of the scenarios discussed below.

The detection statistics of Z in state Ψ_f^δ determines the measured photon observable E , with the effects $E_{k\ell}$:

$$\langle\psi|E_{k\ell}|\psi\rangle = \langle\Psi_f^\delta|Z_{k\ell}|\Psi_f^\delta\rangle. \quad (19.52)$$

This forms the basis for the analysis of all the interferometric schemes discussed in the coming subsections.

Path Detection in Outputs D_1, D_2

With no path marking interaction and with $\delta = 0$, the output state is

$$\Psi_f^o = (\alpha|10\rangle + \beta|01\rangle) \otimes |p_0\rangle.$$

The output of detectors D_1, D_2 with no path marking are represented by the simple pointer observable $Z^{(0)}$ with projections $Z_1^{(0)} = |10\rangle\langle 10| \otimes I, Z_2^{(0)} = |01\rangle\langle 01| \otimes I$, so that the detection probabilities are

$$\begin{aligned}\langle\Psi_f^o|Z_1^{(0)}|\Psi_f^o\rangle &= \langle\psi|10\rangle\langle 10|\psi\rangle = |\alpha|^2, \\ \langle\Psi_f^o|Z_2^{(0)}|\Psi_f^o\rangle &= \langle\psi|01\rangle\langle 01|\psi\rangle = |\beta|^2.\end{aligned}$$

The observable measured by this experiment is then $E^w : k \mapsto E_k^w$, with which we identify with the spectral projections $\frac{1}{2}(I \pm \sigma_z)$ of the Pauli operator σ_z (by putting $|10\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |01\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$):

$$E_1^w = |10\rangle\langle 10| = \frac{1}{2}(I + \sigma_z), \quad E_2^w = |01\rangle\langle 01| = \frac{1}{2}(I - \sigma_z).$$

This confirms that when $\psi = |10\rangle$, the detection probabilities at D_1 and D_2 are 1 and 0, respectively, and similarly, for $\psi = |01\rangle$ they are 0 and 1. Thus, the measured observable is a *which way* observable represented by σ_z .

Interference Detection in D_1, D_2

We now consider the use of the Mach–Zehnder interferometer for an interference measurement. Again we have no interaction with the path marking probe, but now a phase shifter is put in place. We first consider the specific value $\delta = \pi/2$ for the phase shift; then the output state is

$$\Psi_f^{\pi/2} = \frac{(1+i)}{\sqrt{2}} \left[\alpha \frac{1}{\sqrt{2}} (|10\rangle - |01\rangle) + \beta \frac{1}{\sqrt{2}} (|10\rangle + |01\rangle) \right] \otimes |p_0\rangle. \quad (19.53)$$

If a measurement of $Z_1^{(0)} = |10\rangle\langle 10| \otimes I$, $Z_2^{(0)} = |01\rangle\langle 01| \otimes I$ is now applied by observing the outputs of D_1, D_2 , the associated probabilities are

$$\begin{aligned} \langle \Psi_f^{\pi/2} | Z_1^{(0)} \Psi_f^{\pi/2} \rangle &= \frac{1}{2} |\alpha + \beta|^2 = \langle \psi | E_1^i \psi \rangle, \\ \langle \Psi_f^{\pi/2} | Z_2^{(0)} \Psi_f^{\pi/2} \rangle &= \frac{1}{2} |\alpha - \beta|^2 = \langle \psi | E_2^i \psi \rangle, \end{aligned}$$

defining the measured observable $E^i : k \mapsto E_k^i$, with

$$E_1^i = \frac{1}{2}(I + \sigma_x), \quad E_2^i = \frac{1}{2}(I - \sigma_x). \quad (19.54)$$

According to customary practice, a general interference observable is of the form $\sigma_\delta = \cos \delta \sigma_y - \sin \delta \sigma_x$, with $-\pi \leq \delta \leq \pi$, given that the path is represented by σ_z . They are singled out by the condition that the interference contrast can assume its maximum possible value. In this case their eigenstates give equal probabilities of $1/2$ to the path projections $|10\rangle\langle 10|$, $|01\rangle\langle 01|$. Hence the above interference observables σ_δ are complementary to the *which way* observable σ_z .

In the present experiment, the measured input observable is defined by the projections of Eq. (19.54); these are the spectral projections of the operator σ_x , which indeed represents an interference observable.

Interferometric Measurement Preceded by Path Marking

Now consider the case with the path-marking interaction included. We assume the probe states $|p_1\rangle$ and $|p_2\rangle$ to be mutually orthogonal. (In a later application this assumption will be lifted.) The detection statistics for the pointer observable $Z^{(1)}$ with $Z_{k\ell}^{(1)} = |\varphi_k\rangle\langle\varphi_k| \otimes |p_\ell\rangle\langle p_\ell|$ (cf. Eq. (19.51)) in the state Ψ_f^δ determines again the measured observable, $E^{(1)}$, and one obtains

$$\begin{aligned} E_{11}^{(1)} &= \frac{1}{2}(I + \sigma_z) \cos^2 \frac{\delta}{2} & E_{21}^{(1)} &= \frac{1}{2}(I + \sigma_z) \sin^2 \frac{\delta}{2} \\ E_{12}^{(1)} &= \frac{1}{2}(I - \sigma_z) \sin^2 \frac{\delta}{2} & E_{22}^{(1)} &= \frac{1}{2}(I - \sigma_z) \cos^2 \frac{\delta}{2}. \end{aligned} \quad (19.55)$$

These effects are all fractions of one of the two path projections $\frac{1}{2}(I \pm \sigma_z)$.

The marginal observable $F^{(1)}$ associated with the detectors D_1, D_2 is given by

$$\begin{aligned} F_1^{(1)} &= E_{11}^{(1)} + E_{12}^{(1)} = \frac{1}{2}(I + \cos \delta \sigma_z), \\ F_2^{(1)} &= E_{21}^{(1)} + E_{22}^{(1)} = \frac{1}{2}(I - \cos \delta \sigma_z), \end{aligned}$$

and the margin $\mathbf{G}^{(1)}$ measured by the path marker has the effects

$$\begin{aligned} G_1^{(1)} &= E_{11}^{(1)} + E_{21}^{(1)} = \frac{1}{2}(I + \sigma_z), \\ G_2^{(1)} &= E_{22}^{(1)} + E_{12}^{(1)} = \frac{1}{2}(I - \sigma_z). \end{aligned}$$

These represent a sharp path observable.

The observable $\mathbf{F}^{(1)}$ represents a smeared path observable when $\cos \delta \notin \{-1, 1\}$. For $\cos \delta = 1$, say, the outcomes of the path marker observation and the detection at D_1, D_2 are strictly correlated. However, when $\delta \neq 0, \pi$, the interferometer is set to measure an interference observable associated with σ_δ but instead the associated output observable has become a smeared *which way* observable. In the extreme case of a perfect interference setup, $\delta = 0$, the observable $\mathbf{F}^{(1)}$ becomes trivial, $F_1^{(1)} = F_2^{(1)} = \frac{1}{2}I$. The interference is “washed out” as a result of the path-marking interaction. This shows the disturbing effect of the first measurement, the sharp path marking, on the second, interference measurement, in line with Proposition 11.1. After the path marking interaction, all the detectors are able to “see” is a “shadow” of the path information provided by the path marker.

It is possible to define a third marginal observable $\mathbf{H}^{(1)}$,

$$\begin{aligned} H_1^{(1)} &= E_{11}^{(1)} + E_{22}^{(1)} = \cos^2 \frac{\delta}{2} I, \\ H_2^{(1)} &= E_{12}^{(1)} + E_{21}^{(1)} = \sin^2 \frac{\delta}{2} I, \end{aligned}$$

which in the present experiment also turns out to be trivial.

Quantum Erasure

In the previous experiment, each path was correlated with one of two orthogonal marker states. We can now consider a new set of pointer states, which are superpositions of the two orthogonal path-marker states,

$$|\phi_1\rangle = |q_1\rangle = \frac{1}{\sqrt{2}}(|p_1\rangle + e^{i\gamma}|p_2\rangle), \quad |\phi_2\rangle = |q_2\rangle = \frac{1}{\sqrt{2}}(|p_1\rangle - e^{i\gamma}|p_2\rangle).$$

Observing these symmetric states involves outputs for which both $|p_1\rangle$ and $|p_2\rangle$ are equally likely, so no information about the path is recorded.

This setup will be found to yield an implementation of the phenomenon of *quantum erasure*, in which the path information originally stored in the entanglement between probe and photon is erased and interference is fully recovered in the subsequent interferometric measurement on the photon states conditional on the readings q_1, q_2 .

The final state (19.50) in terms of the new pointer states is

$$\begin{aligned} \Psi_f^{\delta,\gamma} = & \frac{1}{2\sqrt{2}} \left[(\alpha(e^{i\delta} + 1) - ie^{-i\gamma}\beta(e^{i\delta} - 1))|10\rangle \right. \\ & \left. + (i\alpha(e^{i\delta} - 1) + e^{-i\gamma}\beta(e^{i\delta} + 1))|01\rangle \right] \otimes |q_1\rangle \\ & + \frac{1}{2\sqrt{2}} \left[(\alpha(e^{i\delta} + 1) + ie^{-i\gamma}\beta(e^{i\delta} - 1))|10\rangle \right. \\ & \left. + (i\alpha(e^{i\delta} - 1) - e^{-i\gamma}\beta(e^{i\delta} + 1))|01\rangle \right] \otimes |q_2\rangle. \end{aligned} \quad (19.56)$$

As before we can find the four joint probabilities for the marker and detector outputs, defined as the expectations of the projections $Z_{k\ell}^{(2)} = |\varphi_k\rangle\langle\varphi_k| \otimes |q_\ell\rangle\langle q_\ell|$, $k, \ell = 1, 2$, cf. Eq. (19.51). The associated input observable $\mathbf{E}^{(2)}$ is determined via the relation

$$\langle \Psi_f^{\delta,\gamma} | Z_{k\ell}^{(2)} | \Psi_f^{\delta,\gamma} \rangle = \langle \psi | E_{k\ell}^{(2)} | \psi \rangle,$$

and we obtain:

$$\begin{aligned} E_{11}^{(2)} &= \frac{1}{4}(I + \sin \delta \cos \gamma \sigma_x + \sin \delta \sin \gamma \sigma_y + \cos \delta \sigma_z) = \frac{1}{4}(I + \mathbf{n}^{(2)} \cdot \boldsymbol{\sigma}), \\ E_{21}^{(2)} &= \frac{1}{4}(I - \sin \delta \cos \gamma \sigma_x - \sin \delta \sin \gamma \sigma_y - \cos \delta \sigma_z) = \frac{1}{4}(I - \mathbf{n}^{(2)} \cdot \boldsymbol{\sigma}), \\ E_{12}^{(2)} &= \frac{1}{4}(I - \sin \delta \cos \gamma \sigma_x - \sin \delta \sin \gamma \sigma_y + \cos \delta \sigma_z) = \frac{1}{4}(I + \mathbf{m}^{(2)} \cdot \boldsymbol{\sigma}), \\ E_{22}^{(2)} &= \frac{1}{4}(I + \sin \delta \cos \gamma \sigma_x + \sin \delta \sin \gamma \sigma_y - \cos \delta \sigma_z) = \frac{1}{4}(I - \mathbf{m}^{(2)} \cdot \boldsymbol{\sigma}). \end{aligned} \quad (19.57)$$

Here we have introduced the unit vectors

$$\mathbf{n}^{(2)} = (\sin \delta \cos \gamma, \sin \delta \sin \gamma, \cos \delta), \quad \mathbf{m}^{(2)} = (-\sin \delta \cos \gamma, -\sin \delta \sin \gamma, \cos \delta).$$

The margin $\mathbf{F}^{(2)}$ associated with the detector outputs is obtained by summing over both probe outputs:

$$\begin{aligned} F_1^{(2)} &= E_{11}^{(2)} + E_{12}^{(2)} = \frac{1}{2}(I + \cos \delta \sigma_z), \\ F_2^{(2)} &= E_{21}^{(2)} + E_{22}^{(2)} = \frac{1}{2}(I - \cos \delta \sigma_z). \end{aligned}$$

This is a smeared path observable. The margin $\mathbf{G}^{(2)}$ associated with the probe outputs is obtained by summing over both detection outputs:

$$\begin{aligned} G_1^{(2)} &= E_{11}^{(2)} + E_{21}^{(2)} = \frac{1}{2} I, \\ G_2^{(2)} &= E_{12}^{(2)} + E_{22}^{(2)} = \frac{1}{2} I. \end{aligned}$$

This is a trivial observable.

The fact that these two margins are a smeared path and a trivial observable, respectively, in the present setup can be understood as follows. The entanglement

between probe and photon is devised to establish a strict correlation between the path states $|10\rangle, |01\rangle$ and the pointer states $|p_1\rangle, |p_2\rangle$, for any photon input state ψ_i . This correlation information is not accessible by measuring a probe observable with the eigenstates $|q_1\rangle, |q_2\rangle$ because these are equal weight superpositions of the path marker states. Further, the reduced state of the photon after the coupling has been established is a mixture of the path states, so that any phase relation between these states has been erased. Accordingly, the detector outputs cannot detect any interference indicative of coherence between the path input states, and the only information remaining available about the input is path information.

The third possible marginal input observable $H^{(2)}$ turns out to be the only one sensitive to the parameter γ :

$$\begin{aligned} H_1^{(2)} &= E_{11}^{(2)} + E_{22}^{(2)} = \frac{1}{2}(I + \sin \delta (\cos \gamma \sigma_x + \sin \gamma \sigma_y)) \\ H_2^{(2)} &= E_{12}^{(2)} + E_{21}^{(2)} = \frac{1}{2}(I - \sin \delta (\cos \gamma \sigma_x + \sin \gamma \sigma_y)). \end{aligned}$$

This is a smeared interference observable, the unsharpness being determined by the term $\sin \delta$ and the associated vector being given by $\pm(\cos \gamma, \sin \gamma, 0)$. By varying γ from 0 to 2π , all possible interference observables can be realised. The erasure scheme presented here constitutes a joint unsharp measurement of the path and interference observables represented by $F^{(2)}$ and $H^{(2)}$, respectively.

As shown in Sect. 14.3, the compatibility of the observables $F^{(2)}$ and $H^{(2)}$, which are unbiased qubit observables, is equivalent to the unsharpness trade-off inequality (14.36) or its explicit form (14.34b). Here the relevant terms are $\text{fuzz}(F_1^{(2)})^2 = \sin^2 \delta$, $\text{fuzz}(H_1^{(2)})^2 = \cos^2 \delta$, and $\tilde{e} = (0, 0, \cos \delta)$, $\tilde{f} = \sin \delta(-\cos \gamma, \sin \gamma, 0)$, hence $|\tilde{e} \times \tilde{f}|^2 = 1 - \cos^2 \delta \sin^2 \delta$. So the inequality assumes the form

$$\text{fuzz}(F_1^{(2)})^2 \text{fuzz}(H_1^{(2)})^2 = \sin^2 \delta \cos^2 \delta \geq \cos^2 \delta \sin^2 \delta,$$

which can be equivalently cast as

$$\text{fuzz}(F_1^{(2)})^2 + \text{fuzz}(H_1^{(2)})^2 = 1 \geq 1.$$

We also see that for $\delta = \pi/2$, all four effects $E_{k\ell}$ are fractions of spectral projections of a sharp interference observable; the margin $\{H_1^{(2)}, H_2^{(2)}\}$ becomes a sharp interference observable and the margin $\{F_1^{(2)}, F_2^{(2)}\}$ becomes a trivial observable. Here we have recovered an observation made in the review [25] of Scully, Englert and Walther, that the detector statistics conditional on the probe output readings display interference with perfect visibility. In this way the path information encoded in the probe through its interaction with the photon has been erased, and interference has been reestablished. This has become possible by measuring a probe observable that is complementary to the probe's path indicating observable.

In fact, we have found that for any photon input state ψ , the conditional probabilities for detections at D_1, D_2 given a probe recording of $|q_1\rangle$ (say) are

$$\begin{aligned}\text{prob}(D_1|q_1) &= \frac{\langle\psi|E_{11}^{(2)}\psi\rangle}{\langle\psi|G_1^{(2)}\psi\rangle} = \langle\psi|\frac{1}{2}(I + \cos\gamma\sigma_x + \sin\gamma\sigma_y)\psi\rangle, \\ \text{prob}(D_2|q_1) &= \frac{\langle\psi|E_{21}^{(2)}\psi\rangle}{\langle\psi|G_1^{(2)}\psi\rangle} = \langle\psi|\frac{1}{2}(I - \cos\gamma\sigma_x - \sin\gamma\sigma_y)\psi\rangle.\end{aligned}$$

For $\gamma = 0$ and the input state $\psi = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle)$, this gives $\text{prob}(D_1|q_1) = 1$ and $\text{prob}(D_2|q_1) = 0$. This corresponds to the case of perfect interference fringes. Similarly, for the detector probabilities conditional on $|q_2\rangle$ and the above input eigenstate of σ_x , we obtain probabilities 0 and 1 for D_1 and D_2 , respectively, which are characteristic of interference anti fringes.

This situation is a consequence of the fact that for the above input and $\delta = \pi/2$, $\gamma = 0$, the state Ψ_e and also the total output state $\Psi_f^{\pi/2}$ is an EPR state:

$$\begin{aligned}\Psi_f^{\pi/2} &= \frac{1+i}{\sqrt{2}}\frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(|10\rangle - |01\rangle)\otimes|p_1\rangle + \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle)\otimes|p_2\rangle\right] \\ &= \frac{1+i}{\sqrt{2}}\frac{1}{\sqrt{2}}[|10\rangle\otimes|q_1\rangle - |01\rangle\otimes|q_2\rangle].\end{aligned}$$

Here it is seen that it is the choice of the probe pointer as the observable either with eigenstates $|p_i\rangle$ or $|q_j\rangle$ that decides whether path information or interference are obtained in the interferometer on detection at D_1, D_2 – and this choice can be made *after* the path marking interaction has taken place. However, this choice is between two complementary path-marking probe observables, and their realisations require mutually exclusive experimental settings.

Quantitative Erasure

We have seen that in the erasure setting, where the probe observable has eigenstates $|q_j\rangle$ and $0 < \delta < \pi/2$, the operator measure $\mathbf{E}^{(2)}$ is a joint observable for an unsharp path and an unsharp interference observable. In this case the experiment provides simultaneous information about these noncommuting quantities. In the limit of $\delta = 0$ ($\delta = \pi/2$), the interference margin $\mathbf{H}^{(2)}$ becomes trivial (sharp) and the path margin $\mathbf{F}^{(2)}$ becomes sharp (trivial).

The possibility of obtaining some joint information about both observables, path and interference, can also be achieved by modifying the path marking coupling in such a way that the correlation between the paths and the probe indicator observable is not perfect. This has been described as *quantitative erasure* (e.g., [26]). We show here that quantitative erasure is again an instance of a joint unsharp measurement.

We take the path-marking interaction to be of the same form as before, Eq. (19.50), but now we specify the marker states $|p_1\rangle, |p_2\rangle$ to be nonorthogonal. As pointer states we choose mutually orthogonal unit vectors $|q_1\rangle, |q_2\rangle$, so that we may write

$$|p_1\rangle = \cos \frac{\theta}{2} |q_1\rangle + \sin \frac{\theta}{2} |q_2\rangle, \quad |p_2\rangle = \sin \frac{\theta}{2} |q_1\rangle + \cos \frac{\theta}{2} |q_2\rangle. \quad (19.58)$$

Note that $\langle p_1 | p_2 \rangle = \sin \theta$. The final state (Eq. (19.50)) after the path-marking interaction and passage of the photon through the interferometer is

$$\begin{aligned} \Psi_f^{\delta, \theta} = & \left[\left(\frac{\alpha}{2} \cos \frac{\theta}{2} (e^{i\delta} + 1) - i \frac{\beta}{2} \sin \frac{\theta}{2} (e^{i\delta} - 1) \right) |10\rangle \right. \\ & \left. + \left(i \frac{\alpha}{2} \cos \frac{\theta}{2} (e^{i\delta} - 1) + \frac{\beta}{2} \sin \frac{\theta}{2} (e^{i\delta} + 1) \right) |01\rangle \right] \otimes |q_1\rangle \\ & + \left[\left(\frac{\alpha}{2} \sin \frac{\theta}{2} (e^{i\delta} + 1) - i \frac{\beta}{2} \cos \frac{\theta}{2} (e^{i\delta} - 1) \right) |10\rangle \right. \\ & \left. + \left(i \frac{\alpha}{2} \sin \frac{\theta}{2} (e^{i\delta} - 1) + \frac{\beta}{2} \cos \frac{\theta}{2} (e^{i\delta} + 1) \right) |01\rangle \right] \otimes |q_2\rangle. \end{aligned} \quad (19.59)$$

We determine the measured observable $E^{(3)}$ with effects $E_{k\ell}^{(3)}$ associated with the output projections $Z_{k\ell}^{(3)} = |\varphi_k\rangle\langle\varphi_k| \otimes |q_\ell\rangle\langle q_\ell|$ (see Eq. (19.51)), again via $\langle \Psi_f^{\delta, \theta} | Z_{k\ell}^{(3)} | \Psi_f^{\delta, \theta} \rangle = \langle \psi | E_{k\ell}^{(3)} | \psi \rangle$,

$$\begin{aligned} E_{11}^{(3)} &= \frac{1}{4} [I(1 + \cos \theta \cos \delta) + \sin \delta \sin \theta \sigma_x + (\cos \delta + \cos \theta) \sigma_z] \\ &= \frac{1}{4} [I(1 + \cos \theta \cos \delta) + \mathbf{m}^{(3)} \cdot \boldsymbol{\sigma}], \\ E_{21}^{(3)} &= \frac{1}{4} [I(1 - \cos \theta \cos \delta) - \sin \delta \sin \theta \sigma_x - (\cos \delta - \cos \theta) \sigma_z] \\ &= \frac{1}{4} [I(1 - \cos \theta \cos \delta) - \mathbf{n}^{(3)} \cdot \boldsymbol{\sigma}], \\ E_{12}^{(3)} &= \frac{1}{4} [I(1 - \cos \theta \cos \delta) + \sin \delta \sin \theta \sigma_x + (\cos \delta - \cos \theta) \sigma_z] \\ &= \frac{1}{4} [I(-\cos \theta \cos \delta) + \mathbf{n}^{(3)} \cdot \boldsymbol{\sigma}], \\ E_{22}^{(3)} &= \frac{1}{4} [I(1 + \cos \theta \cos \delta) - \sin \delta \sin \theta \sigma_x - (\cos \delta + \cos \theta) \sigma_z] \\ &= \frac{1}{4} [I(1 + \cos \theta \cos \delta) - \mathbf{m}^{(3)} \cdot \boldsymbol{\sigma}], \end{aligned} \quad (19.60)$$

where

$$\begin{aligned} \mathbf{m}^{(3)} &= (\sin \delta \sin \theta, 0, \cos \delta + \cos \theta), \\ \mathbf{n}^{(3)} &= (\sin \delta \sin \theta, 0, \cos \delta - \cos \theta). \end{aligned} \quad (19.61)$$

The three marginal observables are determined as before:

$$\begin{aligned} F_1^{(3)} &= E_{11}^{(3)} + E_{12}^{(3)} = \frac{1}{2} (I + \sin \delta \sin \theta \sigma_x + \cos \delta \sigma_z), \\ F_2^{(3)} &= E_{22}^{(3)} + E_{21}^{(3)} = \frac{1}{2} (I - \sin \delta \sin \theta \sigma_x - \cos \delta \sigma_z), \\ G_1^{(3)} &= E_{11}^{(3)} + E_{21}^{(3)} = \frac{1}{2} (I + \cos \theta \sigma_z), \\ G_2^{(3)} &= E_{12}^{(3)} + E_{22}^{(3)} = \frac{1}{2} (I - \cos \theta \sigma_z), \\ H_1^{(3)} &= E_{11}^{(3)} + E_{22}^{(3)} = \frac{1}{2} I (1 + \cos \theta \cos \delta), \\ H_2^{(3)} &= E_{12}^{(3)} + E_{21}^{(3)} = \frac{1}{2} I (1 - \cos \theta \cos \delta). \end{aligned} \quad (19.62)$$

For $\delta = \pi/2$, the first margin (corresponding to the detector statistics) becomes an unsharp interference observable, while the second margin (corresponding to the probe output statistics) is an unsharp path observable. In both cases the unsharpness is determined by the parameter θ . For $\theta = \pi/2$ we have $|p_1\rangle = |p_2\rangle$, so that no path marking arises and the state $\Psi_f^{\delta,\theta}$ is not entangled; accordingly, the only nontrivial margin $F^{(3)}$ is a sharp interference observable. In contrast, for $\theta = 0$ or π , the only nontrivial margin $G^{(3)}$ is a sharp path observable; here $|p_1\rangle, |p_2\rangle$ are orthogonal and the path marking destroys all interference.

19.7 Exercises

1. Determine the phase space observable defined by the sequential measurement scheme for position and momentum in Sect. 19.1, and verify that its margins are the unsharp position and momentum observables given in Eq. (19.5)
2. Verify that the positive operator measure defined by the Arthurs–Kelly model is the phase space observable given in Eq. (19.4) with (19.14) and the instrument (19.11) is given by (19.13).
3. Verify the following equations: $e^{-i\theta N} a e^{i\theta N} = e^{i\theta} a$, $\theta \in \mathbb{R}$, $a | \alpha\rangle = \alpha | \alpha\rangle$, and $\langle \alpha | N | \alpha\rangle = |\alpha|^2$ for all $\alpha \in \mathbb{C}$. Let $\psi, \psi' \in L^2(\mathbb{R})$. Show that, if $\langle \alpha | \psi\rangle = \langle \alpha | \psi'\rangle$ for all $\alpha \in \mathbb{C}$, then $\psi = \psi'$.
4. Determine the approximate values of a spin-1/2 quantity $S_a = \mathbf{a} \cdot \boldsymbol{\sigma} = \frac{1}{2} \varrho_a - \frac{1}{2} \varrho_{-a}$ [see Exercises 1 and 2, Sect. 9.6] obtained in a standard measurement ($L^2(\mathbb{R})$, $e^{iS_a \otimes P}$, ϕ , $\mathbf{Q}$).
5. Show that the set of all possible weak values of a spin quantity S_a exhaust the whole complex plane $\mathbb{C}$.
6. Verify that the description of the state evolution for a Mach–Zehnder interferometer given in Sect. 19.6 entails a unitary operator that effects the transition $\Psi_e \mapsto \Psi_f^\delta$ as given in Eq. (19.50).
7. Verify the various expressions of the interferometric output state (19.50) given in Eqs. (19.53), (19.56) and (19.59).
8. Deduce the three interferometric observables, Eqs. (19.55), (19.57), and (19.60).

References

1. Arthurs, E., Kelly, J.: On the simultaneous measurements of a pair of conjugate observables. *Bell Syst. Tech.* **44**, 725–729 (1965)
2. Busch, P.: Indeterminacy relations and simultaneous measurements in quantum theory. *Int. J. Theor. Phys.* **24**(1), 63–92 (1985) (Based on P. Busch, Doctoral Dissertation, University of Cologne (1982))
3. Raymer, M.G.: Uncertainty principle for joint measurement of noncommuting variables. *Am. J. Phys.* **62**, 986–993 (1994)

4. Törmä, P., Stenholm, S., Jex, I.: Measurement and preparation using two probe modes. *Phys. Rev. A* **52**, 4812–4822 (1995)
5. Taylor, M.E.: *Noncommutative Harmonic Analysis. Mathematical Surveys and Monographs*, vol. 22. American Mathematical Society, Providence (1986)
6. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Physik* **43**, 172–198 (1927)
7. Loudon, R.: *The Quantum Theory of Light*. Oxford University Press, Oxford (1973)
8. Perina, J.: *Coherence of Light*. D. Reidel Publishing Company, Dordrecht (1985)
9. Leonhardt, U.: *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge (1997)
10. Busch, P., Grabowski, M., Lahti, P.: *Operational Quantum Physics*, 2nd edn. *Lecture Notes in Physics Monographs*, vol. 31. Springer, Berlin (1997)
11. Dirac, P.A.M.: The quantum theory of the emission and absorption of radiation. *Proc. R. Soc. Lond. Ser. A* **114**(767), 243–265 (1927)
12. Schumaker, B.L.: Quantum mechanical pure states with gaussian wave functions. *Phys. Rep.* **135**(6), 317–408 (1986)
13. Lahti, P., Pellonpää, J.-P., Schultz, J.: Realistic eight-port homodyne detection and covariant phase space observables. *J. Mod. Opt.* **57**(13), 1171–1179 (2010)
14. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
15. Kiukas, J., Lahti, P.: A note on the measurement of phase space observables with an eight-port homodyne detector. *J. Mod. Opt.* **55**(12), 1891–1898 (2008)
16. Werner, R.: Quantum harmonic analysis on phase space. *J. Math. Phys.* **25**(5), 1404–1411 (1984)
17. Lahti, P., Pellonpää, J.-P.: Covariant phase observables in quantum mechanics. *J. Math. Phys.* **40**(10), 4688–4698 (1999)
18. Pellonpää, J.-P., Schultz, J., Paris, M.G.A.: Balancing efficiencies by squeezing in realistic eight-port homodyne detection. *Phys. Rev. A* **83**, 043818 (2011)
19. Carmeli, C., Heinosaari, T., Pellonpää, J.-P., Toigo, A.: Optimal covariant measurements: the case of a compact symmetry group and phase observables. *J. Phys. A* **42**(14), 145304, 18 (2009)
20. Pellonpää, J.-P., Schultz, J.: Measuring the canonical phase with phase-space measurements. *Phys. Rev. A* **88**, 012121 (2013)
21. Albini, P., De Vito, E., Toigo, A.: Quantum homodyne tomography as an informationally complete positive-operator-valued measure. *J. Phys. A* **42**(29), 295302, 12 (2009)
22. Lahti, P., Pellonpää, J.-P.: Continuous variable tomographic measurements. *Phys. Lett. A* **373**(38), 3435–3438 (2009)
23. Pellonpää, J.-P.: Quantum tomography, phase-space observables and generalized Markov kernels. *J. Phys. A* **42**(46), 465303, 18 (2009)
24. Busch, P., Shilladay, C.: Complementarity and uncertainty in Mach-Zehnder interferometry and beyond. *Phys. Rep.* **435**(1), 1–31 (2006)
25. Scully, M.O., Englert, B.-G., Walther, H.: Quantum optical tests of complementarity. *Nature* **351**, 111–116 (1991)
26. Englert, B.-G., Bergou, J.A.: Quantitative quantum erasure. *Opt. Commun.* **179**(1–6), 337–355 (2000)

Part IV

Foundations

Chapter 20

Bell Inequalities and Incompatibility

The term *Bell inequalities* is a collective name for a class of inequalities involving probabilities that are used to demonstrate the nonlocal nature of correlations between quantum systems in certain entangled states. From a historical point of view, Bell inequalities are derived from an assumption of *local realism* which is characteristic of classical physical systems. A violation of a Bell inequality by a given physical system will then be an indication that one of the assumptions of realism or locality must be dropped. In the case of quantum mechanical violations of Bell inequalities, it is widely understood that the assumption of realism is compatible with quantum mechanics, so that the conclusion is drawn that this theory possesses nonlocal features.

A demonstration of violations of Bell inequalities in quantum mechanics typically involves the use of incompatible observables. There is an intriguing interplay between the degree of Bell inequality violation and the degree of incompatibility of the observables involved, which is the subject of this chapter.

20.1 Bell Inequalities and Compatibility: General Observations

Experimental demonstrations of Bell inequalities for entangled quantum systems are based on sets of measurements that are incompatible with each other. In the original formulation of his inequality, John Bell [1] considered a triple of dichotomic observables in the context of an Einstein–Podolsky–Rosen [2] experiment, where each pair selected from these observables has a joint probability distribution in a given state. The original Bell inequality then follows from the condition that there is a joint probability distribution for all three observables in the given state such that the pair distributions arise as margins. This condition is, in turn, a consequence of an assumption of *locality* in a situation where the pairs of observables are being

measured at separate locations, so that the measurements should not have any direct influence on each other. This assumption is just one of the basic composition rules of compound systems, see Sect. 9.5.

Clauser, Horne, Shimony and Holt [3] extended Bell's consideration to one that can be tested experimentally; thus, at each of two separate measurement locations, one of a choice of two dichotomic observables is measured. For the resulting four pair probability distributions to possess a common joint probability over all four observables involved, they showed that a suitably adapted form of Bell inequality has to be satisfied. This inequality is known as the *CHSH inequality*.

In the classic Einstein–Podolsky–Rosen–Bell (EPR–Bell) experiment, the pair probabilities under consideration arise from an entangled quantum state, normally the singlet state of a pair of qubits. The violation of a Bell inequality is then taken as an indication of *nonclassical correlations* and, since the experiments are performed in separated locations, of *quantum nonlocality*.

Remark 20.1 It was with respect to the question of joint probabilities that as early as 1854, George Boole [4, 5] discovered forms of Bell inequalities as a consequence of what he referred to as *conditions of the possibility of experience* (with a bow to Immanuel. Kant). This historic fact was noted in 1989 by I. Pitowsky [6, 7]. $\triangleleft$

Fine [8, 9] proved in 1982 that a complete set of CHSH inequalities is in fact necessary *and sufficient* for a family of pair probabilities involving a total of four dichotomic random variables to be embeddable into a quadruple joint probability. One can ask the analogous question of what additional conditions must be satisfied for four pairwise compatible dichotomic quantum observables to be jointly measurable. A necessary condition is that the joint observables for the pairs obey a set of operator Bell-CHSH inequalities. However, these inequalities are no longer sufficient for the overall joint measurability; this is related to the fact that the required joint probabilities must satisfy the additional constraint of being of the correct quantum mechanical form for each state.

Thus the joint measurability of a whole set of four quantum observables in a Bell experiment implies the satisfaction of Bell inequalities for all pairs of margins and for all quantum states. This, in turn, means that the detection of Bell inequality violations requires the inclusion of observables that are not jointly measurable. We will see that there is a converse (in a certain sense) to this implication: as soon as the set of observables in a Bell experiment contains incompatible elements, it is possible to achieve maximal violations of some CHSH inequalities [10].

The idea of hypothetically considering the existence of a joint observable for a set of four observables in an EPR–Bell experiment was formulated and explored by Abu-Zeid and de Muynck in 1984 [11], with the conclusion that the violation of Bell inequalities reflects the nonexistence of such a joint observable in the case of noncommuting sharp spin- $\frac{1}{2}$ observables. The issue of formulating and exploring the meaning and role of operator forms of Bell inequalities for unsharp spin observables was addressed in [12]. Moreover, it was noted there that it is possible to suppress

the violation of a given Bell inequality even for incompatible observables if these observables are chosen to be sufficiently unsharp. This observation was taken up and generalised by Kar and Roy [13, 14].

20.2 Bell Inequalities and Joint Probabilities

We begin with an explicit formulation of a complete set of Bell inequalities as a necessary and sufficient condition for the embeddability of a family of probabilities. We consider four probability spaces $(\Omega_i, \mathcal{A}_i, \lambda_i)$, $i = 1, \dots, 4$, and four random variables (i.e. measurable functions) $X_i : \Omega_i \rightarrow \{-1, 1\}$. These random variables induce four probability measures $\mu_i : 2^{\{-1, 1\}} \rightarrow [0, 1]$ via $\mu_i(A) = \lambda_i(X_i^{-1}(A))$, $A \subset \{-1, 1\}$. (The random variables are actually superfluous at this stage; we may regard the measures μ_i as primitive entities in themselves.) Assume that $\mu_{13}, \mu_{14}, \mu_{23}, \mu_{24}$ are probability measures defined on the power set $2^{\{-1, 1\} \times \{-1, 1\}}$ of $\{-1, 1\} \times \{-1, 1\}$ such that $\mu_1, \dots, \mu_4$ are their marginal measures in the sense that $\mu_{ij}(A \times \{-1, 1\}) = \mu_i(A)$ and $\mu_{ij}(\{-1, 1\} \times A) = \mu_j(A)$ for $i = 1, 2, j = 3, 4$ and $A \subset \{-1, 1\}$. The measures μ_{ij} are of course not uniquely determined: the product measure gives just one possible choice. In our earlier notation and terminology μ_{ij} is an element of the set $\Gamma(\mu_i, \mu_j)$ of all couplings (joint probabilities) of μ_i and μ_j .

We ask under what conditions these probability measures μ_{ij} are marginal probabilities of some probability measure on the Cartesian product set $\{-1, 1\}^4$.

For any probability measure μ defined on the power set of $\{-1, 1\}^4$ and any $i = 1, 2, j = 3, 4$ we consider certain marginal measures $\theta_1(\mu), \theta_2(\mu), \theta_3(\mu), \theta_4(\mu), \theta_{13}(\mu), \theta_{14}(\mu), \theta_{23}(\mu), \theta_{24}(\mu)$ writing e.g.

$$[\theta_1(\mu)](A) = \mu(A \times \{-1, 1\} \times \{-1, 1\} \times \{-1, 1\})$$

for $A \subset \{-1, 1\}$ and

$$[\theta_{13}(\mu)](A) = \mu(\{(x, y, z, w) \in \{-1, 1\}^4 \mid (x, z) \in A\})$$

for $A \subset \{-1, 1\}^2$. Clearly $\theta_{ij} \in \Gamma(\theta_i(\mu), \theta_j(\mu))$.

For the rest of this section we introduce some notation to get an easily readable appearance for our formulas. We write $\mu_i(\{1\}) = p_i$, $\mu_i(\{-1\}) = p_{\bar{i}} = 1 - p_i$, $\mu_{ij}(\{(1, 1)\}) = p_{ij}$, $\mu_{ij}(\{(1, -1)\}) = p_{i\bar{j}}$, $\mu_{ij}(\{(-1, 1)\}) = p_{\bar{i}j}$, $\mu_{ij}(\{(-1, -1)\}) = p_{\bar{i}\bar{j}}$. Here e.g. the notation $p_{\bar{i}j}$ for $\mu_{ij}(\{-1, 1\})$ is a unit with no separate role for $\bar{i}$ or j . Throughout this section we retain this notational convention connecting any μ_i, μ_{ij} with probabilities denoted by $p_i, p_{\bar{i}}, p_{ij}, p_{i\bar{j}}, p_{\bar{i}j}, p_{\bar{i}\bar{j}}$ as above.

Theorem 20.1 *Let $\mu_1, \dots, \mu_4 : 2^{\{-1,1\}} \rightarrow [0, 1]$ be probability measures with fixed couplings $\mu_{13}, \mu_{14}, \mu_{23}, \mu_{24}$ (with the p -notations for their probabilities introduced above). The following two conditions are equivalent.*

- (i) *There is a probability measure μ defined on the power set of $\{-1, 1\}^4$ such that $\mu_1 = \theta_1(\mu), \dots, \mu_4 = \theta_4(\mu)$ and $\mu_{ij} = \theta_{ij}(\mu)$ for $i = 1, 2, j = 3, 4$.*
- (ii) *The following set of Bell-CHSH inequalities holds:*

$$\begin{aligned} 0 &\leq p_{1\bar{3}} + p_{\bar{1}4} - p_{24} + p_{23} \leq 1, \\ 0 &\leq p_{1\bar{4}} + p_{\bar{1}3} - p_{23} + p_{24} \leq 1, \\ 0 &\leq p_{2\bar{3}} + p_{\bar{2}4} - p_{14} + p_{13} \leq 1, \\ 0 &\leq p_{2\bar{4}} + p_{\bar{2}3} - p_{13} + p_{14} \leq 1, \end{aligned}$$

or equivalently:

$$\begin{aligned} 0 &\leq p_1 + p_4 - p_{13} - p_{14} - p_{24} + p_{23} \leq 1, \\ 0 &\leq p_1 + p_3 - p_{13} - p_{14} - p_{23} + p_{24} \leq 1, \\ 0 &\leq p_2 + p_4 - p_{23} - p_{14} - p_{24} + p_{13} \leq 1, \\ 0 &\leq p_2 + p_3 - p_{13} - p_{23} - p_{24} + p_{14} \leq 1. \end{aligned}$$

This is a well-known result, originating from [8]. A somewhat modified elementary proof using direct calculation may be found e.g. in [15]. One proof method is based on so-called correlation polytopes [16].

The formulation of Bell inequalities given here is based on the pair probabilities directly. There is an equivalent formulation, used by Clauser, Horne, Shimony and Holt, in terms of expectation values. If we use random variables, there is the issue of where they are defined. We take $Y_1, Y_2, Y_3, Y_4 : \{-1, 1\}^4 \rightarrow \{-1, 1\}$ to be simply the coordinate projections. We consider the expectation functional $\langle \cdot \rangle$ with respect to a fixed probability measure μ which appears in the statement of Theorem 20.1 (i). For example for the pair Y_1, Y_3 we get

$$\langle Y_1 Y_3 \rangle = p_{13} - p_{1\bar{3}} - p_{\bar{1}3} + p_{\bar{1}\bar{3}}. \quad (20.1)$$

It is straightforward to obtain an equivalent set of CHSH inequalities to the above sets. They are of the form

$$|\langle Y_1 Y_3 \rangle + \langle Y_1 Y_4 \rangle + \langle Y_2 Y_3 \rangle - \langle Y_2 Y_4 \rangle| \leq 2. \quad (20.2)$$

This inequality is equivalent to the second of the second set given in the theorem. The remaining three inequalities are obtained by swapping the location of the minus sign.

There is an interesting alternative assumption which already leads to Bell's inequalities.

Theorem 20.2 Let $\mu_1, \mu_2, \mu_3, \mu_4$ and $\mu_{13}, \mu_{14}, \mu_{23}, \mu_{24}$ be as in Theorem 20.1. Suppose that κ, ν are probability measures defined on the power set of $\{-1, 1\}^3$ such that

$$\begin{aligned}\mu_{13}(A) &= \kappa(\{(x, y, z) \in \{-1, 1\}^3 \mid (x, z) \in A\}) \\ \mu_{23}(A) &= \kappa(\{(x, y, z) \in \{-1, 1\}^3 \mid (y, z) \in A\}) \\ \mu_{14}(A) &= \nu(\{(x, y, z) \in \{-1, 1\}^3 \mid (x, z) \in A\}) \\ \mu_{24}(A) &= \nu(\{(x, y, z) \in \{-1, 1\}^3 \mid (y, z) \in A\})\end{aligned}$$

for any $A \subset \{-1, 1\}^2$. Then the CHSH inequalities of the form (20.2) must hold.

Proof In the proof we let $Y_i, i = 1, \dots, 4$, be the coordinate variables. Writing all expressions of the form (20.1) appearing in (20.2) in terms of the available pair probabilities, and then using the given triple probabilities (extending our earlier notation), we obtain after some algebra

$$\begin{aligned}\langle Y_1 Y_3 \rangle + \langle Y_2 Y_3 \rangle &= 2(p_{123} + p_{\bar{1}\bar{2}3} - p_{1\bar{2}3} - p_{\bar{1}23}) \\ &= 4(p_{123} + p_{\bar{1}\bar{2}3}) - 2(p_{12} + p_{\bar{1}2}), \\ \langle Y_1 Y_4 \rangle - \langle Y_2 Y_4 \rangle &= 2(p_{1\bar{2}4} - p_{\bar{1}\bar{2}4} - p_{124} + p_{\bar{1}24}) \\ &= 4(p_{1\bar{2}4} + p_{\bar{1}24}) - 2(p_{12} + p_{\bar{1}2}).\end{aligned}$$

Adding these expressions gives

$$\begin{aligned}\langle Y_1 Y_3 \rangle + \langle Y_1 Y_4 \rangle + \langle Y_2 Y_3 \rangle - \langle Y_2 Y_4 \rangle \\ = 4(p_{123} + p_{\bar{1}\bar{2}3} + p_{1\bar{2}4} + p_{\bar{1}24}) - 2 \in [-2, 2].\end{aligned}$$

The bounds ± 2 are obtained due to the fact that the expression in brackets has a value in the interval $[0, 1]$:

$$0 \leq p_{123} + p_{\bar{1}\bar{2}3} + p_{1\bar{2}4} + p_{\bar{1}24} \leq p_{12} + p_{\bar{1}2} + p_{12} + p_{\bar{1}2} = 1. \quad \square$$

Example 20.1 Consider sharp observables A, B and C, D of two qubit systems S_1 and S_2 , respectively. For any state ρ of the compound system $S_1 + S_2$, let ρ_1 and ρ_2 be the reduced states of S_1 and S_2 , respectively. Consider the probability measures $\mu_1 = A_{\rho_1}, \mu_2 = B_{\rho_1}, \mu_3 = C_{\rho_2}, \mu_4 = D_{\rho_2}$ with the couplings $\mu_{13} = (A \otimes C)_\rho, \mu_{14} = (A \otimes D)_\rho, \mu_{23} = (B \otimes C)_\rho$, and $\mu_{24} = (B \otimes D)_\rho$. It is an easy exercise to choose A, B, C, D and ρ such that the inequalities of Theorem 20.1 are violated. Verify also that if ρ is separable, that is, a mixture of decomposable states $\rho_1 \otimes \rho_2$ then these inequalities are satisfied.

20.3 Bell Inequality Violation and Nonlocality

Before we turn to the connection between joint measurability and Bell inequalities in quantum mechanics, we recall the connection between Bell inequality violations and the issue of nonlocality in the context of EPR–Bell experiments.

In such experiments one considers a bipartite system composed of two subsystems S_a and S_b which are prepared together and then allowed to propagate until they are measured in separate locations. There is a choice of two (dichotomic) measurements on each site S_a and S_b , with values $a_k, b_\ell \in \{-1, 1\}$ for $k, \ell = 1, 2$, respectively. We let $\mu^{(k\ell)}$ be the corresponding probability measures with the respective margins $\mu_a^{(k)}$ and $\mu_b^{(\ell)}$ as, for instance, in Example 20.1. Following Einstein, Podolsky and Rosen (EPR), one may adopt the assumption of *realism*, which means that properties that can be ascertained with certainty without changing the system are actually realised. This can be modelled by assuming that there is a hypothetical *hidden variable*, with associated probability (distribution) space (Λ, Σ, ν) , whose value determines the outcomes of measurements. The observed randomness in the occurrence of pairs of outcomes (a_k, b_ℓ) of the joint measurements labelled by (k, ℓ) is hypothetically explained by the existence of the probability measure ν together with functions

$$\Lambda \ni \lambda \mapsto f^{(k\ell)}(a_k, b_\ell | \lambda) \in [0, 1]$$

such that the joint probabilities $\mu^{(k\ell)}(a_k, b_\ell)$ are obtained as

$$\mu^{(k\ell)}(a_k, b_\ell) = \int_{\Lambda} f^{(k\ell)}(a_k, b_\ell | \lambda) d\nu(\lambda), \quad k, \ell = 1, 2.$$

The reason for the applicability of the EPR reality criterion lies in the fact that the measurements on the two systems are spatially well separated, so that, classically, it is reasonable to assume that the measurement outcomes and their statistics at location a do not depend on the choice of measurements made at location b . This *locality assumption* is implemented by requiring the factorisability condition

$$f^{(k\ell)}(a_k, b_\ell | \lambda) = f_a^{(k)}(a_k | \lambda) f_b^{(\ell)}(b_\ell | \lambda),$$

where $f_a^{(k)}$ and $f_b^{(\ell)}$ are the a and b margins of $f^{(k,\ell)}$, that is, for instance $f_a^{(k)}(a_k) = f^{(k\ell)}(a_k, 1 | \lambda) + f^{(k\ell)}(a_k, -1 | \lambda)$. This allows one to define the product $f = f_a^{(1)} f_a^{(2)} f_b^{(1)} f_b^{(2)}$ which induces a probability measure μ on the power set of $\{-1, 1\}^4$, with the values

$$\mu(\{(a_1, a_2, b_1, b_2)\}) = \int_{\Lambda} f(a_1, a_2, b_1, b_2 | \lambda) d\nu(\lambda).$$

It is immediate to check that $\mu^{(k\ell)}, k, \ell = 1, 2$ are margins of μ .

Thus, the assumption that the reality criterion and locality apply to an experiment as described here entails the quadruple joint probability and, via Theorem 20.1, the necessary validity of Bell inequalities. Conversely, if an experiment shows a violation of a Bell inequality, this forces the conclusion that the systems under investigation display nonlocal correlations if the realism assumption is held up.

As is well known, EPR–Bell experiments were carried out with qubits and were found to display violations of Bell inequalities; and these violations were in line with predictions deduced from the quantum mechanical description of the systems at hand as given, see Example 20.1. Since quantum mechanics admits *ideal* measurements (Sect. 10.6) that do not disturb eigenstates of the measured quantities, the reality criterion is fulfilled, and so the conclusion is that entangled quantum systems can show nonlocal behaviour—albeit in a way that does not allow the nonlocal correlations to be used for superluminal signalling.

20.4 Bell Inequality Violation and Incompatibility

We consider here only finite level systems S_a and S_b . For simplicity we assume both systems to be of similar types (but not identical), with the associated Hilbert spaces $\mathcal{H}_d$ having dimension d . For example, they could both carry spin of magnitude $s = (d - 1)/2$. The total system is prepared in such a way that the state in $\mathcal{H}_d \otimes \mathcal{H}_d$ is an entangled state. The observables being measured are chosen from a pair of binary observables with values ± 1 in $\mathcal{H}_d$ for each subsystem, say A_1, A_2 for S_a and B_1, B_2 for S_b . We identify the observables with the sets of effects associated with the possible outcomes ± 1 , so that we can write for $k, \ell = 1, 2$

$$\begin{aligned} A_k &= \{A_{k+}, A_{k-}\}, \quad B_\ell = \{B_{\ell+}, B_{\ell-}\}, \\ \text{where } A_{k+} + A_{k-} &= I, \quad B_{\ell+} + B_{\ell-} = I. \end{aligned} \tag{20.3}$$

We are now ready to introduce the *Bell operator* $\mathbb{B}$ for the operators

$$A_k = A_{k+} - A_{k-}, \quad B_\ell = B_{\ell+} - B_{\ell-}, \quad k, \ell = 1, 2,$$

namely,

$$\begin{aligned} \mathbb{B} &= A_1 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_1 - A_2 \otimes B_2 \\ &= A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2). \end{aligned}$$

Theorem 20.3 (Tsirelson Bound) *The Bell operator satisfies the following operator inequality:*

$$-2\sqrt{2} I \leq \mathbb{B} \leq 2\sqrt{2} I.$$

In the proof we will make use of the following result.

Lemma 20.1 *Let E, F be two effect operators in a Hilbert space $\mathcal{H}$. The operators $\hat{E}, \hat{F}$ in $\mathcal{H} \otimes \mathbb{C}^2$, defined as*

$$\begin{aligned}\hat{E} &= E \otimes \frac{1}{2}(I + \sigma_3) + (I - E) \otimes \frac{1}{2}(I - \sigma_3) + E^{1/2}(I - E)^{1/2} \otimes \sigma_1, \\ \hat{F} &= F \otimes \frac{1}{2}(I + \sigma_3) + (I - F) \otimes \frac{1}{2}(I - \sigma_3) + F^{1/2}(I - F)^{1/2} \otimes \sigma_1,\end{aligned}\quad (20.4)$$

(where σ_1, σ_3 are Pauli operators) are Naimark dilations of E, F , with the projection $P = I \otimes \frac{1}{2}(I + \sigma_3)$ and the identification $\mathcal{H} \equiv \mathcal{H} \otimes [\varphi_3]$, with $[\varphi_3]$ denoting the range of the projection $\frac{1}{2}(I + \sigma_3)$.

The proof of this lemma is left as an exercise.

Proof (of Theorem 20.3) Consider Naimark dilations $\tilde{\mathbf{A}}_k$ of the observables $\mathbf{A}_k$ in a common Hilbert space $\mathcal{H}_a \supset \mathcal{H}_d$ and similarly Naimark dilation $\tilde{\mathbf{B}}_\ell$ of $\mathbf{B}_\ell$ in $\mathcal{H}_b \supset \mathcal{H}_d$, with projections P and Q , respectively. Let

$$\tilde{\mathbb{B}} = \tilde{A}_1 \otimes (\tilde{B}_1 + \tilde{B}_2) + \tilde{A}_2 \otimes (\tilde{B}_1 - \tilde{B}_2)$$

denote the Bell operator associated with the operators $\tilde{A}_k = \tilde{A}_{k+} - \tilde{A}_{k-}$ and $\tilde{B}_\ell = \tilde{B}_{\ell+} - \tilde{B}_{\ell-}$. Then the following equation holds:

$$\langle \Phi | \tilde{\mathbb{B}} \Phi \rangle = \left\langle \Phi | \tilde{\mathbb{B}} \Phi \right\rangle \quad \text{for all } \Phi \in \mathcal{H}_d \otimes \mathcal{H}_d \subset \mathcal{H}_a \otimes \mathcal{H}_b.$$

It follows that

$$\|\tilde{\mathbb{B}}\| \leq \|\tilde{\mathbb{B}}\| = 2 \left[1 + \frac{1}{4} \|[\tilde{A}_1, \tilde{A}_2]\| \|[\tilde{B}_1, \tilde{B}_2]\| \right]^{\frac{1}{2}} \leq 2\sqrt{2}. \quad (20.5)$$

The expression for the norm of $\tilde{\mathbb{B}}$ is a straightforward consequence of the fact that $\tilde{A}_k^2 = I$ and $\tilde{B}_\ell^2 = I$, which entails [17]:

$$\tilde{\mathbb{B}}^2 = 4I \otimes I - [\tilde{A}_1, \tilde{A}_2] \otimes [\tilde{B}_1, \tilde{B}_2].$$

The bound $2\sqrt{2}$ for the Bell operator is the Tsirelson bound. This bound is a consequence of the fact that $\|[\tilde{A}_1, \tilde{A}_2]\| \leq 2\|\tilde{A}_1\| \|\tilde{A}_2\| \leq 2$, and similarly $\|[\tilde{B}_1, \tilde{B}_2]\| \leq 2$. $\square$

Remark 20.2 The Tsirelson bound can be saturated by choosing, for example, operators $\Sigma_1, \Sigma_2, \Sigma_3$ with the properties of the Pauli matrices, that is, $\Sigma_1^2 = \Sigma_2^2 = I$ and $[\Sigma_1, \Sigma_2] = 2i\Sigma_3$. A realisation of such operators in $\mathcal{H}_d$ with $d \geq 2$ is obtained by choosing a 2-dimensional subspace in which one defines the standard Pauli operators $\sigma_1, \sigma_2, \sigma_3$ and putting $\Sigma_k = \sigma_k \oplus I_{d-2}$, where I_{d-2} is the projection onto the orthocomplement of the chosen 2-dimensional subspace. We then have $\|[\Sigma_1, \Sigma_2]\| = 2\|\Sigma_3\| = 2$. $\triangleleft$

The inequality (20.5) shows that satisfaction of the CHSH inequality is guaranteed as soon as only one of the local pairs of observables are compatible. Violation of Bell inequalities requires choices of incompatible pairs on both sides. This is in agreement with the observation made in Theorem 20.2.

Theorem 20.4 *For pairs of dichotomic observables A_1, A_2 and B_1, B_2 associated with systems S_a and S_b , the CHSH inequality is satisfied if one of the pairs is compatible. If A_1, A_2 are compatible, then the triples A_1, A_2, B_ℓ ($\ell = 1, 2$) are compatible.*

Proof If, say, A_1 and A_2 are compatible, these two observables have a joint Naimark extension such that $[A_1, \tilde{A}_2] = 0$; it follows from (20.5) that $\|\mathbb{B}\| \leq \|\tilde{\mathbb{B}}\| = 2$, hence the CHSH inequality is satisfied. Furthermore, then the triples A_1, A_2, B_ℓ are compatible as well, and a joint observable is given as $G_{12\ell} = \{A_{i,j}^{(12)} \otimes B_n^\ell : i, j, n = \pm\}$, where $A_{12} = \{A_{i,j}^{(12)} : i, j = \pm\}$ is a joint observable for A_1, A_2 . $\square$

Using this theorem, one can infer a bound on the degree of incompatibility of any pair of effects (or dichotomic observables) in $\mathcal{H}_d$ from the Tsirelson bound. As shown in Remark 11.4, any pair A_1, A_2 can be smeared by way of mixing each observable with a trivial observable, such that the resulting observables are compatible. The degree of smearing required can be taken as a measure of the degree of incompatibility between A_1, A_2 . We define their smeared versions $A_1^{(\lambda)}, A_2^{(\lambda)}$, where

$$A_{k\pm}^{(\lambda)} = \lambda A_{k\pm} + (1 - \lambda) \frac{1}{2} I.$$

For sufficiently small but positive λ , these observables are compatible. This is trivially the case if $\lambda \leq \frac{1}{2}$, see Proposition 11.6. A tight bound is established in the result of Banik et al. [18] quoted in Remark 11.4: for any pair of dichotomic observables A_1, A_2 in a finite dimensional Hilbert space there exists a value $\lambda \geq 1/\sqrt{2}$ such that the smeared observables $A_1^{(\lambda)}, A_2^{(\lambda)}$ are compatible.

Applying this result, we obtain an alternative derivation of the Tsirelson bound. Indeed, suppose that dichotomic observables A_1, A_2 have compatible smearings with a particular value of $\lambda \geq 1/\sqrt{2}$. Then, the pairs $A_1^{(\lambda)}, A_2^{(\lambda)}$ and B_1, B_2 obey the CHSH inequality with associated Bell operator $\mathbb{B}^{(\lambda)}$. But this operator is related to the Bell operator $\mathbb{B}$ for the pairs A_1, A_2 and B_1, B_2 as

$$\mathbb{B}^{(\lambda)} = A_1^{(\lambda)} \otimes (B_1 + B_2) + A_2^{(\lambda)} \otimes (B_1 - B_2) = \lambda \mathbb{B} \leq 2I$$

since $A_k^{(\lambda)} = A_{k+}^{(\lambda)} - A_{k-}^{(\lambda)} = \lambda(A_{k+} - A_{k-})$. It follows that $\mathbb{B} \leq (2/\lambda)I \leq 2\sqrt{2}I$.

Remark 20.3 We have seen that the use of incompatible dichotomic observables is necessary for a violation of the CHSH inequality. In [10] it is shown that conversely, for any pair of incompatible dichotomic observables A_1, A_2 a pair of dichotomic observables B_1, B_2 can be found such that a maximal violation of the CHSH inequality arises.

One may ask how much smearing is enough to eliminate Bell inequality violations without necessarily going all the way to compatible pairs of local observables. It is not hard to find the answer for the case of pairs of qubits.

We will consider unbiased effects which all have the same degrees of unsharpness. In the qubit case these are of the form $E^{(1,\lambda\mathbf{n})} = \frac{1}{2}(I + \lambda\mathbf{n} \cdot \boldsymbol{\sigma})$, where $\mathbf{n}$ is a unit vector. First we recall the criterion (14.34c) of the joint measurability of two effects $E^{(1,\lambda\mathbf{n}_1)}, E^{(1,\lambda\mathbf{n}_2)}$:

$$\lambda(\|\mathbf{n}_1 + \mathbf{n}_2\| + \|\mathbf{n}_1 - \mathbf{n}_2\|) \leq 2 \quad (20.6)$$

The term in brackets has maximal value $2\sqrt{2}$, which is assumed for $\mathbf{n}_1 \perp \mathbf{n}_2$. Hence for (20.6) to be satisfied for all pairs of directions $\mathbf{n}_1, \mathbf{n}_2$ the value of λ need not be smaller than

$$\lambda_{opt} = \frac{1}{\sqrt{2}}. \quad (20.7)$$

Now assume the pairs of observables $A_1^{(\lambda)}, A_2^{(\lambda)}$ and $B_1^{(\lambda)}, B_2^{(\lambda)}$ associated with qubits S_a, S_b , respectively, are unbiased observables and thus arise as smearings of sharp qubit observables A_1, A_2 and B_1, B_2 . Let $\mathbb{B}$ and $\mathbb{B}'$ be the Bell operators associated with the sharp observables and the smeared observables, respectively. Then it is straightforward to see that $\mathbb{B}' = \lambda^2 \mathbb{B}$. Satisfaction of the CHSH inequalities for the smeared observables is equivalent to

$$-\frac{2}{\lambda^2} I \leq \mathbb{B} \leq \frac{2}{\lambda^2} I.$$

By virtue of the Tsirelson bound, this is satisfied for all $\lambda \leq 1/\sqrt[4]{2}$. In fact, the smearing parameter λ can be chosen equal to this value to ensure that the CHSH inequalities are satisfied for the smeared versions of all choices of sharp qubit observables A_1, A_2, B_1, B_2 . This shows that Bell inequality violations disappear before the smeared observables may have become compatible as $1/\sqrt[4]{2} > 1/\sqrt{2}$.

Remark 20.4 We noted the necessity and sufficiency of incompatibility for Bell inequality violations, and thus for nonlocality in EPR–Bell type experiments involving spatially separated systems. However, we also saw that some Bell inequalities can still be satisfied even when incompatible observables are used. It has been known for some time that Bell nonlocality is a stronger form of nonseparability than a bipartite quantum state can have than the property of steerability, and the latter is strictly stronger than entanglement [19]. Recent work has shown that there is a correspondence between incompatibility and steerability that is tighter than the connection between incompatibility and Bell nonlocality [20–22].

20.5 Exercises

1. Prove Theorem 20.1.
2. Deduce a complete set of CHSH inequalities of the form (20.2) from the corresponding inequalities for pair probabilities given in Theorem 20.1.
3. Complete the calculations required in the proof of Theorem 20.2.
4. Verify that the operators given in Eq. (20.4) are projections.

References

1. Bell, J.S.: On the Einstein Podolsky Rosen paradox. *Physics* **1**(3), 195–200 (1964)
2. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935)
3. Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969)
4. Boole, G.: The laws of thought. *Great Books in Philosophy*. Prometheus Books, Amherst, NY (2003) (orig. 1854)
5. Boole, G.: On the theory of probabilities. *Phil. Trans. R. Soc. Lond.* **152**, 225–252 (1862)
6. Pitowsky, I.: From George Boole to John Bell—the origins of Bell’s inequality. In: *Bell’s Theorem, Quantum Theory and Conceptions of the Universe*, pp. 37–49. Springer (1989)
7. Pitowsky, I.: George Boole’s “conditions of possible experience” and the quantum puzzle. *Br. J. Philos. Sci.* **45**(1), 95–125 (1994)
8. Fine, A.: Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.* **48**, 291–295 (1982)
9. Fine, A.: Joint distributions, quantum correlations, and commuting observables. *J. Math. Phys.* **23**(7), 1306–1310 (1982)
10. Wolf, M.M., Perez-Garcia, D., Fernandez, C.: Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Phys. Rev. Lett.* **103**, 230402 (2009)
11. de Muynck, W.M., Abu-Zeid, O.: On an alternative interpretation of the Bell inequalities. *Phys. Lett. A* **100**(9), 485–489 (1984)
12. Busch, P.: Elements of unsharp reality in the EPR experiment. In: *Symposium on the Foundations of Modern Physics* (Joensuu, 1985), pp. 343–357. World Sci. Publishing, Singapore (1985)
13. Kar, G., Roy, S.: Objectification problem, CHSH inequalities for a system of two spin-1/2 particles. *Pramana* **46**(1), 9–15 (1996)
14. Kar, G., Roy, S.: Unsharp observables and objectification problem in quantum theory. *Riv. Nuovo Cimento Soc. Ital. Fis.* (4), **22**(3):49 (1999)
15. Busch, P.: EPR-Bell tests with unsharp observables and relativistic quantum measurement. In: Placek, T., Butterfield, J. (eds.) *Non-locality and Modality*, pp. 175–193. Springer, New York (2002)
16. Pitowsky, I.: *Quantum Probability-Quantum Logic*. Lecture Notes in Physics, vol. 321. Springer, Berlin (1989)
17. Landau, L.J.: Experimental tests of general quantum theories. *Lett. Math. Phys.* **14**(1), 33–40 (1987)
18. Banik, M., Gazi, M.R., Ghosh, S., Kar, G.: Degree of complementarity determines the nonlocality in quantum mechanics. *Phys. Rev. A* **87**(5), 052125 (2013)
19. Wiseman, H.M., Jones, S.J., Doherty, A.C.: Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox. *Phys. Rev. Lett.* **98**, 140402 (2007)

20. Quintino, M.T., Vértesi, T., Brunner, N.: Joint measurability, Einstein-Podolsky-Rosen steering, and Bell nonlocality. *Phys. Rev. Lett.* **113**(16), 160402 (2014)
21. Uola, R., Moroder, T., Gühne, O.: Joint measurability of generalized measurements implies classicality. *Phys. Rev. Lett.* **113**, 160403 (2014)
22. Uola, R., Budroni, C., Gühne, O., Pellonpää, J.-P.: One-to-one mapping between steering and joint measurability problems. *Phys. Rev. Lett.* **115**, 230402 (2015)

Chapter 21

Measurement Limitations Due to Conservation Laws

The realisability of measurements of an observable depends on the availability of appropriate measurement couplings between object system and apparatus or probe system. It is a well-known fact that the fundamental interactions between physical systems are subject to symmetries. Formally, the Hamiltonian operator of a quantum system commutes with the generators of the relevant symmetry transformations, which are the elements of the Galilei group in the case of nonrelativistic quantum mechanics and the Poincaré group in relativistic quantum theory. This property of the symmetry generators, which thus are conserved quantities, forms a constraint on the forms of interaction Hamiltonians available in nature [1, 2].

Limitations of the measurability of quantum observables in the presence of additive conserved quantities were first exhibited by Wigner in his original study of 1952 [3] and further investigated in two subsequent papers of Araki and Yanase [4, 5].

Here we present an extended and strengthened version [6] of the so-called Wigner–Araki–Yanase (WAY) theorem, highlighting its crucial assumptions. We also describe the evolution of this theorem from a strict no-go verdict for certain, highly idealised, precise measurements into a quantitative constraint on the accuracy and approximate repeatability of imprecise measurements. Finally we show that a WAY-type limitation also applies to the measurement of continuous, unbounded observables, such as the measurement of position in the presence of momentum conservation [7].

21.1 Measurement of Spin Versus Angular Momentum Conservation

We begin by briefly recalling Wigner’s original presentation of the conflict between angular momentum conservation and repeatable measurements of a spin component. We assume that a probe system is used to measure the spin- x component s_x of a

particle, and that (say) the z -component $s_z \otimes I + I \otimes J_z$, in short, $s_z + J_z$, of the total angular momentum of object and probe is conserved. Here J_z is the probe's angular momentum operator acting on the probe Hilbert space $\mathcal{K}$ (of suitable finite or infinite dimension). Let $\phi \in \mathcal{K}$ denote the initial vector state of the probe, and $\phi_{\pm} \in \mathcal{K}$ two orthogonal pointer states (which can be considered to be elements of an orthonormal basis of $\mathcal{K}$), defining a selfadjoint pointer observable $Z = P[\phi_+] - (I - P[\phi_+])$. In this chapter we choose units where $\hbar = 1$.

Choosing a unitary coupling $U : \mathbb{C}^2 \otimes \mathcal{K} \rightarrow \mathbb{C}^2 \otimes \mathcal{K}$ with

$$U(\varphi_+ \otimes \phi) = \varphi_+ \otimes \phi_+, \quad U(\varphi_- \otimes \phi) = \varphi_- \otimes \phi_-, \quad (21.1)$$

the measurement scheme $(\mathcal{K}, \phi, Z, U)$ defines a repeatable measurement of s_x .

The evolution of the eigenstates $\psi_{\pm} = (\varphi_+ \pm \varphi_-)/\sqrt{2}$ of s_z is then

$$\begin{aligned} \psi_+ \otimes \phi &\longrightarrow \frac{1}{2} [\psi_+ \otimes (\phi_+ + \phi_-) + \psi_- \otimes (\phi_+ - \phi_-)], \\ \psi_- \otimes \phi &\longrightarrow \frac{1}{2} [\psi_+ \otimes (\phi_+ - \phi_-) + \psi_- \otimes (\phi_+ + \phi_-)]. \end{aligned} \quad (21.2)$$

This violates angular momentum conservation, since the expectations of the observable $s_z + J_z$ agree on the right hand sides of (21.2) but differ by one unit on the left hand sides. Since spin component measurements are practically possible, it should be possible to model an approximate realisation of the measurement; following Wigner [3], modifying the pointer observable and the unitary coupling to be of the form

$$Z' = P[\phi_+] - P[\phi_-], \quad (21.3)$$

$$\begin{aligned} U'(\varphi_+ \otimes \phi) &= \alpha \varphi_+ \otimes \phi_+ + \sqrt{1 - \alpha^2} \varphi_- \otimes \eta, \\ U'(\varphi_- \otimes \phi) &= \alpha \varphi_- \otimes \phi_- - \sqrt{1 - \alpha^2} \varphi_+ \otimes \eta, \end{aligned}$$

with $\alpha \in (0, 1)$. There are now three relevant mutually orthogonal pointer states, representing a three-outcome measurement, the third, $\eta \in \mathcal{K}$, corresponding to an uncertain outcome where the apparatus cannot identify a definite spin. As shown by Wigner [3], one can choose ϕ and η such that the conservation of $s_z + J_z$ is satisfied.

The measured observable has now three outcomes represented by the effects $E_{\pm} = \alpha^2 P[\varphi_{\pm}]$ and $E_0 = (1 - \alpha^2)I$. This observable, which still does not commute with the conserved quantity, is an unsharp version of the spin observable one intends to measure, and the repeatability has now been compromised since even if an outcome “+” or “-” is shown, there is a nonzero probability that an immediate repetition will yield the “uncertain” outcome.

As shown in [3], if the probe allows for a large number (n) of eigenvalues of J_z , the probability $1 - \alpha^2$ of the “uncertain” outcome can be made arbitrarily small for sufficiently large n , whilst maintaining the conservation of the angular momentum component $s_z + J_z$, and accurate and repeatable measurements are recovered to a very good approximation.

We note that the large size of the probe is used here only as a sufficient condition to achieve good measurement accuracy. As we will see below, a quantitative reformulation of the Wigner–Araki–Yanase theorem stipulates the large size of the probe as the limitation that must be accepted if an approximately repeatable and accurate measurement of an observable not commuting with an additive conserved quantity is to be achieved.

21.2 The Yanase Condition

To the extent that the above example of a measurement limitation due to a conservation law points to a universal restriction of measurability, one must expect this limitation also to hold for the measurement of the pointer observable itself. The requirement of stable and unambiguous recordings of measurement outcomes now leads to the condition that the measurement of a pointer observable should be repeatable. Thus, the pointer observable should be chosen (discrete and) so as to commute with any additive conserved quantity since otherwise the repeatability of its measurement cannot be achieved. This condition—the commutativity of the pointer observable with the probe part of the conserved quantity governing a measurement interaction—is known as the *Yanase condition* [5].

21.3 The Wigner–Araki–Yanase Theorem

The first general statement of the Wigner–Araki–Yanase theorem is due to Araki and Yanase [4] who showed that an observable that permits a repeatable sharp measurement in the presence of a conservation law must commute with the object part of that conserved quantity. Here we give a somewhat stronger version of the proof which is obtained by a simple rephrasing of the original proof.

Theorem 21.1 (WAY Theorem) *Let $\mathcal{M} = (\mathcal{K}, U, \phi, Z, f)$ be a measurement of a sharp observable $\mathbf{M}$ associated with a discrete selfadjoint operator M in $\mathcal{H}$, and let L_1 and L_2 be bounded selfadjoint operators on $\mathcal{H}$ and $\mathcal{K}$, respectively, such that $[U, L_1 \otimes I + I \otimes L_2] = 0$. Assume that $\mathcal{M}$ is repeatable or satisfies the Yanase condition $[L_2, Z] = 0$. Then $[L_1, M] = 0$.*

Proof We choose orthonormal bases $\{\varphi_{ij}\}$ and $\{\phi_{ik}\}$ of eigenstates of M and Z , respectively (with j, k as degeneracy indices). The most general unitary coupling U that constitutes a measurement of $\mathbf{M}$ then takes the form

$$U(\varphi_{ij} \otimes \phi) = \sum_k \psi_{ijk} \otimes \phi_{ik}, \quad (21.4)$$

where $\{\psi_{ijk}\}$ in $\mathcal{H}$ is an arbitrary set of vectors such that

$$\sum_k \langle \psi_{ijk} | \psi_{ij'k} \rangle = \delta_{jj'}$$

for all i .¹ Implementing the conservation law, that is, $L = U^*LU$, we may write the matrix elements of $L = L_1 \otimes I + I \otimes L_2$ in the following way:

$$\langle \varphi_{i'j'} \otimes \phi | L \varphi_{ij} \otimes \phi \rangle = \sum_{k,k'} \langle \psi_{i'j'k'} \otimes \phi_{i'k'} | L \psi_{ijk} \otimes \phi_{ik} \rangle. \quad (21.5)$$

The additivity of L entails that (21.5) can be written

$$\begin{aligned} & \langle \varphi_{i'j'} | L_1 \varphi_{ij} \rangle + \langle \varphi_{i'j'} | \varphi_{ij} \rangle \langle \phi | L_2 \phi \rangle \\ &= \sum_{k,k'} [\langle \psi_{i'j'k'} | L_1 \psi_{ijk} \rangle \langle \phi_{i'k'} | \phi_{ik} \rangle + \langle \psi_{i'j'k'} | \psi_{ijk} \rangle \langle \phi_{i'k'} | L_2 \phi_{ik} \rangle]. \end{aligned} \quad (21.6)$$

By the orthogonality of pointer eigenstates, examination of each of the remaining terms in the sum in the above expression tells us that these vanish if one of the following conditions holds:

- (a) $\langle \psi_{i'j'k'} | \psi_{ijk} \rangle = 0$ for $i \neq i'$;
- (b) $\langle \phi_{i'k'} | L_2 \phi_{ik} \rangle = 0$ for $i \neq i'$.

If $\mathcal{M}$ is repeatable, then it follows (see, for instance, Theorem 10.3) that condition (a) is satisfied. On the other hand, if Z commutes with L_2 then condition (b) is satisfied. If either of these is satisfied, then the right hand side of (21.6) is zero, and thus the left hand side must vanish also. Clearly the second term on the left hand side vanishes due to the orthogonality of the eigenstates of M , and the first vanishes if and only if L_1 leaves M -eigenspaces invariant, i.e. if and only if $[L_1, M] = 0$. $\square$

We interpret the theorem as follows: if $\mathcal{M}$ is a measurement of M and $[L_1, M] \neq 0$, then the conservation of $L_1 \otimes I + I \otimes L_2$ entails that $\mathcal{M}$ must violate both repeatability and the Yanase condition.

Remark 21.1 As demonstrated in footnote 4 of [4], the case of L_2 being unbounded can be incorporated into the proof in a natural way. This is achieved by using the unitary operators $V(t) = \exp(itL)$ and $V_i(t) = \exp(itL_i)$ (with $i = 1, 2, t \in \mathbb{R}$) and noting that $V(t) = V_1(t) \otimes V_2(t)$. Then one can follow the previous line of argument, replacing the original operators with their exponentiated forms, and exploiting the boundedness of L_1 . It is not known whether the result can be extended to the case where L_1 is unbounded as well. $\triangleleft$

Remark 21.2 A result of the form of the above theorem—i.e. using the weakened form (a) of repeatability or the Yanase condition (b) to deduce the commutativity

¹Recall that if M is of rank ∞ then (21.4) does not necessarily extend to a unitary operator.

of the observable to be measured with the conserved quantity—has been proved by Beltrametti et al. in 1990 [8] for a minimal unitary measurement $\mathcal{M}$, where the discrete pointer observable is, by definition, nondegenerate.

Further extensions and refinements of the Wigner–Araki–Yanase theorem have been obtained by several research groups over the years; a survey of these developments can be found in [6]. $\triangleleft$

It should be noted that the Wigner–Araki–Yanase theorem as formulated above does not apply to the measurement of observables with continuous value spaces, such as position, for which repeatable measurements do not exist (Theorem 10.4). This raises the question whether conservation laws constrain measurements in such cases. It turns out that the Wigner–Araki–Yanase theorem can be rephrased in the form of a quantitative limitation where approximate and approximately repeatable measurements are possible only at the expense of a “large size” of the probe system, quantified by the magnitude of the variance of the probe part of the conserved quantity.

21.4 A Quantitative Version of the WAY Theorem

A Wigner–Araki–Yanase-type limitation in the form of trade-off inequalities follows as an application of the Cauchy–Schwarz inequality, as shown by Ozawa [9].

Given a measurement $\mathcal{M}$ that is to serve as an approximate determination of an observable M , we recall from Eq. (13.36) that the measurement noise $\epsilon_{mn}(Z, M; \varphi)$ is defined as the square root of the expectation of the squared difference operator $(M_{\text{out}} - M_{\text{in}})^2$ in the state $\varphi \otimes \phi$. Denoting this difference operator by N , write $\epsilon(\varphi) = \epsilon_{mn}(Z, M; \varphi)$, and assume that N is selfadjoint. Clearly $\epsilon(\varphi)^2 \geq (\Delta N)^2$, where ΔN is the standard deviation of $E_{\varphi \otimes \phi}^N$. The global noise is defined as the supremum over all input (vector) states φ of $\epsilon(\varphi)^2$, i.e. $\epsilon^2 = \sup_{\varphi} \epsilon(\varphi)^2$ (cf. Eq. (13.38)). This quantity should be finite for any measurement that would qualify as an approximate determination of M . Then the Cauchy–Schwarz inequality entails

$$\epsilon^2 \geq \epsilon(\varphi)^2 = \frac{(\Delta N)^2 (\Delta L)^2}{(\Delta L)^2} \geq \frac{1}{4} \frac{|\langle [N, L] \rangle|^2}{(\Delta L)^2}, \quad (21.7)$$

where the variance $(\Delta L)^2 = (\Delta_{\varphi \otimes \phi} L)^2 = (\Delta_{\varphi} L_1)^2 + (\Delta_{\phi} L_2)^2$ and $\langle [N, L] \rangle = \langle [N, L] \rangle_{\varphi \otimes \phi} = \langle \varphi \otimes \phi | [N, L] \varphi \otimes \phi \rangle$. We recall that the measurement is accurate if and only if $\epsilon = 0$.

If the Yanase condition, $[Z, L_2] = 0$, is satisfied and the interaction obeys the conservation law, then all that remains in the numerator is $|\langle [M, L_1] \rangle_{\varphi}|^2$:

$$\epsilon^2 \geq \sup_{\varphi} \frac{1}{4} \frac{|\langle [M, L_1] \rangle_{\varphi}|^2}{(\Delta_{\varphi} L_1)^2 + (\Delta_{\phi} L_2)^2}, \quad (21.8)$$

where the supremum is extended over all vector states in the domain of the commutator $ML_1 - L_1M$. If this commutator is zero, then there is no lower bound on the measurement noise, in accordance with the Wigner–Araki–Yanase theorem.

In general, however, the lower bound of inequality (21.8) depends on the system state φ . In the case that $|\langle [M, L_1] \rangle_\varphi|^2$ is non-zero but finite, clearly if $(\Delta L)^2$ becomes large the lower bound on the global noise decreases. Furthermore, since the initial system state φ is arbitrary, only by fixing ϕ such that $(\Delta_\phi L_2)^2$ is nonzero can one prevent the denominator from becoming zero. Moreover, we see that decreasing ϵ as an error bound requires making $(\Delta_\phi L_2)^2$ large, thus establishing the necessity of a large apparatus variance for good measurements.

It is worthwhile investigating the case of a measurement scheme $\mathcal{M}$ that satisfies neither of the commutativity conditions $[Z, L_2] = 0$ and $[M, L_1] = 0$ but is such that the bound on the right hand side of (21.7) vanishes in all states φ ; thus, $[M_{\text{out}}, L_1 + L_2] = [M, L_1] = U^*[Z, L_2]U$, by the conservation law. This is clearly satisfied if $\mathcal{M}$ happens to be accurate, $\epsilon = 0$. Such a measurement scheme allows for perfectly accurate transfer of information from system to apparatus, and demonstrates the necessary violation of the Yanase condition for this to be achieved.

21.5 Position Measurements Obeying Momentum Conservation

It is possible to adapt the inequality (21.7) to obtain a general argument pointing to a Wigner–Araki–Yanase-type of limitation for position measurements obeying momentum conservation in the case where the Yanase condition is satisfied. In that case,

$$\begin{aligned} \epsilon^2 &\geq \sup_{\varphi} \frac{1}{4} \frac{|\langle \varphi | [Q, P] \varphi \rangle|^2}{(\Delta_{\varphi} P)^2 + (\Delta_{\phi} P_{\mathcal{A}})^2} \\ &= \frac{1}{4} \frac{1}{\inf_{\varphi} (\Delta_{\varphi} P)^2 + (\Delta_{\phi} P_{\mathcal{A}})^2} = \frac{1}{4(\Delta_{\phi} P_{\mathcal{A}})^2}, \end{aligned}$$

where the supremum and infimum are taken over the unit vectors in the domain of the commutator $[Q, P] = iI$, which contains the Schwartz space of rapidly decreasing functions for which $\inf_{\varphi} (\Delta_{\varphi} P)^2 = 0$. This bound allows for an increase in accuracy only when $(\Delta_{\phi} P_{\mathcal{A}})^2$ is large, establishing the necessity of a large variance of the conserved quantity for good measurements.

Notice that the non-zero lower bound to the accuracy arises after explicit implementation of the Yanase condition, $[Z, P_{\mathcal{A}}] = 0$. If we relinquish this condition, there is nothing that would prevent $[Z_{\text{out}} - Q, P + P_{\mathcal{A}}]$ from vanishing. Indeed, this would be the case in any model where one could choose the pointer observable as the apparatus position, $Q_{\mathcal{A}}$.

In the position-momentum case, the role of the Yanase condition must be considered very carefully. Previously—in the case where the WAY theorem certainly applied to the probe system—we argued for the Yanase condition by applying the theorem to the measurement of the pointer, of which we demanded accurate and repeatable measurements. However, since no such theorem has been proved in the continuous/unbounded case, it is as yet unknown whether this condition must be met for a measurement to comply with the conservation law. The above model-independent relation points in the direction of a Wigner–Araki–Yanase type limitation if the Yanase condition is satisfied. If it is not, there is no such obstruction.

We present two models of approximate position measurements obeying momentum conservation that highlight the significance of the Yanase condition.

The first model, introduced by Ozawa in 1991 [10], has the Yanase condition built in. This model involves a probe system consisting of three reference particles, with the Hilbert spaces $\mathcal{K}_i$, $i = 1, 2, 3$. The coupling operator is given as (for ease of notation, we omit the tensor product signs together with the relevant identity operators)

$$\begin{aligned} U &= e^{-i \frac{\lambda}{\sqrt{2}} (Q - Q_1)(Q_2 - Q_3)} = e^{-i \frac{\lambda}{\sqrt{2}} Q(Q_2 - Q_3)} e^{i \frac{\lambda}{\sqrt{2}} Q_1(Q_2 - Q_3)} \\ &= \iint Q(dx) Q_1(dy) e^{-i \frac{\lambda}{\sqrt{2}} (x-y)(Q_2 - Q_3)}. \end{aligned}$$

We take $\mathcal{H}$ and $\mathcal{K}_1$ to be the position representation spaces $L^2(\mathbb{R})$ whereas for $\mathcal{K}_2$ and $\mathcal{K}_3$ we choose the momentum representation spaces $L^2(\mathbb{R})$ so that $\mathcal{K}_2 \otimes \mathcal{K}_3 \cong L^2(\mathbb{R}^2)$ with $\mathbb{R}^2$ is parametrised by (p_2, p_3) . By using the orthogonal transformation $(p_2, p_3) \mapsto (u, v) = 2^{-1/2}(p_2 - p_3, p_2 + p_3)$ we see that $\mathcal{K}_2 \otimes \mathcal{K}_3$ is unitarily equivalent to $\tilde{\mathcal{K}}_2 \otimes \tilde{\mathcal{K}}_3 \cong L^2(\mathbb{R}^2)$, with coordinates (u, v) , where e.g. the operator $P_2 - P_3$ is the multiplication by $\sqrt{2}u$ and behaves covariantly under the action generated by $Q_2 - Q_3$. We use $P_3 - P_2$ as the pointer observable and denote its spectral measure by Z .

Assuming that the initial probe state is the vector state $\phi_1 \otimes \phi_2 \otimes \phi_3 \in \mathcal{K}_1 \otimes \tilde{\mathcal{K}}_2 \otimes \tilde{\mathcal{K}}_3$ the measured observable $\mathbf{E}$ is determined to be

$$\begin{aligned} \mathbf{E}_\varphi(X) &= \langle U(\varphi \otimes \phi_1 \otimes \phi_2 \otimes \phi_3) | I_{\mathcal{H}} \otimes I_{\mathcal{K}_1} \otimes Z(\lambda X) \otimes I_{\tilde{\mathcal{K}}_3} U(\varphi \otimes \phi_1 \otimes \phi_2 \otimes \phi_3) \rangle \\ &= \iint Q_\varphi(dx) Q_{1, \phi_1}(dy) Z_{\phi_2}(\lambda(X - x + y)) \\ &= \int Q_\varphi(dx) (Q_{1, \Pi \phi_1} * Z_{\phi_2})(\lambda(X - x)) \\ &= (Q_\varphi * (Q_{1, \Pi \phi_1} * Z_{\phi_2}))(\lambda X). \end{aligned} \tag{21.9}$$

In this model the Yanase condition holds, $[Z, P_2 + P_3] = 0$, and the square of the measurement accuracy is bounded from below by the sum

$$\text{Var}(Q_{1, \Pi \phi_1}) + \frac{1}{\lambda^2} \text{Var}(Z_{\phi_2}),$$

which is always strictly positive, showing that the measurement serves only as an approximative position measurement. In order to have this bound small, the coupling parameter λ has to be large and the variance of the first probe particle's position must be made small.

The second model is a modification of the standard model for position measurements [11]. The unitary coupling between object particle and probe particle is given as (again, we omit the tensor product notations and the relevant identity operators)

$$U = e^{-i\frac{\lambda}{2}((Q-Q_1)P_1 + P_1(Q-Q_1))}. \quad (21.10)$$

(We assume that $\lambda > 1$.) This evolution operator commutes with the total momentum $P + P_1$.

In the present case it seems most convenient to evaluate the action of U by solving the Schrödinger equation for the Hamiltonian $H = \frac{\lambda}{2}((Q - Q_1)P_1 + P_1(Q - Q_1))$ and the initial condition $\Psi_0 = \varphi\phi$. It is straightforward to verify that the solution takes the form

$$\Psi_t(x, x_1) = e^{\lambda t/2} \varphi(x) \phi((1 - e^{\lambda t})x + e^{\lambda t} x_1).$$

(The form of the trial solution $\Psi_t(x, x_1) = f(t)\varphi(x)\phi(a(t)x + b(t)x_1)$ can be guessed by taking a clue from the solution of the Heisenberg equations for Q and Q_1 .) After a time $t = 1/2$ the coupling is switched off and the pointer observable is measured on the state $\Psi_{1/2} = \Psi_f = U\Psi_0$.

If the pointer observable and pointer function are the position Q_1 and $h(x) = (1 - e^{-\lambda})^{-1}x$, respectively, the measured observable is a smeared position $\mu^{(\lambda)} * Q$, where $\mu^{(\lambda)}$ has the density $e^{(\lambda)}$,

$$e^{(\lambda)}(x) = (e^\lambda - 1) |\phi(-x(e^\lambda - 1))|^2. \quad (21.11)$$

Here we see that one may keep the probe preparation fixed and still make the accuracy distribution as sharply peaked as one wishes by tuning the scaling parameter λ appropriately. It is noteworthy that compared to the standard model, the inaccuracy scales with $e^{-\lambda}$ rather than λ^{-1} . This shows that if the Yanase condition is dropped and the pointer chosen as the position of the probe, arbitrarily accurate measurements of position can be realised within the constraints of momentum conservation, without having to pay the price of large variance of the probe momentum.

This conclusion remains also true with respect to the repeatability properties of the measurements as we show next.

We shall confine the probe's initial state to a bounded subset of the real line. This is not an overly stringent requirement from a physical perspective. In the Ozawa model this simply amounts to the initial state vectors ϕ_1 and ϕ_2 having bounded supports; in the second model it means that the probe state ϕ is concentrated in a finite interval. Thus we can think of the inaccuracy measure as being concentrated on an interval $[-d, d]$ in either model.

One way of quantifying the degree of approximate repeatability [12, 13] in the case of a position measurement is as follows: A measurement is said to be approximately, or δ -repeatable (for $\delta > 0$) if given an outcome in a set X , the outcome of an immediate subsequent control measurement will be found, with probability 1, in a suitably enlarged set X_δ (where X_δ is the set of points not more than a distance $\delta > 0$ away from X). This can be written symbolically as a conditional probability of finding the particle's position $x \in X_\delta$ given that the pointer was found to have a value $u \in h^{-1}(X)$:

$$P_{\psi_f}(x \in X_\delta | u \in h^{-1}(X)) = 1 \quad (21.12)$$

for all sets X . Considering the control measurement to be accurate, for this to be satisfied in the Ozawa model we must have $E_\varphi(X) = 0$ whenever φ is supported outside of X_δ , and this follows if $\delta \geq d$. If the initial apparatus states ϕ_1 and ϕ_2 are concentrated on intervals $[-\ell, \ell]$ and $[-m, m]$ respectively, we have that $d = \ell + m/\lambda$. Therefore even as the coupling strength λ becomes large, δ is bounded below by the width of the reference system state ϕ_1 , and in order to recover good repeatability properties (i.e. a small δ), ϕ_1 must carry a large spread of momentum.

In the alternative model we see similar behaviour, with a fundamental difference; we again have that $\delta \geq d$ enables approximate repeatability in the sense of a condition analogous to (21.12). However, in contrast to the Ozawa model, simply letting λ be large allows for arbitrarily good repeatability; if ϕ is concentrated on $[-n, n]$, then $d = n/(e^\lambda - 1)$.

Thus under violation of the Yanase condition, arbitrarily accurate and repeatable information transfer from the system to a quantum probe is feasible without any size constraint (n can be arbitrarily large, allowing the spread of the probe momentum to be small).

The last example, showing no measurement obstruction due to the conservation law, contrasts, perhaps somewhat surprisingly, with the WAY theorem for accurate measurements: Within the realm of that theorem, it is not sufficient to violate the Yanase condition in order to lift the obstruction against perfect accuracy and repeatability.

It is not known whether, under violation of the Yanase condition, there exist measurements of position that are fully accurate, and repeatable to a good approximation. It is also an open problem whether, again with giving up the Yanase condition, approximate spin measurements obeying angular momentum conservation are possible with good repeatability properties, without any constraint on the size of the apparatus.

21.6 A Measurement-Theoretic Interpretation of Superselection Rules

In the same year (1952) that Wigner investigated the measurement limitation due to conservation laws he also coauthored (with Wick and Wightman) the so-called “W³” paper [14] that introduced *superselection rules*. Here we explore the logical

relation between the two kinds of limitations. While the Wigner–Araki–Yanase theorem applies in the presence of additive conserved quantities and precludes repeatable measurements as well as barring the satisfaction of the Yanase condition, the statement of a superselection rule is unconditional. That is, no operator measure that does not commute with a given superselection charge N can represent an observable. Hence any observable $\mathbf{E}$ must commute with N , that is, for any $X \in \mathcal{A}$,

$$e^{i\theta N} \mathbf{E}(X) e^{-i\theta N} = \mathbf{E}(X).$$

This constraint is often also described by saying that phase relations between charge sectors are *unobservable* as there is no “phase-sensitive” observable.

That this statement is stronger than that of the Wigner–Araki–Yanase theorem can be seen by casting it in measurement theoretic terms as a “strong” Wigner–Araki–Yanase theorem (where stronger assumptions yield a stronger conclusion).²

Proposition 21.1 *Suppose $N_{\text{tot}} = N \otimes I + I \otimes N_{\mathcal{A}}$ is strongly conserved, in the sense that for any measurement of any observables $\mathbf{E}$, the coupling U preserves $N \otimes I$ and $I \otimes N_{\mathcal{A}}$ separately so that, in particular, $U^* N \otimes I U = N \otimes I$. Then*

$$e^{i\theta N} \mathbf{E}(X) e^{-i\theta N} = \mathbf{E}(X) \quad (21.13)$$

for all value sets X and all $\theta \in [0, 2\pi)$.

Proof It is sufficient to consider measurement schemes where the initial pointer state is a vector state ϕ and the pointer function is the identity map. For U to serve as a measurement of $\mathbf{E}$ with some pointer observable $\mathbf{Z}$, one has

$$\langle U\varphi \otimes \phi | I \otimes \mathbf{Z}(X) U\varphi \otimes \phi \rangle = \langle \varphi | \mathbf{E}(X) \varphi \rangle$$

for all object initial states φ and all X . Replacing φ with $e^{-iN\theta} \varphi$ does not change the left hand side, therefore equating the right hand sides immediately gives (21.13). $\square$

There is a kind of converse to Proposition 21.1.

Proposition 21.2 *Assume that a superselection rule applies for system and probe, $\mathcal{S} + \mathcal{A}$, with the charge $N_{\text{tot}} = N \otimes I + I \otimes N_{\mathcal{A}}$, and suppose that the measurement couplings U are made up of observables of the system and probe, so that each U is of the form $U = \sum_{jk} c_{jk} A_j \otimes B_k$ (or a limit of such terms). Then $U^* e^{iN\theta} \otimes I U = e^{iN\theta} \otimes I$ for all $\theta \in [0, 2\pi)$.*

Proof Since N and $N_{\mathcal{A}}$ are superselection charges for $\mathcal{S}$ and $\mathcal{A}$, respectively, the operators A_j and B_k commute with these respective charges. Then it is clear that U commutes with N and hence the strong conservation law follows. $\square$

²A similar statement was proved by S. Tanimura [15] on the basis of stronger assumptions.

21.7 Exercises

1. Show that the coupling defined in Eq. (21.3) extends to a unitary map that commutes with the total angular momentum component $s_z + J_z$, and determine the measured observable.
2. In the first position measurement model of Sect. 21.5, the coupling U is based on a Hamiltonian

$$H = \lambda(Q - Q_A)(Q_B - Q_C).$$

Instead of working out the action of U directly as done in the text, solve the associated Schrödinger equation with the initial condition $\Psi_0(x, y, u, v) = \varphi(x)\phi_1(y)\phi_2(u)\phi_3(v)$ (with the probe Hilbert spaces as given in the text). Thus show that

$$\Psi_t(x, y, u, v) = \varphi(x)\phi_1(y)\phi_2(u + 2\lambda t(x - y))\phi_3(v).$$

Putting $t = 1/\sqrt{2}$ to obtain $\Psi_{1/\sqrt{2}} = \Psi_f$, verify the measured observable, Eq. (21.9).

3. Verify the solution of the Schrödinger equation given in the context of the second position measurement model of Sect. 21.5.

References

1. Mackey, G.W.: Axiomatics of particle interactions. In: Proceedings of the International Quantum Structures Association, Part I (Castiglione, 1992), vol. 32, pp. 1643–1659 (1993)
2. Mackey, G.W.: Weyl's program and modern physics. In: Differential Geometrical Methods in Theoretical Physics (Como, 1987). NATO Advanced Science Institutes Series C Mathematical and Physical Sciences, vol. 250, pp. 11–36. Kluwer Acad. Publ., Dordrecht (1988)
3. Wigner, E.P.: Die Messung quantenmechanischer Operatoren. Z. Physik **133**, 101–108 (1952). English Translation at [arXiv:1012.4372](https://arxiv.org/abs/1012.4372)
4. Araki, H., Yanase, M.M.: Measurement of quantum mechanical operators. Phys. Rev. **2**(120), 622–626 (1960)
5. Yanase, M.M.: Optimal measuring apparatus. Phys. Rev. **123**, 666–668 (1961)
6. Loveridge, L., Busch, P.: 'Measurement of quantum mechanical operators' revisited. Eur. Phys. J. D **62**(2), 297–307 (2011)
7. Busch, P., Loveridge, L.: Position measurements obeying momentum conservation. Phys. Rev. Lett. **106**, 110406 (2011)
8. Beltrametti, E., Cassinelli, G., Lahti, P.: Unitary measurements of discrete quantities in quantum mechanics. J. Math. Phys. **31**(1), 91–98 (1990)
9. Ozawa, M.: Conservation laws, uncertainty relations, and quantum limits of measurements. Phys. Rev. Lett. **88**, 050402 (2002)
10. Ozawa, M.: Does a conservation law limit position measurements? Phys. Rev. Lett. **67**, 1956–1959 (1991)
11. Busch, P., Grabowski, M., Lahti, P.: Operational Quantum Physics. Lecture Notes in Physics Monographs, vol. 31. Springer, 2nd corrected edition (1997)
12. Davies, E.B.: On the repeated measurements of continuous observables in quantum mechanics. J. Funct. Anal. **6**, 318–346 (1970)

13. Busch, P., Lahti, P.J.: Some remarks on unsharp quantum measurements, quantum nondemolition, and all that. *Ann. Physik* **47**(5), 369–382 (1990)
14. Wick, G.C., Wightman, A.S., Wigner, E.P.: The intrinsic parity of elementary particles. *Phys. Rev.* **88**, 101–105 (1952)
15. Tanimura, S.: Superselection rules from measurement theory (2011). [arXiv:1112.5701](https://arxiv.org/abs/1112.5701)

Chapter 22

Measurement Problem

As shown in Chap. 10, a measurement of an observable $E \in \mathcal{O}(\Omega, \mathcal{A}, \mathcal{H})$ is modelled by a measurement scheme $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$, which reproduces the measurement outcome probabilities of E in any initial state ϱ of the object system as the corresponding distribution of the pointer observable $\mathbf{Z}$ in the final probe state σ^f , that is, $E_\varrho = \mathbf{Z}_{\sigma^f}$. We now make precise the so-called quantum measurement problem—the question whether an account can be provided, in terms of quantum mechanics, of the occurrence of a definite value of the pointer observable at the end of a measurement, and what inference, if any, may be possible on the value of the measured object observable E after the measurement.

22.1 Preliminaries

Every measurement has a finite resolution. The simplest way of accommodating this fact is obtained by describing the recording of outcomes as a measurement of a discretised pointer observable, using a reading scale as defined in Remark 10.2. Thus, let $\mathcal{R}$ be a partition of Ω into disjoint measurable subsets $X_i, i \in \mathbb{I} \subset \mathbb{N}$; this induces the corresponding discrete observables, $\mathbf{Z}^\mathcal{R} : i \mapsto \mathbf{Z}_i = \mathbf{Z}(X_i)$ and $E^\mathcal{R} : i \mapsto E_i = E(X_i)$. Introducing the pointer function $r : \Omega \rightarrow \mathbb{I}$, with $r^{-1}(\{i\}) = X_i$, the “discretised” measurement scheme defining $E^\mathcal{R}$ is then $\mathcal{M}^\mathcal{R} = (\mathcal{K}, \mathbf{Z}, \sigma, U, r)$. In any real situation, $\mathbb{I}$ is finite, but we allow here the case $\mathbb{I} = \mathbb{N}$.

Given the reading scale $\mathcal{R}$, the normalised final component states, as defined in Sect. 10.2, are (for X_i such that $E_\varrho(X_i) \neq 0$)

$$\begin{aligned}
\varrho_i^f &= \varrho^f(X_i) = \mathbb{E}_\varrho(X_i)^{-1} \operatorname{tr}_{\mathcal{K}}[I \otimes Z(X_i)^{1/2} U(\varrho \otimes \sigma) U^* I \otimes Z(X_i)^{1/2}] \\
&= \mathbb{E}_\varrho(X_i)^{-1} \mathfrak{I}(X_i)(\varrho) \\
\sigma_i^f &= \sigma^f(X_i) = \mathbb{E}_\varrho(X_i)^{-1} \operatorname{tr}_{\mathcal{H}}[I \otimes Z(X_i)^{1/2} U(\varrho \otimes \sigma) U^* I \otimes Z(X_i)^{1/2}] \\
&= \mathbb{E}_\varrho(X_i)^{-1} Z_i^{1/2} \sigma^f Z_i^{1/2}.
\end{aligned}$$

For those X_i for which $\mathbb{E}_\varrho(X_i) = 0$ we will put $\varrho_i^f = 0$ and $\sigma_i^f = 0$. The unconditional final object state is a mixture of the ϱ_i^f with the weights $\mathbb{E}_\varrho(X_i)$:

$$\varrho^f = \mathfrak{I}(\mathbb{R})(\varrho) = \mathfrak{I}(\cup X_i)(\varrho) = \sum_i \mathbb{E}_\varrho(X_i) \varrho_i^f; \quad (22.1)$$

but it is important to note that the corresponding decomposition of $\sigma^f = \operatorname{tr}_{\mathcal{H}}[U(\varrho \otimes \sigma) U^*]$,

$$\sigma^f = \sum_i \mathbb{E}_\varrho(X_i) \sigma_i^f, \quad (22.2)$$

will only be applicable under suitable additional assumptions, some of which will be indicated later. Moreover,

$$\operatorname{tr}[\sigma_i^f Z_i] = \mathbb{E}_\varrho(X_i)^{-1} \operatorname{tr}[\sigma^f Z_i^2] \leq \mathbb{E}_\varrho(X_i)^{-1} \operatorname{tr}[\sigma^f Z_i] = 1,$$

and generally, it will be the case that $\operatorname{tr}[\sigma_i^f Z_i] < 1$. If the pointer observable is projection valued, $Z_i^2 = Z_i$, then $\operatorname{tr}[\sigma_i^f Z_i] = 1$ and the final component states σ_i^f are pointer eigenstates, $Z_i \sigma_i^f = \sigma_i^f$ whenever $\mathbb{E}_\varrho(X_i) \neq 0$. It is worth noting that this does not imply that $\operatorname{tr}[\varrho_i^f E_i] = 1$.

To justify that the pointer observable does assume a definite value (with respect to a reading scale) after the measurement, that is, in the final probe state σ^f , one should be able to conclude that (1) the decomposition (22.2) is valid, and (2) the component states σ_i^f are pointer eigenstates. In addition, one should be able to justify that (3) the decomposition is somehow physically uniquely distinguished, a requirement which seems impossible in the light of the nonunique decomposability of mixed states, Theorem 9.2. If this all turned out to be possible, one could then go on to investigate under what conditions the measured observable can be said to have assumed a definite value after the measurement.

We begin by studying the conditions under which a measurement of an observable satisfies the first two requirements (Sect. 22.2). The measurement problem is formulated in Sect. 22.3 together with a brief outline of some attempts to overcome it.

22.2 Reading of Pointer Values

We say that a measurement $\mathcal{M} = (\mathcal{K}, \mathbf{Z}, \sigma, U)$ of an observable $\mathbf{E}$ satisfies the *pointer value-definiteness condition* with respect to a reading scale $\mathcal{R}$ if

$$\mathrm{tr}[\sigma_i^f Z_i] = 1 \quad (22.3)$$

for all $i \in \mathbb{I}$ and all initial states ϱ of $\mathcal{S}$ for which $\mathbf{E}_\varrho(X_i) \neq 0$. As noted above, this condition holds if $\mathbf{Z}$ is sharp. In addition, we say that $\mathcal{M}$ satisfies the *pointer mixture condition* with respect to $\mathcal{R}$ if the decomposition (22.2) is valid for each initial state ϱ of $\mathcal{S}$. The following result connects these two properties with the mutual orthogonality of the final component states ϱ_i^f of the object system.

Theorem 22.1 *Let $\mathcal{M}$ be a measurement of an observable $\mathbf{E}$ and $\mathcal{R}$ a reading scale. For any initial state ϱ of the object system, condition (i) implies conditions (ii) and (iii), where*

- (i) $\varrho_i^f \cdot \varrho_j^f = 0$ for all $i \neq j$;
- (ii) $\sigma^f = \sum_i \mathbf{E}_\varrho(X_i) \sigma_i^f$;
- (iii) $\mathrm{tr}[\sigma_i^f Z_i] = 1$ for all i .

Moreover, if $\sigma = P[\phi]$, then for any $\varrho = P[\varphi]$ the conditions (ii) and (iii) together imply (i).

Proof (i) $\Rightarrow$ (ii) & (iii): For each i , let F_i be the support projection of ϱ_i^f , that is, the smallest projection P such that $P\varrho_i^f = \varrho_i^f$. Using Exercise 1 one now gets (for $i \neq j$):

$$\begin{aligned} \varrho_i^f \cdot \varrho_j^f = 0 &\Leftrightarrow F_i \varrho_j^f = 0 \\ &\Leftrightarrow \mathrm{tr}[F_i \otimes Z_j^{1/2} U(\varrho \otimes \sigma) U^* F_i \otimes Z_j^{1/2}] = 0 \\ &\Leftrightarrow F_i \otimes Z_j^{1/2} U(\varrho \otimes \sigma) U^* F_i \otimes Z_j^{1/2} = 0 \\ &\Leftrightarrow F_i \otimes Z_j^{1/2} (U(\varrho \otimes \sigma) U^*)^{1/2} = 0 \\ &\Leftrightarrow F_i \otimes Z_j (U(\varrho \otimes \sigma) U^*)^{1/2} = 0. \end{aligned} \quad (22.4)$$

Since $(I - F_i)\varrho_i^f = 0$, a similar line of calculation also gives

$$F_i \otimes Z_i (U(\varrho \otimes \sigma) U^*)^{1/2} = I \otimes Z_i (U(\varrho \otimes \sigma) U^*)^{1/2}. \quad (22.5)$$

Combining (22.4) and (22.5) and using the fact that $\sum Z_i = I$ yields

$$\begin{aligned} F_i \otimes I (U(\varrho \otimes \sigma)U^*)^{1/2} &= F_i \otimes Z_i (U(\varrho \otimes \sigma)U^*)^{1/2} \\ &= I \otimes Z_i (U(\varrho \otimes \sigma)U^*)^{1/2}. \end{aligned} \quad (22.6)$$

From this one obtains $I \otimes Z_i U(\varrho \otimes \sigma)U^* = I \otimes Z_i^2 U(\varrho \otimes \sigma)U^*$, which gives (iii).

Using (iii), one shows similarly that

$$I \otimes Z_i (U(\varrho \otimes \sigma)U^*)^{1/2} = I \otimes Z_i^{1/2} (U(\varrho \otimes \sigma)U^*)^{1/2}. \quad (22.7)$$

Multiplying the expression on the left hand side of (22.6) with its adjoint and summing over i , one obtains using (22.7):

$$\sum F_i \otimes I U(\varrho \otimes \sigma)U^* F_i \otimes I = \sum I \otimes Z_i^{1/2} U(\varrho \otimes \sigma)U^* I \otimes Z_i^{1/2}. \quad (22.8)$$

Taking the partial trace with respect to $\mathcal{H}$ finally yields (ii).

(ii) & (iii) $\Rightarrow$ (i): Let now $\sigma = P[\phi]$. For any $\varrho = P[\varphi]$ we let $U(\varphi \otimes \phi) = \sum_{nk} c_n \varphi_{nk} \otimes \phi_{nk}$, with distinct $c_n > 0$, be the Schmidt decomposition (Theorem 3.15). Then $\sigma^f = \sum_{nk} |c_n|^2 P[\phi_{nk}]$ is the spectral decomposition, with the spectral projections $\sum_k P[\phi_{nk}]$.

We assume that (ii) and (iii) hold. Note that for those i with $\sigma_i^f = 0$, i.e., $E_\varrho(X_i) = 0$, one also has $\varrho_i^f = 0$, and the equation in (i) is satisfied for all j . It remains to show that (i) holds also for those i for which $\sigma_i^f \neq 0$. For all these i , condition (ii) implies that the Z_i have eigenvalue 1 and one has $Z_j \sigma_i^f = 0$ for $j \neq i$.

Let $Z_i^{(1)}$ be the spectral projection of Z_i associated with eigenvalue 1. Then we have $Z_j^{(1)} \sigma_i^f = \delta_{ij} \sigma_i^f$. It follows that all $Z_i^{(1)}$ commute with all σ_j^f , hence with σ^f and so with all the spectral projections $\sum_k P[\phi_{nk}]$. Moreover, the support subspace of σ^f is contained in the subspace $\oplus_j Z_j^{(1)}(\mathcal{K})$.

Thus, one can choose the orthonormal system $\{\phi_{nk}\}$ such that $Z_i^{(1)} \phi_{nk} = \phi_{nk}$ or $Z_i^{(1)} \phi_{nk} = 0$. Hence there is a rearrangement of this system, $\{\phi_{nk}\} = \{\phi_{i\ell}\}$, such that $Z_i^{(1)} \phi_{i\ell} = \phi_{i\ell}$. It follows that there are corresponding rearrangements $\{\varphi_{i\ell}\} = \{\varphi_{nk}\}$ and $\{d_{i\ell}\} = \{c_n\}$ such that $U(\varphi \otimes \phi) = \sum d_{i\ell} \varphi_{i\ell} \otimes \phi_{i\ell}$. Then $\varrho_i^f = \sum_\ell |d_{i\ell}|^2 P[\varphi_{i\ell}]$. Since the subsets of vectors $\varphi_{i\ell}$ with different values of i are mutually disjoint and therefore orthogonal, one concludes that (i) holds. $\square$

The pointer mixture condition (ii), together with the pointer value-definiteness condition (iii), ensure a decomposition of the final apparatus state in terms of pointer eigenstates. For a projection valued pointer observable this also means that the final apparatus state σ^f is conditionalised with respect to the reading scale $\mathcal{R}$ [1].

The mutual orthogonality of the component states of $\mathcal{S}$, $\varrho_i^f \cdot \varrho_j^f = 0$, corresponds to the optimal distinguishability of these states, which is mandatory for making unique inferences from the pointer readings on properties of the measured object.

22.3 The Problem of Objectification

We still consider a measurement $\mathcal{M}$ of an observable $\mathbf{E}$ with a reading scale $\mathcal{R}$. Assume first that the resulting measurement $(\mathcal{K}, \mathbf{Z}, \sigma, U, r)$ is repeatable; then all of the following conditions are satisfied for any initial state ϱ of the object system:

$$\varrho^f = \sum \mathbf{E}_i(X_i)\varrho_i^f, \quad (22.9)$$

$$\sigma^f = \sum \mathbf{E}_i(X_i)\sigma_i^f, \quad (22.10)$$

$$\mathbf{E}_i\varrho_i^f = \varrho_i^f \text{ for all } i, \quad (22.11)$$

$$\mathbf{Z}_i\sigma_i^f = \sigma_i^f \text{ for all } i, \quad (22.12)$$

$$\varrho_i^f \cdot \varrho_j^f = 0 \text{ for all } i \neq j. \quad (22.13)$$

Note that condition (22.11) expresses the repeatability of the measurement, see Exercise 10 of Sect. 10.9.

Given that the final states (22.9) and (22.10) are both mixtures of eigenvalue-1 eigenstates of the effects $\mathbf{E}_i$ and $\mathbf{Z}_i$, respectively, it is tempting to consider that the discretised pointer observable and the discretised measured observable have assumed one of the values $i \in \mathbb{I}$, with the probabilities $\mathbf{E}_\varrho(X_i)$ indicating the lack of our knowledge about the actual value.

However suggestive this interpretation might appear, it is rendered questionable by Theorem 9.2: although the decompositions (22.9) and (22.10) of ϱ^f and σ^f appear desirable with respect to the measurement, there are, in general, uncountably many other decompositions for these states. It would require compelling physical reasons to explain why these decompositions should reflect the actual state of affairs. Such an explanation appears impossible to obtain given that ϱ^f and σ^f are in fact reduced states of a—generally entangled—state of the compound system consisting of object and probe as it results from the measurement coupling. This compound state does not appear to represent a situation in which it would be admissible to assert that an actual pointer value, say, k had been obtained. Consequently there is no justification of the claim that the probe would then be in the state σ_k^f ; or that the object system would be in the corresponding state ϱ_k^f in which the measured observable would have (in the sense of probabilistic certainty) the corresponding value k . This justification gap is the *measurement problem* of quantum mechanics.

This problem, also known as the *objectification problem*, has been one of the major issues in the foundations of quantum mechanics since the work of von Neumann [2]. It has been studied by many authors, including Wigner [3], Shimony [4], Fine [5], and Brown [6], who contributed to increasingly rigorous and stringent formulations. The perhaps sharpest statement is in the form of a no-go theorem conceived originally by Wigner. This is based on an extension of the pointer mixture condition which, in its currently most general form, amounts to the statement

$$U(\varrho \otimes \sigma)U^* = \sum I \otimes Z_i^{1/2} U(\varrho \otimes \sigma)U^* I \otimes Z_i^{1/2}. \quad (22.14)$$

This would be a necessary feature of the final state if it were possible to argue that the value of the pointer is definite in that state. Clearly, such a decomposition could never be valid for all q unless the initial probe state σ is a mixed state. That this does not help is the content of the following statement, known as the *insolubility theorem* (for the quantum measurement problem) [7].

Theorem 22.2 *Let E be a (nontrivial) observable of a quantum system S . There is no measurement scheme $(\mathcal{K}, \mathbf{Z}, \sigma, U, g)$ for E that satisfies the pointer value-definiteness condition (22.12) and the pointer mixture condition (22.14) for the whole system $S + \mathcal{P}$ for some (nontrivial) reading scale $\mathcal{R}$ and all initial states q of S .*

This theorem shows the assumptions that are left open for challenge.

(A) *Measurement scheme provides pointer value-definiteness and pointer mixture property for $S + \mathcal{P}$.* Some approaches try to dissolve the objectification problem by redefining the notion of objectivity in such a way that the pointer mixture property for $S + \mathcal{P}$ need not be stipulated in order to be allowed to attribute definite values to the pointer. This poses the task to formulate new interpretations of quantum mechanics, such as, for example, some version of the many-worlds interpretation or a form of modal interpretation. This line of approach is systematically analysed in a monograph on interpretations of quantum mechanics by J. Bub [8], which provides a classification of a whole range of possible weakened conceptions of quantum reality, formalised by appropriate value-attribution rules.

Alternatively, one could argue that pointers, being macroscopic quantities, are *genuinely unsharp observables*; that is, the pointer value-definiteness may only be approximately realisable such that the probability $\text{tr}[\sigma_i^f \mathbf{Z}(g^{-1}(X_i))]$ is never equal to unity but rather equal to $1 - \varepsilon$ for some small but nonzero number ε . Hence there may be room for a resolution of the quantum measurement problem by relaxing the requirement of objectification into *unsharp objectification*. This proposal, the implications of which were explored in [9], would, of course, be obsolete if the statement of Theorem 22.2 could be extended to cover such genuinely unsharp pointer observables.

(B) *Measurement coupling is unitary.* New possibilities arise indeed if this assumption is given up. In fact, regarding (tacitly) both S and $\mathcal{P}$ as proper quantum systems, Wigner [3] concluded that the linearity of the quantum mechanical dynamics cannot be maintained if objectification is to be achieved. This argument is taken up in various attempts to introduce *modified quantum dynamics*, thus giving room for spontaneous, autonomous processes leading to the objectivity of some macroscopic observables. We do not discuss such approaches here. For reviews of these so-called continuous spontaneous collapse models, see e.g. [10–12]. Another attempt at making plausible a modification of the dynamical law for measurements was developed in recent years by P. Hajicek [13].

(C) *Apparatus is a proper quantum system.* Instead of requiring the equality (22.14) one could just assume that the two states are equivalent in the sense that they give the same probabilities for all physically admissible effects $A \otimes B$ of the object-probe system, that is,

$$\begin{aligned}\mathrm{tr}[U(\varrho \otimes \sigma)U^*A \otimes B] &= \sum \mathrm{tr}[I \otimes Z_i^{1/2} U(\varrho \otimes \sigma)U^*I \otimes Z_i^{1/2}A \otimes B] \\ &= \sum \mathrm{tr}[U(\varrho \otimes \sigma)U^*A \otimes Z_i^{1/2}BZ_i^{1/2}]\end{aligned}$$

for all A, B and every ϱ . Such a requirement would induce strong reductions of the sets $\mathcal{S}(\mathcal{K})$ and $\mathcal{E}(\mathcal{K})$ rendering smaller sets of possible states and effects of the probe. In particular, it would amount to the probe displaying some *classical observables*. Although such an assumption would justify the pointer mixture condition for $\mathcal{S}+\mathcal{P}$, it raises new difficulties.

We illustrate this consideration for the case of a minimal unitary measurement $\mathcal{M}^m = (\mathcal{K}, \mathbf{Z}, \phi, U_\phi)$ of a discrete sharp observable $\mathbf{A}$ of Example 10.2. If φ is an initial (vector) state of $\mathcal{S}$, then the pointer mixture condition for $\mathcal{S}+\mathcal{P}$ implies that for any $R \in \mathcal{P}(\mathcal{K})$,

$$\langle U(\varphi \otimes \phi) | (P \otimes R) U(\varphi \otimes \phi) \rangle = \sum \langle U(\varphi \otimes \phi) | (P \otimes P[\phi_i] R P[\phi_i]) U(\varphi \otimes \phi) \rangle \quad (22.15)$$

for all $P \in \mathcal{P}(\mathcal{H})$ and for all $\varphi \in \mathcal{H}$, which is to say that R must commute with all $P[\phi_i]$. This simple observation shows that in the present situation, the commutativity of the pointer observable with any other probe observable is a necessary and sufficient condition for the pointer objectification. Since classical observables are always objective, that is, they have a definite though possibly unknown value, we can formulate this result as follows [14].

Theorem 22.3 *Let $\mathcal{M}^m$ be a minimal unitary measurement of an observable $\mathbf{A}$ performed on a proper quantum system $\mathcal{S}$. The pointer mixture condition for $\mathcal{S}+\mathcal{P}$ is obtained if and only if the pointer observable $\mathbf{Z}$ is a classical observable.*

This theorem can be extended to more general measurement situations [10]. It highlights some issues that arise with the assumption of a classical pointer. In particular, one is forced to realise that the classical nature of $\mathbf{Z}$ cannot be reconciled with the assumption that the unitary measurement coupling U arises from a generating Hamiltonian observable H of $\mathcal{S} + \mathcal{P}$ via the relation $U = e^{iH}$. In fact, if H commutes with the pointer observable, then there is no measurement, that is, the measured observable is trivial: $\mathbf{E}(X) = \mathbf{Z}_\phi(X) I$. Since on the other hand the classical nature of $\mathbf{Z}$ is inevitable for the pointer objectification and thus for the measurement, one arrives at the puzzling conclusion that the unitary operator U represents a measuring coupling only if H , the generator of time evolution, is not an observable.

The ‘divorce’ between the algebras of dynamical generators and actual observables (beables) is at the heart of Belavkin’s *nondemolition (or causality) principle*, a form of superselection rule with which quantum dynamics is generalised to encompass indeterministic evolutions, represented mathematically by the so-called stochastic filtering equation. This makes room for the description of measurements with definite yet random outcomes (see [15] for a review of this approach and a survey of its key papers).

To conclude, we note that in recent years there has been a new surge of investigations into modelling quantum measurements and the ‘collapse’ process, with particular emphasis on the macroscopic nature of the measuring apparatus and the environment (see e.g. [16–19]).

22.4 Exercises

1. To prove Theorem 22.1 check the following results: (a) Let $A \geq O$ be a trace class operator. Then $\text{tr}[A] = 0$ if and only if $A = O$. (b) Let $A, B \geq O$ be bounded operators. Then $ABA = O \Leftrightarrow AB^{1/2} = O \Leftrightarrow A^2B = O$.
2. Investigate whether the last claim of Theorem 22.1 is valid for general states ϱ, σ .
3. Prove Theorem 22.2. *Hint.* Recall that for any nontrivial observable E there exist pairs of mutually orthogonal unit vectors φ, ψ such that $E_\varphi \neq E_\psi$ and exploit the ‘inheritance of superpositions’ from the initial object state to the final object-apparatus state under a unitary coupling.

References

1. Cassinelli, G., Zanghì, N.: Conditional probabilities in quantum mechanics. II. Additive conditional probabilities. *Nuovo Cim. B* (11) **79**(2), 141–154 (1984)
2. J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Die Grundlehren der mathematischen Wissenschaften, Band 38. Springer-Verlag, Berlin, 1968. (Reprint of the 1932 original). English translation: *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton (1955, 1996)
3. Wigner, E.P.: The problem of measurement. *Amer. J. Phys.* **31**, 6–15 (1963)
4. Shimony, A.: Approximate measurement in quantum mechanics. ii. *Phys. Rev. D* **9**, 2321–2323 (1974)
5. Fine, A.: Insolubility of the quantum measurement problem. *Phys. Rev. D* **3**(2), 2783–2787 (1970)
6. Brown, H.: The insolubility proof of the quantum measurement problem. *Found. Phys.* **16**(9), 857–870 (1986)
7. Busch, P., Shimony, A.: Insolubility of the quantum measurement problem for unsharp observables. *Stud. Hist. Philos. Mod. Phys.* **27**(4), 397–404 (1996)
8. Bub, J.: *Interpreting the Quantum World*. Cambridge University Press, Cambridge (1999)
9. Busch, P., Jaeger, G.: Unsharp quantum reality. *Found. Phys.* **40**(9–10), 1341–1367 (2010)
10. Busch, P., Lahti, P., Mittelstaedt, P.: *The Quantum Theory of Measurement*. Lecture Notes in Physics Monographs, vol. 2, 2nd edn. Springer, Berlin (1996)
11. Ghirardi, G.: Collapse theories. In: E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. CSLI, Stanford University, winter 2011 edition (2011)
12. Bassi, A., Lochan, K., Satin, S., Singh, T.P., Ulbricht, H.: Models of wave-function collapse, underlying theories, and experimental tests. *Rev. Mod. Phys.* **85**(2), 471–527 (2013)
13. Hajicek, P.: Incompleteness of Measurement Apparatuses. [arXiv:1411.5524](https://arxiv.org/abs/1411.5524) (2014)
14. Beltrametti, E., Cassinelli, G., Lahti, P.: Unitary measurements of discrete quantities in quantum mechanics. *J. Math. Phys.* **31**(1), 91–98 (1990)
15. Belavkin, V.: On the Dynamical Solution of Quantum Measurement Problem. [arXiv:0512187](https://arxiv.org/abs/0512187) (2005)

16. Sewell, G.: On the mathematical structure of quantum measurement theory. *Rep. Math. Phys.* **56**(2), 271–290 (2005)
17. Allahverdyan, A.E., Balian, R., Nieuwenhuizen, T.M.: Understanding quantum measurement from the solution of dynamical models. *Phys. Rep.* **525**(1), 1–166 (2013)
18. Narnhofer, H., Wreszinski, W.F.: On reduction of the wave-packet, decoherence, irreversibility and the second law of thermodynamics. *Phys. Rep.* **541**(4), 249–278 (2014)
19. Liuzzo-Scorpo, P., Cuccoli, A., Verrucchi, P.: Parametric description of the quantum measurement process. *EPL (Europhys. Lett.)* **111**(4), 40008 (2015)

Chapter 23

Axioms for Quantum Mechanics

In this final chapter we address the question of justifying the Hilbert space formulation of quantum mechanics. We show that this theory can essentially be derived from physically plausible assumptions using the general frame of statistical dualities sketched briefly already in the Introduction (Sect. 23.1). In line with the central theme of the book, the axioms to be presented are motivated by properties and limitations of preparations and measurements.

In Sect. 23.1 the statistical duality and its representations are discussed in greater detail and two basic sets of axioms will there be stated, the weaker ones, Axioms 1 and 2, and the stronger ones, Axioms 1' and 2'. This general frame will be further specified along different but closely related approaches.

The first approach (Sect. 23.2), starting with Axioms 1 and 2, builds on the *orthogonality postulate*, which expresses the requirement of the existence of a sufficiently regular subset of observables (equivalence classes of measurements). The structures emerging from these axioms have been studied extensively in variants of the quantum logic approach to quantum mechanics.

The second approach starts with the stronger Axioms 1' and 2' adding to them the *projection postulate* (Sect. 23.3) which formalises the assumption of the existence of a sufficiently large set of pure ideal first kind yes-no measurements and identifies this set with the resulting subset of decision effects. This approach was developed in the so-called convexity or operational approaches.

Both sets of postulates lead to the same kind of ortho-order structure, which admits a 'vector space coordinatisation' in terms of a family of subspaces of an appropriate vector space (Sect. 23.4). This derives from the fundamental representation theorem of projective geometries and from the fact that a Hermitian space is orthomodular (in a geometric sense) exactly when the lattice of its closed subspaces is orthomodular (with respect to the ortho-order structure). The final step in identifying the relevant vector space as an infinite-dimensional Hilbert space is obtained in Sect. 23.4 from a theorem due to Maria Solér. In Sect. 23.5 the physical content of the conditions required by this theorem are analysed in terms of symmetries.

In closing the chapter, we briefly discuss some arguments in favour of the choice of the complex field over the real numbers or quaternions (Sect. 23.6).

The historic paper [1] of Birkhoff and von Neumann, entitled *The Logic of Quantum Mechanics*, marks the beginning of the investigations on the mathematical and conceptual foundations of quantum mechanics which go under the title quantum logic. The literature of the field is very rich. In addition to the influential lecture notes of Mackey [2] we mention here only some representative monographs to indicate the diversity of the field [3–10].

The papers of Ludwig [11], Mielnik [12, 13], Davies and Lewis [14], and Edwards [15–17] have strongly influenced the development of the convexity or operational approach to quantum mechanics. In addition to these original papers, the monumental work of Ludwig [18] as well as the books [19] and [20] are valuable sources for the physical and mathematical ideas underlying the convexity approaches. A closely related approach is the empirical logic framework developed by Foulis and Randall for the study of the manuals of physical operations (e.g., [21, 22]).

In recent years, quantum information theory has renewed interest in the foundations of quantum mechanics, and some new ideas have been proposed for an axiomatic foundation of quantum mechanics, see, e.g., [23, 24] and the many references given therein. Though interesting, these investigations deal only with the so-called finite level systems, which are only a part of quantum mechanics. Thus the more general question is left open of how to derive quantum mechanics in infinite-dimensional Hilbert space, which is needed, for instance, to describe physical systems in four-dimensional spacetime. The present chapter is devoted to this question, and the presentation follows mostly the survey article [25].

23.1 Statistical Duality and Its Representation

Statistical Duality

A general probabilistic formulation of a physical theory builds on the concepts of states and observables and on the idea of statistical causality merging the two entities into a probability measure. *States* are understood as equivalence classes of preparations of a physical system, *observables* as equivalence classes of measurements on it, and the requirement of *statistical causality* asserts that any state α and any observable E determine a probability measure $p(\alpha, E, \cdot)$ defined on a σ -algebra $\mathcal{A}$ of subsets of a (nonempty) set Ω , with $p(\alpha, E, X)$ giving the probability that a result is registered in the set X when a measurement of E is performed on the system prepared in state α . An observable E thus comes with a value space Ω together with a collection of test sets, $X \in \mathcal{A}$, within which the results are counted. To make this explicit, we may also write $(E, \Omega, \mathcal{A})$ for E .

Let S and O be the sets of all states and all observables of the system. We call the pair (S, O) together with the probability function $p : (\alpha, E) \mapsto p(\alpha, E, \cdot)$ a *statistical duality*. In an axiomatic approach one aims to introduce physically motivated

structures for the sets $\mathcal{S}$ and $\mathcal{O}$ so that the form of the probability measures $p(\alpha, \mathbf{E}, \cdot)$, $\alpha \in \mathcal{S}$, $\mathbf{E} \in \mathcal{O}$, gets determined.

State Space

Convex Structure

The set $\mathcal{S}$ of states can be equipped with a convex structure reflecting the possibility of combining preparations (and thus states) into new preparations (and thus states) by statistically mixing them. Indeed, for any $\alpha, \beta \in \mathcal{S}$ and $0 \leq \lambda \leq 1$ one may define a state $\langle \lambda, \alpha, \beta \rangle$ through the λ -convex combination of the probability measures $p(\alpha, \mathbf{E}, \cdot)$ and $p(\beta, \mathbf{E}, \cdot)$, $\mathbf{E} \in \mathcal{O}$,

$$p(\langle \lambda, \alpha, \beta \rangle, \mathbf{E}, \cdot) = \lambda p(\alpha, \mathbf{E}, \cdot) + (1 - \lambda) p(\beta, \mathbf{E}, \cdot). \quad (23.1)$$

The existence of such a state $\langle \lambda, \alpha, \beta \rangle$ is built in the assumption that preparations can be statistically mixed to produce new preparations, and the uniqueness of $\langle \lambda, \alpha, \beta \rangle$ follows from the statistical completeness built in the notions of states (as equivalence classes of preparations) and observables (as equivalence classes of measurements).

The existence of the function $[0, 1] \times \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$ with the property (23.1) defines what is known as a *convex structure*, and it is then a mathematical convenience to consider $\mathcal{S}$ as properly placed in a real vector space U so that we may simply write $\langle \lambda, \alpha, \beta \rangle = \lambda\alpha + (1 - \lambda)\beta$ [26, Theorem 2], see also [27, 28].

Using σ -convex combinations of probability measures, one may also introduce σ -convex combinations of states (α_i) with weights (λ_i) , $\lambda_i \geq 0$, $\sum \lambda_i = 1$, through $\sum \lambda_i p(\alpha_i, \mathbf{E}, \cdot)$, $\mathbf{E} \in \mathcal{O}$, with the obvious requirement that the series $\sum \lambda_i p(\alpha_i, \mathbf{E}, X)$ is convergent for each $\mathbf{E}$ and X . If there is an $\alpha \in \mathcal{S}$ such that $p(\alpha, \mathbf{E}, \cdot) = \sum \lambda_i p(\alpha_i, \mathbf{E}, \cdot)$ for all $\mathbf{E} \in \mathcal{O}$, we say that α is a σ -convex combination of the states (α_i) with weights (λ_i) and we write $\alpha = \sum \lambda_i \alpha_i$. Again, if such a state exists it is unique. As will be seen from Theorem 23.1, it is a mathematically convenient idealisation to assume that the set of states is closed under σ -convex combinations.

The convex structure of $\mathcal{S}$ allows the distinction between the *pure* and *mixed* states, corresponding to the extreme elements and nonextreme elements of $\mathcal{S}$, respectively. We let $\mathcal{P} = \text{ex}(\mathcal{S})$ denote the set of all pure states in $\mathcal{S}$. The existence of pure preparations, and thus of pure states, is another natural assumption supported equally well by everyday experience and sophisticated quantum experiments;

We summarise the above discussion in the first axiom.

Axiom 1 The set of all states of a physical system described by the statistical duality $(\mathcal{S}, \mathcal{O}, p)$ forms a convex subset of a real vector space.

We specify later the assumptions concerning σ -convex combinations of states and the existence of a sufficiently large set of pure states.

$\mathcal{S}$ as a Base for a Generating Cone

Let $K = \{\lambda\alpha \mid \lambda \in \mathbb{R}^+, \alpha \in \mathcal{S}\} = \cup_{\lambda \geq 0} \lambda\mathcal{S} \subset U$ be the cone defined by $\mathcal{S}$. We assume now that this cone is a proper cone, that is, $K \cap -K = \{0\}$, and that each $\gamma \in K$,

$\gamma \neq 0$, has a unique representation as $\gamma = \lambda\alpha$ for some $\lambda > 0$ and $\alpha \in S$. Let $V = K - K$ be the real vector space generated by K (possibly a subspace of U). Then $K = V^+ = \{v \in V \mid v \geq 0\}$ and S is a base of the cone K . The existence of a base for a generating cone of a real vector space V is known to be equivalent to the existence of a strictly positive linear functional on V [29, Lemma 2]. We let $e : V \rightarrow \mathbb{R}$ be the functional defined by S so that

$$S = \{\alpha \in V^+ \mid e(\alpha) = 1\}, \quad (23.2)$$

and call it the *intensity* functional. The number $e(\alpha)$ will be informally referred to as the intensity of the state α ; this terminology is borrowed from the language of particle physics experimentation, where the beam intensity decreases in proportion to the probability that the particles may enter one of the available (scattering or decay) channels open to them.

The physical interpretation attached to S can thus be extended to the positive cone V^+ : any $\lambda\alpha$, $\lambda \in \mathbb{R}^+$, $\alpha \in S$, represents a state of the system obtained from α by changing its intensity. The elements of S will be distinguished as normalised states, whereas the term ‘state’ will be extended to refer to all elements of V^+ , including the empty state, the null element of V . The linear operations $(\alpha, \beta) \mapsto \alpha + \beta$ and $(\lambda, \alpha) \mapsto \lambda\alpha$, $\alpha, \beta \in V^+$, $\lambda \in \mathbb{R}^+$, preserve their original interpretation as mixing and intensity changing, respectively. In particular, the term *pure state* can thus refer to an element of $\text{ex}(S)$ or of $\{\lambda\alpha \mid \lambda \in \mathbb{R}^+, \alpha \in \text{ex}(S)\}$. This extension will be accepted as a mathematically convenient way of speaking, with no physical implications.

Let $\text{conv}(S \cup -S)$ denote the convex hull of the set $S \cup -S$. This set is convex, absorbing (that is, for any $v \in V$ there is a $\lambda > 0$ such that $v \in \lambda \text{conv}(S \cup -S)$), and balanced (that is, $\lambda v \in \text{conv}(S \cup -S)$ for all $-1 \leq \lambda \leq 1$ and $v \in \text{conv}(S \cup -S)$). Therefore its Minkowski functional (or gauge) $p_S : V \rightarrow \mathbb{R}$, defined as

$$p_S(v) = \inf \{\lambda > 0 \mid v \in \lambda \text{conv}(S \cup -S)\},$$

is a seminorm [30, Theorem II.1.4]. Clearly, $p_S(\alpha) = e(\alpha)$ for all $\alpha \in V^+$, and we note that for all $v \in V$,

$$p_S(v) = \inf \{e(\alpha) + e(\beta) \mid \alpha, \beta \in V^+, v = \alpha - \beta\}.$$

If this seminorm is a norm, then (V, V^+, S) or just (V, S) is a *base norm space*. The following result, due to Edwards and Gerzon [31], is now crucial:

Theorem 23.1 *If the set S of all states of the system forms a base for a generating cone of a vector space V and is σ -convex, then its Minkowski functional $p_S : V \rightarrow \mathbb{R}$ is a norm with respect to which V is a Banach space.*

There is another technical point to be noted. According to a proposition of Ellis [32], if (V, V^+, S) is a base norm space, then also $(V, \overline{V^+}, \overline{S})$ (where $\overline{A}$ denotes the norm closure of a subset A of V) is a base norm space and its norm coincides with

the norm of $(V, \mathcal{S})$. We note, in addition, that $\mathcal{S}$ is closed if and only if V^+ is closed. With the risk of adding some nonphysical elements in the set $\mathcal{S}$ we now formulate an alternative stronger assumption concerning the set of states of a statistical duality, as first formulated in [14].

Axiom 1' The set of all states of a physical system described by the statistical duality $(\mathcal{S}, \mathcal{O}, p)$ is represented by a norm closed generating cone V^+ of a base norm Banach space $(V, \mathcal{S})$.

We have thus formulated two axioms reflecting basic assumptions on the set of states, the weaker axiom serving as a starting point for the quantum logic approach (Sect. 23.2) and the stronger axiom defining the beginnings of the convexity or state space approach (Sect. 23.3).

Experimental Functions

Affine Maps $\mathcal{S} \rightarrow [0, 1]$

With every observable $(\mathbf{E}, \Omega, \mathcal{A}) \in \mathcal{O}$ one may consider the family of all ordered pairs $\langle \mathbf{E}, X \rangle$, $X \in \mathcal{A}$, called experimental pairs. With each such pair we can associate the statement ‘a measurement of $\mathbf{E}$ yields a result in the set X ’, denoted by $(\mathbf{E}, X)$ and called an *experimental statement*. Then the number $p(\alpha, \mathbf{E}, X)$, $\alpha \in \mathcal{S}$, gives the probability for the statement $(\mathbf{E}, X)$ to be true in the state α . Experimental statements $(\mathbf{E}, X)$ and $(\mathbf{F}, Y)$ are said to be *equivalent* if for all $\alpha \in \mathcal{S}$,

$$p(\alpha, \mathbf{E}, X) = p(\alpha, \mathbf{F}, Y). \quad (23.3)$$

This defines an equivalence relation in the set of all experimental statements $\{(\mathbf{E}, X) \mid \mathbf{E} = (\mathbf{E}, \Omega, \mathcal{A}) \in \mathcal{O}, X \in \mathcal{A}\}$. Let $\mathbf{E}_0$ denote the set of all equivalence classes $|(\mathbf{E}, X)|$ of the statements $(\mathbf{E}, X)$. A given element of $\mathbf{E}_0$ is denoted by a letter a and is called an (*experimental*) *proposition*. The experimental proposition a is a set of experimental statements equivalent among themselves with respect to p .

There is a fundamental difference between $(\mathbf{E}, X)$ and $|(\mathbf{E}, X)|$. Namely, $(\mathbf{E}, X)$ is nothing more than a statement saying that a measurement of $\mathbf{E}$ yields a result in X , and it does not depend on p ; in contrast, $|(\mathbf{E}, X)|$ represents the proposition that any two statements $(\mathbf{E}_1, X_1)$ and $(\mathbf{E}_2, X_2)$ from $|(\mathbf{E}, X)|$ are equivalent to $(\mathbf{E}, X)$, that is, for every $\alpha \in \mathcal{S}$, $p(\alpha, \mathbf{E}_1, X_1) = p(\alpha, \mathbf{E}_2, X_2) = p(\alpha, \mathbf{E}, X)$. The experimental proposition $|(\mathbf{E}, X)|$ clearly depends on p ; it should be written more exactly $|(\mathbf{E}, X)|_p$. Hence to specify the proposition $|(\mathbf{E}, X)|_p$ one has to measure all the experimental statements $(\mathbf{F}, Y)$ and find those equivalent to $(\mathbf{E}, X)$. Hence the experimental proposition $a = |(\mathbf{E}, X)|_p$ is a subset of all experimental statements depending on p , typically much bigger than the one-element set $\{(\mathbf{E}, X)\}$.

Each experimental proposition $a \in E_0$ defines a real valued function, viz., an *experimental function*

$$f_a : S \rightarrow [0, 1], f_a(\alpha) = p(\alpha, E, X), \quad (E, X) \in a, \quad (23.4)$$

and we let $E \subset [0, 1]^S$ denote the set of all such functions. If $f \in E$, then $f(\alpha), \alpha \in S$, is the probability that the proposition $a = |(E, X)|$, with $f = f_a$, is true in state α , that is, an E -measurement in state α yields a result in X .

Consistently with the definition (23.1) of mixed states, it is natural to assume that the experimental functions are affine; thus we have the following axiom.

Axiom 2 The set E of experimental functions of a statistical duality (S, O, p) is a subset of the set of affine functions $S \rightarrow [0, 1]$.

We let 0 and e denote the constant zero and one functions $S \rightarrow [0, 1]$, respectively. Clearly, $0, e \in E$, and if $f \in E$ then also $f^\perp = e - f \in E$ so that for any $\alpha \in S$, $f(\alpha) + f^\perp(\alpha) = 1$. Moreover, as a set of real valued functions, E is partially ordered in a natural way, that is, for any $f, g \in E$, $f \leq g$ if and only if $f(\alpha) \leq g(\alpha)$ for all $\alpha \in S$, the operational content being given by (23.4).

The set E has the order bounds 0 and e and the mapping $E \ni f \mapsto f^\perp \in E$ is an order reversing involution. However, it is not an orthocomplementation, that is, the greatest lower bound of a pair $(f, f^\perp)$ need not be 0.

Any observable E can now be represented as an E -valued set function $E : X \mapsto f_{|(E, X)|}$ such that for each $\alpha \in S$, the set function $X \mapsto \alpha(f_{|(E, X)|}) = f_{|(E, X)|}(\alpha) = p(\alpha, E, X)$ is a probability measure. Moreover, any $f \in E$ is in the range of some observable E . In this sense we can say that the set O of all observables is a *surjective* set of E -valued set functions.

We will develop two distinct approaches towards specifying further the structure on E , for both of which the following notions of orthogonality will be important.

Definition 23.1 A sequence (finite or countably infinite) of experimental functions $f_1, f_2, \dots$ is *orthogonal* if there is an experimental function g such that $g + f_1 + f_2 + \dots = e$.

Definition 23.2 A sequence (finite or countably infinite) of experimental functions $f_1, f_2, \dots$ is *pairwise orthogonal* if $f_i + f_j \leq e$ for all $i, j \in \mathbb{N}$, $i \neq j$.

An orthogonal sequence is pairwise orthogonal but in general not conversely. Still the concepts of orthogonality and pairwise orthogonality have the following common property: if for some state α one of the experimental statements f_i is true, that is, $f_i(\alpha) = 1$, then all the other experimental statements are false in that state, that is, $f_j(\alpha) = 0$ for all $j \neq i$. This supports the hypothesis that the statistical duality (S, O, p) might have a substructure where the two notions coincide; we return to this in Sect. 23.2.

Positive Unit Bounded Functionals on V

Further properties of experimental functions can be obtained under the assumption of Axiom 1'. Indeed, in this case any $f \in E$ has a unique extension to a positive continuous linear functional on V (again denoted by f) that is bounded by e . The set E of experimental functions is then a subset of the order interval $[0, e]$ of the order unit Banach space $(V^*, [0, e])$. At the risk of adding some new elements in the set O of all observables one may assume that actually $E = [0, e]$.¹ We express this stronger assumption concerning experimental functions as a further possible axiom.

Axiom 2' The set E of all experimental functions coincides with the order interval $[0, e]$ of the dual Banach space of (V, S) .

This axiom has a simple but important consequence: for any two $f, g \in E$, if $f \leq g^\perp$, then also $f + g \in E$. Clearly, then $f, g \leq f + g$, but this does not mean that their least upper bound $f \vee g$ would exist in E , and even if it did exist, it need not be equal to $f + g$.

Another important structure of the set E arising from Axiom 2' is its convexity: for any $f, g \in E$ and $0 \leq \lambda \leq 1$, $\lambda f + (1 - \lambda)g \in E$. Since the order interval $E \subset V^*$ is also compact (by the Banach–Alaoglu theorem), the Krein–Milman theorem applies and the closure of the convex hull of the extremal elements of E is the whole set of experimental functions, that is, $\overline{\text{co}}(\text{ex}(E)) = E$.

Operations

The number $p(\alpha, E, X) = f_{(E, X)}(\alpha)$ is the probability that a measurement of $E \in O$ in the state $\alpha \in S$ leads to a result in the set $X \in \mathcal{A}$. The occurrence of a particular outcome is generally accompanied by an unavoidable change of the system. Indeed, a measurement may destroy the system or cause a change in its state. This measurement dynamics is in addition to the autonomous (or free) time evolution of the system.

In order to describe the state changes of a system due to the process of measurement, we now build on Axioms 1' and 2'. Allowing for the possibility that the intensity of a state may change in the process, the change of the system from an input state α to an output state α' will be assumed to be given by a map $\phi : V^+ \rightarrow V^+$. The case $\alpha' = 0$ is taken to represent the fact the system is destroyed in the intervention. As illustrated in this book, realistic physical measurement situations are modelled by various types of state changes $\alpha \mapsto \phi(\alpha)$.

In line with the intended probabilistic interpretation of the intensity of a state, we anticipate that state changes are intensity non-increasing, so that for a normalised state α , the number $e(\phi(\alpha)) \in [0, 1]$ may be taken to represent the probability of the measurement outcome associated with the state change described by ϕ . Thus, we assume that for each state $\alpha \in V^+$,

$$e(\phi(\alpha)) \leq e(\alpha). \quad (23.5)$$

¹Clearly, this assumption could also have been included in Axiom 1 but we refrained from doing so.

Next we consider the effect a measurement can have on a mixed state $\beta = \lambda_1\alpha_1 + \lambda_2\alpha_2$, with $\lambda_1, \lambda_2 \in \mathbb{R}^+$ and $\alpha_1, \alpha_2 \in V^+$. In a change ϕ the state β transforms to $\phi(\beta)$ whereas the states α_1, α_2 transform to $\phi(\alpha_1)$ and $\phi(\alpha_2)$ of which one may form the mixture $\lambda_1\phi(\alpha_1) + \lambda_2\phi(\alpha_2)$. We assume that measurements induce only changes that are represented by maps ϕ for which these states are always the same, that is, for any $\lambda_1, \lambda_2 \in \mathbb{R}^+, \alpha_1, \alpha_2 \in V^+$,

$$\phi(\lambda_1\alpha_1 + \lambda_2\alpha_2) = \lambda_1\phi(\alpha_1) + \lambda_2\phi(\alpha_2). \quad (23.6)$$

We leave it as a simple exercise to check that any map $\phi : V^+ \rightarrow V^+$ with the properties (23.5) and (23.6) has a unique extension to a positive linear contracting mapping of V into V . We denote this extension by the same letter ϕ and we call such a mapping an *operation*.

Remark 23.1 There are physical situations, such as detectors with memory effects, that are modelled by nonlinear maps ϕ , so that the states $\phi(\lambda_1\alpha_1 + \lambda_2\alpha_2)$ and $\lambda_1\phi(\alpha_1) + \lambda_2\phi(\alpha_2)$ are different. Equally, an amplifier is a device that typically increases the intensity of at least some states. To describe such devices condition (23.5) should be relaxed. The paper of Mielnik [12] contains an extensive analysis of possible state changes, including some nonlinear processes. This highlights the fact that the above restriction to affine maps ϕ is a nontrivial assumption on the nature of the experimental setups with which statistics are collected.

Let $\mathcal{O}$ be the set of all operations. The sequential application of any two operations defines a new operation giving $\mathcal{O}$ the structure of a noncommutative semigroup. Another physically relevant structure of $\mathcal{O}$ is that of convexity: for any two $\phi_1, \phi_2 \in \mathcal{O}$ and any $0 \leq \lambda \leq 1$, $\lambda\phi_1 + (1 - \lambda)\phi_2 \in \mathcal{O}$. This allows one to single out the *extremal* operations, that is, operations that cannot be obtained as nontrivial mixings of any other operations.

An operation $\phi \in \mathcal{O}$, when combined with the intensity functional e , defines an experimental function $e \circ \phi \in \mathcal{E}$. On the other hand, if $f \in \mathcal{E}$, then fixing a $\beta \in \mathcal{S}$ and defining $\phi(\alpha) = f(\alpha)\beta$, $\alpha \in V$, one observes that $\phi \in \mathcal{O}$ and $e \circ \phi = f$. The set of functionals $e \circ \phi$, $\phi \in \mathcal{O}$, thus coincides with the set of experimental functions. Due to this coincidence, an experimental function is also called *effect*: $f \in \mathcal{E}$ is the effect of any operation $\phi \in \mathcal{O}$ such that $f = e \circ \phi$. The extremal elements of $\mathcal{E}$ are called *decision effects* or *sharp effects*; hence, an effect is either sharp (extremal) or unsharp (non-extremal).

Let f be a decision effect. If $f = e \circ (\lambda\phi_1 + (1 - \lambda)\phi_2)$ for some $\phi_1, \phi_2 \in \mathcal{O}$, $\lambda \in [0, 1]$, then $e \circ \phi_1 = e \circ \phi_2$, that is, the operations ϕ_1 and ϕ_2 are *isotonic*. Actually, the relation $e \circ \phi = e \circ \psi$, $\phi, \psi \in \mathcal{O}$, defines an equivalence relation in $\mathcal{O}$ and one may immediately confirm that there is a one-to-one onto correspondence between the set of effects and the isotony classes of operations. We let $[\phi]_f$ denote the isotony class of operations $\phi \in \mathcal{O}$ associated with the effect $f \in \mathcal{E}$.

One may now define an *instrument* as an operation valued set function $\mathcal{A} \ni X \mapsto \phi_X \in \mathcal{O}$ for which $X \mapsto e \circ \phi_X$ is an observable, that is, $X \mapsto e(\phi_X(\alpha))$ is a probability measure for each $\alpha \in \mathcal{S}$. By definition, any instrument defines an observable,

but the converse holds also: any observable E arises from some instrument such that $E(X) = e \circ \phi_X$. Calling two instruments isotonic if they define the same observable one again has that the isotony classes of instruments are in one-to-one onto correspondence with the observables of the system.

Preparing the system in a state $\alpha \in S$, acting on it by an operation $\phi \in \mathcal{O}$, and detecting the (probabilistic) effect $e \circ \phi \in E$ —these steps comprise the main elements in the general probabilistic description of a physical experiment. Accordingly, the operational approach towards a justification of the specific structures of the quantum probabilistic description builds on the statistical duality of states and observables. To specify further the structures of the *probabilistic model* $(S, \mathcal{O}, E)$, one may proceed in various ways by posing additional conditions on either of the sets S , $\mathcal{O}$, or E . The remark below gives an indication of how to reach classical descriptions out of this general probabilistic model. Our aim is to pose conditions which lead to the quantum probabilistic model.

Remark 23.2 The space V is partially ordered by the cone V^+ defined by the base S . If this order is a lattice order, then S is a (Choquet) simplex [33], a structure considered to be characteristic of classical descriptions. If $V = V^+ - V^+$ is a vector lattice, then also its dual V^* is a vector lattice. In this case, the order interval E is a lattice as well, and the set $\text{ex}(E)$ of extremal effects is a Boolean lattice with $f \mapsto f^\perp = e - f$ as the orthocomplementation [34]—another characteristic of classical descriptions.

In the next two sections we shall follow two alternative ways of specifying further the statistical duality $(S, \mathcal{O}, p)$. We start with an approach based on Axioms 1 and 2. In Sect. 23.3 we build on the stronger Axioms 1' and 2'.

23.2 Quantum Logic

Mackey's approach to quantum logic proceeds by endowing the statistical duality $(S, \mathcal{O}, p)$ with additional structure. In particular it is assumed that there is a large enough subset of observables $\tilde{\mathcal{O}} \subset \mathcal{O}$ such that the order structure of the associated subset of experimental functions becomes sufficiently regular. Here we make precise these assumptions and the resulting properties. Note that our only starting assumption is that the set of states is convex, and we allow the possibility that the set of experimental functions is a subset of the set of affine functions from S to $[0, 1]$.

Orthogonality Postulate

Any subset of the set $\mathbf{O}$ of all observables defines corresponding subsets of the sets $\mathbf{E}_0$ and $\mathbf{E}$ of propositions and experimental functions. The basic assumption in Mackey's approach to quantum logic [2, Axiom V] can now be stated as the requirement of the existence of a (nonempty) subset $\tilde{\mathbf{O}} \subset \mathbf{O}$ such that in the resulting subset $\mathbf{L} \subset \mathbf{E}$ of the experimental functions the two notions of orthogonality given in Definitions 23.1 and 23.2 coincide. We call this assumption the *orthogonality postulate* and formulate it as our next axiom.

Axiom 3 The set $\mathbf{O}$ of observables of the statistical duality $(S, \mathbf{O}, p)$ contains a (nonempty) subset $\tilde{\mathbf{O}}$ such that in the resulting set $\mathbf{L}$ of experimental functions a sequence $f_1, f_2, \dots \in \mathbf{L}$ is orthogonal (in $\mathbf{L}$) if and only if it is pairwise orthogonal (in $\mathbf{L}$).

This axiom has important implications concerning the order structure of the set $\mathbf{L}$. First of all, it guarantees that for any two mutually orthogonal elements $f, g \in \mathbf{L}$ also $f + g \in \mathbf{L}$. Moreover, it implies that the map $f \mapsto f^\perp$ is an orthocomplementation and it turns $(\mathbf{L}, \leq, {}^\perp)$ into an orthomodular σ -orthocomplemented partially ordered set, with $0, e \in \mathbf{L}$ as the order bounds.

We observe that for any two $f, g \in \mathbf{L}$, the set of their lower (upper) bounds in $\mathbf{L}$ is smaller than the corresponding set in $\mathbf{E}$. Therefore, $f \wedge g$ may exist in $\mathbf{L}$ without existing in $\mathbf{E}$.

Theorem 23.2 ([35]) *Let $(S, \mathbf{O}, p)$ be a statistical duality and let $\tilde{\mathbf{O}}$ be a (nonempty) subset of $\mathbf{O}$ such that the associated set $\mathbf{L}$ of experimental functions satisfies Axiom 3. The set $\mathbf{L}$ is an orthocomplemented orthomodular σ -orthocomplete partially ordered set with respect to the natural order of real functions and the complementation $f^\perp = e - f$.*

Proof Clearly, $\mathbf{L}$ is partially ordered by $\leq$ and $0, e \in \mathbf{L}$. With $f \in \mathbf{L}$, one also has $f^\perp = e - f \in \mathbf{L}$, and $\mathbf{L} \ni f \mapsto f^\perp \in \mathbf{L}$ is an order reversing involution.

Let $f_1, f_2 \in \mathbf{L}$, and assume that $f_1 + f_2 \leq e$. Then by Axiom 3, $f = f_1 + f_2 \in \mathbf{L}$. To show that f is the least upper bound of f_1 and f_2 in $\mathbf{L}$, assume that $g \in \mathbf{L}$ is such that $f_1 \leq g$ and $f_2 \leq g$. Then also $f_1 + g^\perp \leq e$ and $f_2 + g^\perp \leq e$, and thus $f_1 + f_2 + g^\perp \in \mathbf{L}$, so that $f_1 + f_2 \leq g$, that is $f_1 + f_2 = f_1 \vee_L f_2$. By induction one shows that $f_1 + \dots + f_n = f_1 \vee_L \dots \vee_L f_n$ for any pairwise orthogonal set $\{f_1, \dots, f_n\} \subset \mathbf{L}$. Let (f_i) be a sequence of mutually orthogonal elements in $\mathbf{L}$, so that by assumption $f = f_1 + f_2 + \dots \in \mathbf{L}$. Clearly, $f_i \leq f$ for each i . Let $g \in \mathbf{L}$ be such that $f_i \leq g$ for all i . Since for any n , $f_1 + \dots + f_n = f_1 \vee_L \dots \vee_L f_n$ we thus have $f_1 + \dots + f_n \leq g$ for any $n = 1, 2, \dots$ and therefore $f = f_1 + f_2 + \dots \leq g$.

For any $f \in \mathbf{L}$, $f + f^\perp \leq e$ and thus $e = f + f^\perp = f \vee_L f^\perp$. By De Morgan's laws we also have $f \wedge_L f^\perp = 0$ for any $f \in \mathbf{L}$. This concludes the proof that $\mathbf{L}$ is orthocomplemented and σ -orthocomplete.

To show orthomodularity, we need to show that for $f \leq g$, $f, g \in \mathbf{L}$, one has $g = f \vee_L (g \wedge_L f^\perp)$. If $f \leq g$, then $f \vee_L g^\perp = f + g^\perp = f + (e - g)$ and $h = (f \vee_L$

$g^\perp)^\perp = g - f \in L$. Hence $f + h = g \leq e$, and thus $f \vee_L h = f + h$, so that $f \vee_L (g \wedge_L f^\perp) = f \vee_L (f \vee_L g^\perp)^\perp = f + h = g$. $\square$

We call L the *logic* of p . Henceforth we simply write $f \wedge g$ for $f, g \in L$ instead of $f \wedge_L g$, and similarly for $f \vee g$, whenever the meet (join) exists in L .

Remark 23.3 For any $f \in L$ other than 0, it is the case that $\lambda f \notin L$ if $0 < \lambda < 1$. To see this, assume to the contrary that there is an $f \in L$, $0 \neq f$ and some $\lambda \in (0, 1)$ such that $\lambda f \in L$. Since $\lambda f \leq f$, then $\lambda f + (e - f) \in L$, and hence $e - (\lambda f + (e - f)) = (1 - \lambda)f \in L$. Furthermore, $\lambda f + (1 - \lambda)f = f \leq e$, which entails that λf and $(1 - \lambda)f$ are also pairwise orthogonal, so that their sum should equal their least upper bound in L , which is a contradiction. In particular, L is not convex. $\triangleleft$

Each observable $E = (E, \Omega, \mathcal{A}) \in \tilde{\mathcal{O}}$ determines a unique L -valued measure $M_E : \mathcal{A} \rightarrow L$ defined by $M_E(X) = f_{\{E(X)\}}$. By Theorem 23.2, M_E is in fact a σ -homomorphism; this implies that $M_E(\mathcal{A})$ is a Boolean sub- σ -algebra of L . We identify M_E with E .

Each state $\alpha \in S$ determines a unique probability measure $m_\alpha : L \rightarrow [0, 1]$ defined by $m_\alpha(f) = f(\alpha)$, which means, in particular, that for any pairwise orthogonal sequence (f_i) in L , $m_\alpha(\vee f_i) = \sum_i m_\alpha(f_i)$. Again, we identify m_α with α .

The family of L -valued measures $M_E, E \in \tilde{\mathcal{O}}$, is *surjective* (that is, any $f \in L$ is of the form $f = M_E(X)$ for some $M_E(X)$), and the family of probability measures $m_\alpha, \alpha \in S$, is *order determining*, that is, for any $f, g \in L, f \leq g$ if and only if $m_\alpha(f) \leq m_\alpha(g)$ for all $\alpha \in S$. For each $\alpha \in S, E \in \tilde{\mathcal{O}}, X \in \mathcal{A}$ we have

$$p(\alpha, E, X) = m_\alpha(M_E(X)).$$

We note that also the converse result is true: If L is an arbitrary orthocomplemented partially ordered set admitting an order determining set of probability measures S , and $\tilde{\mathcal{O}}$ is a surjective set of L -valued measures, then the function p , defined as $p(\alpha, M, X) = \alpha(M(X))$ for all $\alpha \in S, M \in \tilde{\mathcal{O}}, X \in \mathcal{A}$, is a probability function satisfying the orthogonality postulate and the logic of p is isomorphic to L [35].

The sets L_0 and L of experimental propositions and functions are in one-to-one onto correspondence, and one may naturally carry over the order and complementation of L to L_0 : for any $a, b \in L_0, a \leq b$ if and only if $f_a \leq f_b$, and we let $a^\perp$ stand for the proposition corresponding to the function $e - f_a$. Thus, under the assumption of Axiom 3, we may equally well consider L_0 as an orthocomplemented σ -orthocomplete orthomodular partially ordered set (of propositions), with S as an order determining set of probability measures of L_0 . From now on, we do not distinguish between L_0 and L , and we also consider the elements of L interchangeably as functions on S and the elements of S as functions on L : $a(\alpha) = \alpha(a)$. Also, in using the symbol L , we always mean the structure $(L, \leq, ^\perp)$, with the order bounds 0 and e corresponding to the absurd (always false) and trivial (always true) propositions, respectively. Moreover, we view the observables $(E, \Omega, \mathcal{A}) \in \tilde{\mathcal{O}}$ as L -valued measures and we recall that for each $a \in L$ there is an observable $E \in \tilde{\mathcal{O}}$ and a set $X \in \mathcal{A}$ such that $a = E(X)$.

An important technical assumption concerning the structure of $\mathbf{L}$ is the *separability of $\mathbf{L}$* ; this is the property that any pairwise orthogonal family (a_i) in $\mathbf{L}$ is at most countably infinite. This structure has the following measurement theoretical justification. The range $\mathbf{E}(\mathcal{A})$ of any observable $\mathbf{E} \in \tilde{\mathcal{O}}$ is a Boolean sub- σ -algebra of $\mathbf{L}$. If the value space $(\Omega, \mathcal{A})$ of $\mathbf{E} \in \tilde{\mathcal{O}}$ is a subspace of the real Borel space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, for some $n \in \mathbb{N}$, then the Boolean σ -algebra $\mathbf{E}(\mathcal{A})$ is separable. By the classic Loomis–Sikorski theorem, any separable Boolean sub- σ -algebra $\mathcal{B} \subset \mathbf{L}$ is the range of some (real valued) observable $\mathbf{E} : \mathcal{B}(\mathbb{R}) \rightarrow \mathbf{L}$ [10]. If the logic $\mathbf{L}$ is separable then any Boolean sub- σ -algebra of $\mathbf{L}$ is also separable and thus appears as the range of an observable. With this motivation we pose the following assumption:

Separability of the logic: Any orthogonal family (a_i) in $\mathbf{L}$ is at most countably infinite.

With slight abuse of language we occasionally call also the pair $(\mathbf{S}, \mathbf{L})$ the *logic* of the statistical duality $(\mathbf{S}, \mathcal{O}, p)$ associated with a subset $\tilde{\mathcal{O}}$ of observables satisfying the orthogonality postulate, Axiom 3. We also assume that $\mathbf{L}$ is separable.

Further Specifications

The set $\mathbf{S}$ of states is convex and it determines the order on $\mathbf{L}$. We now assume that pure states exist. Moreover, we assume that they are sufficiently many so that each $a \in \mathbf{L}$, $a \neq 0$, can be realised in some pure state, that is, there is an $\alpha \in \mathbf{P}$ such that $\alpha(a) = 1$.

Sufficiency of pure states: For any $a \in \mathbf{L}$, $a \neq 0$, there is an $\alpha \in \mathbf{P}$ such that $\alpha(a) = 1$.

We assume further that the logic $(\mathbf{S}, \mathbf{L})$ has the *Jauch–Piron property*,² which stipulates that if any two experimental propositions of $\mathbf{L}$ are (probabilistically, i.e. with probability one) true in a given state then there is also a stronger experimental proposition in $\mathbf{L}$ which is true in that state. Clearly, such a claim could not be expected to hold for arbitrary experimental propositions of $\mathbf{E}$.

Jauch–Piron property: For any $a, b \in \mathbf{L}$, if $\alpha(a) = \alpha(b) = 1$ for some $\alpha \in \mathbf{S}$, then there exists a $c \in \mathbf{L}$ such that $c \leq a$, $c \leq b$ and $\alpha(c) = 1$.

These two assumptions have strong structural implications. To state the relevant result we recall that an element $a \in \mathbf{L}$ is the *support* of the state $\alpha \in \mathbf{S}$ if $\alpha(a) = 1$ and for any $b \in \mathbf{L}$ the condition $\alpha(b) = 1$ implies $b \geq a$; that is, a , if it exists, is the smallest proposition which is true (in the sense of probabilistic certainty) in the state α . If the support of α exists, it is unique and we denote it by $s(\alpha)$.

Theorem 23.3 *If the set $\mathbf{S}$ of states of the logic $(\mathbf{S}, \mathbf{L})$ contains a sufficient set $\mathbf{P}$ of pure states and satisfies the Jauch–Piron property then $\mathbf{L}$ is a complete orthocomplemented orthomodular lattice. Each state $\alpha \in \mathbf{S}$ has its support $s(\alpha) \in \mathbf{L}$ and each $a \in \mathbf{L}$, $a \neq 0$, is a support of some state $\alpha \in \mathbf{S}$.*

²This property has been introduced independently in [36] and [37] and it is known to be equivalent to the requirement that each $\alpha \in \mathbf{S}$ has a (unique) support in $\mathbf{L}$ [3, Theorem 11.4.3].

Proof ³We show first that each $\alpha \in \mathbf{S}$ has a support in $\mathbf{L}$. If $\alpha(a) \neq 0$ for any $a \in \mathbf{L}$, then $\alpha(a) < 1$ for each $a \neq 1$, meaning that $s(\alpha) = 1$. If $\{a \in \mathbf{L} \mid \alpha(a) = 0\} \neq \{0\}$, we choose by Zorn's lemma a maximal orthogonal family in this set. By the separability of $\mathbf{L}$, this family is at most countably infinite. Hence there is a maximal orthogonal sequence $(a_i)_{i \geq 1}$ with $\alpha(a_i) = 0$ for all i . Let $a = \bigvee_i a_i$ and observe that $\alpha(a) = 0$. To establish that $a^\perp = s(\alpha)$, we show that for any $x \in \mathbf{L}$, $\alpha(x) = 0$ if and only if $x \perp a^\perp$, that is, $x \leq a$. If $x \leq a$, then $\alpha(x) \leq \alpha(a) = 0$. To show the converse, assume that $\alpha(x) = 0$. By the (dual) Jauch–Piron property there is a $c \in \mathbf{L}$ such that $x \leq c$, $a \leq c$ and $\alpha(c) = 0$. If $x \not\leq a$ then $c \neq a$ (since otherwise $a = c \geq x$) and thus, by orthomodularity $c = a \vee (c \wedge a^\perp)$. Therefore, $\alpha(c) = \alpha(a) + \alpha(c \wedge a^\perp)$ and thus $\alpha(c \wedge a^\perp) = 0$. Since $c \wedge a^\perp$ is orthogonal to each a_i we may expand the maximal orthogonal sequence (a_i) , which is a contradiction. Hence, $x \leq a$, showing that $a^\perp = s(\alpha)$.

We show next that each $a \in \mathbf{L}$, $a \neq 0$, is the support of some $\alpha \in \mathbf{S}$. Let $W(a) = \{x \in \mathbf{L} \mid x = s(\alpha) \text{ for some } \alpha \text{ such that } \alpha(a) = 1\}$. By the sufficiency of $\mathbf{P}$ this is a nonempty set. Moreover, if $x \in W(a)$, $x = s(\alpha)$ and $\alpha(a) = 1$, then $x \leq a$. Let (x_i) be a maximal (countable) orthogonal sequence in $W(a)$ and define $b = \bigvee_i x_i$ (so that $b \leq a$). As above, if $b \neq a$, then $a \wedge b^\perp$ would be an element in $W(a)$ pairwise orthogonal with each x_i , which is not possible. Thus $\bigvee_i x_i = a$. Since any x_i is the support of some α_i (for which $\alpha_i(a) = 1$), then a is the support of all the convex combinations $\sum w_i x_i$ (with all $w_i > 0$).

It remains to be shown that $\mathbf{L}$ is a complete lattice. Let $a, b \in \mathbf{L}$, $a \neq 0 \neq b$ (if a or b is 0 the supremum and infimum exist trivially). Let $\alpha, \beta \in \mathbf{S}$ be such that $s(\alpha) = a$, $s(\beta) = b$, and consider the state $\gamma = \lambda\alpha + (1 - \lambda)\beta$ for some $0 \neq \lambda \neq 1$. Clearly $s(\gamma) = a \vee b$. By De Morgan's laws one gets the dual result. It is well known that every separable orthomodular σ -orthocomplete lattice is complete, see, e.g., [8, Lemma 2.5.2f]. $\square$

There are three further important properties the logic $(\mathbf{S}, \mathbf{L})$ must possess in order to provide a geometric representation of the elements of $\mathbf{L}$ as subspaces of a vector space. The first property is the atomicity: $\mathbf{L}$ is *atomic* if every $a \in \mathbf{L}$, $a \neq 0$, contains an atom. We recall that an element $p \in \mathbf{L}$ is an atom if for any $a \in \mathbf{L}$, $a \neq 0$, the condition $a \leq p$ implies $a = p$. We let $\text{At}(\mathbf{L})$ denote the set of atoms in $\mathbf{L}$.

To incorporate and justify the atomicity of $\mathbf{L}$ we pose the following assumption concerning the identification of pure states. In Sect. 23.3 this assumption is formulated in terms of operations and it forms part of the projection postulate.

Identification of pure states: Let $\alpha \in \mathbf{P}$. For any $\beta \in \mathbf{S}$, if $\beta(s(\alpha)) = 1$ then $\beta = \alpha$.

Proposition 23.1 *With the assumptions of Theorem 23.3, the identification of pure states implies that the support of any pure state is an atom. Moreover, $\mathbf{L}$ is atomic and the map $\mathbf{P} \ni \alpha \mapsto s(\alpha) \in \text{At}(\mathbf{L})$ is a bijection.*

³Our proof is an adaption of the corresponding results in [3] Another source leading to this conclusion is provided by the results of Sect. 2.5.2 of [8].

Proof Let $p = s(\alpha)$ be the support of $\alpha \in \mathbf{P}$ and let $a \in \mathbf{L}$, $a \neq 0$, be such that $a \leq p$. Since $a \neq 0$, there is a $\beta \in \mathbf{P}$ such that $\beta(a) = 1$. From $a \leq p$ one then gets $\beta(p) = 1$, which means that $\beta = \alpha$. Since $a \geq s(\beta)$ and $s(\beta) = s(\alpha) = p$, one has $a \geq p$ and thus $a = p$, that is, p is an atom.

For $a \in \mathbf{L}$, $a \neq 0$, there is a $\alpha \in \mathbf{P}$ such that $\alpha(a) = 1$. Therefore $s(\alpha) \leq a$, showing that $\mathbf{L}$ is atomic.

We leave it as an exercise to show that the mapping $\mathbf{P} \ni \alpha \rightarrow s(\alpha) \in \text{At}(\mathbf{L})$ is injective and surjective. $\square$

The second ingredient required to establish the vector space realisation is the *covering property*: for any $a \in \mathbf{L}$ and $p \in \text{At}(\mathbf{L})$, if $a \wedge p = 0$, then $a \vee p$ covers a (that is, for any $b \in \mathbf{L}$, if $a \leq b \leq a \vee p$, then $b = a$ or $b = a \vee p$). Since $\mathbf{L}$ is an atomic lattice, the covering property can be formalised equivalently as follows: for any $a \in \mathbf{L}$, $p \in \text{At}(\mathbf{L})$, the element $(a \vee p) \wedge a^\perp$ is either an atom or 0 [8, Prop. 3.2.17]. To obtain the covering property for $(\mathbf{S}, \mathbf{L})$ we stipulate the possibility of actualising potential properties with *minimal disturbance*. This feature is part of the projection postulate, which we present in Sect. 23.3 together with an elucidation of its physical motivation. The ideality assumption (I1) of a filter given there corresponds to the following minimal disturbance requirement.

Minimal disturbance: If $\alpha \in \mathbf{P}$, $a \in \mathbf{L}$, and $\alpha(a) \neq 0$, then there exists a pure state $\beta \in \mathbf{P}$ such that $\beta(a) = 1$, that is, $s(\beta) \leq a$, and $\alpha(s(\beta)) = \alpha(a)$.

Proposition 23.2 ([38]) *With the assumptions of Theorem 23.3 and the identification of the pure states, the minimal disturbance implies the covering property.*

Proof Let $p \in \text{At}(\mathbf{L})$, $a \in \mathbf{L}$ and $p = s(\alpha)$. Let α_1 and α_2 be the pure states such that $\alpha(a) = \alpha(s(\alpha_1))$ and $\alpha(a^\perp) = \alpha(s(\alpha_2))$ as given by the minimal disturbance. Clearly, $\alpha(s(\alpha_1) \vee s(\alpha_2)) = 1$, so that $p \leq s(\alpha_1) \vee s(\alpha_2)$ and $p \vee a^\perp \leq s(\alpha_1) \vee a^\perp$. Hence $(p \vee a^\perp) \wedge a \leq (s(\alpha_1) \vee a^\perp) \wedge a = s(\alpha_1)$, which means that $(p \vee a^\perp) \wedge a$ is either the atom $s(\alpha_1)$ or 0, that is, the covering property holds in $\mathbf{L}$. $\square$

We collect the above assumptions concerning the set of states of the logic in the form of an axiom.

Axiom 4 The set $\mathbf{S}$ of states of the logic $(\mathbf{S}, \mathbf{L})$, with a separable $\mathbf{L}$, has a sufficient set of pure states, the Jauch–Piron property, and it allows the identification of pure states and obeys the minimal disturbance.

An atomic lattice with the covering property is often referred to as an AC lattice. We may thus conclude that Axioms 3 and 4 imply that $\mathbf{L}$ is a complete orthomodular AC lattice and that the support function gives a bijective correspondence between the sets $\mathbf{P}$ and $\text{At}(\mathbf{L})$.

The final ingredient required to establish a geometric representation of $(\mathbf{S}, \mathbf{L})$ is the irreducibility of $\mathbf{L}$.

The Classical Case Excluded

There are various features of quantum mechanics that have been elevated to the status of fundamental principles of the theory. These include the notions of superposition, complementarity, uncertainty, entanglement, nonunique decomposability and purification of mixed states, and irreducibility of probabilities, which are unquestionably among the most widely discussed characteristic traits of quantum mechanics. In Remark 23.2 the unique decomposability of mixed states into its pure components is seen to be closely related to a classical description and it goes hand in hand with the Boolean structure of the set of decision effects. Here we discuss briefly the notions of superposition and complementarity to show that for a proper quantum system the logic L is far from being Boolean.

Superpositions

There are several formulations of the notion of superposition in quantum logic. We adopt the following definition taken from [10, p. 53] as a formalisation of the intuitive ideas of Dirac [39].

Definition 23.3 A pure state $\alpha \in P$ is a *superposition* of pure states $\alpha_1, \alpha_2 \in P$ if and only if $\alpha_1(a) = \alpha_2(a) = 1$ implies $\alpha(a) = 1$ for every $a \in L$. Equivalently, a pure state α is a superposition of pure states α_1 and α_2 if and only if $s(\alpha) \leq s(\alpha_1) \vee s(\alpha_2)$.

Instead of stating directly a superposition principle we give the following definition:

A physical system with the structure (S, L) is a *proper quantum system* if for every two pure states $\alpha, \beta \in P, \alpha \neq \beta$, there exists a third one $\gamma \in P, \alpha \neq \gamma \neq \beta$, which is their superposition.

It is then a simple but important consequence that the logic of a proper quantum system is *irreducible*, that is, the *centre* of L ,

$$\text{Cent}(L) = \{c \in L \mid a = (a \wedge c) \vee (a \wedge c^\perp) \text{ for any } a \in L\},$$

contains only the trivial elements 0 and e . (For a proof, see e.g. [8, Corollary 3.2.4].)

Remark 23.4 Note that if $c \in \text{Cent}(L)$, then for any $a \in L$, there is an observable $E \in \tilde{O}$ such that $c = E(X)$ and $a = E(Y)$ for some value sets X and Y . Thus for a proper quantum system there is no (nontrivial) proposition (or property) that could be measured together with every other proposition (or property). By contrast, if $\text{Cent}(L) = L$, the set L forms a Boolean σ -algebra for which the theorems of Stone [40] and Loomis [41] and Sikorski [42] give a representation as a σ -algebra $\mathcal{A}$ of subsets of a set Ω . In this case all properties can be measured jointly. $\triangleleft$

Remark 23.5 If a pure state α is a superposition of pure states β and γ , then β is also a superposition of α and γ , and likewise γ is a superposition of α and β . This is the exchange property and it is often included in the notion of superposition of states. In the present context this property is equivalent to the covering property, stated above as part of the projection postulate. For a proof, see, e.g., [8, Prop. 3.2.17]. $\triangleleft$

Complementarity

The existence of pairs of complementary observables is another fundamental feature of quantum mechanics. Following the ideas of Bohr [43], we say that two observables are *complementary* if all the experimental arrangements which unambiguously define these observables are mutually exclusive. Again, there are various ways of formalising this intuitive idea. We adopt the following definition appropriate to the logic (S, L) .

Definition 23.4 Properties $a, b \in L$ are *complementary* if they are disjoint, that is, $a \wedge b = 0$, but not orthogonal, that is, $a \not\leq b^\perp$. Equivalently, $a, b \in L$ are *complementary* if for any $\alpha \in S$, the condition $\alpha(a) = 1$ implies $0 \neq \alpha(b) \neq 1$, and the condition $\alpha(b) = 1$ implies $0 \neq \alpha(a) \neq 1$.

In a Boolean logic, the conditions $a \wedge b = 0$ and $a \leq b^\perp$ are equivalent. This means that if there are complementary properties in L then L cannot be Boolean.

It is a simple exercise to show that L is irreducible if for any $a \in L, a \neq 0, e$, there is a $b \in L$ such that a and b are complementary. As an alternative to the previous definition, we could call a physical system with the structure (S, L) a *proper quantum system* if for any $a \in L, a \neq 0, e$, there is a $b \in L$ such that a and b are complementary. It then follows that for a proper quantum system the logic L is irreducible.

We summarise the main result of this section.

Corollary 23.1 *Let (S, O, p) be a statistical duality satisfying Axioms 1 and 2. If the logic (S, L) defined by a (nonempty) subset $\tilde{O}$ of observables satisfies Axioms 3 and 4 and the physical system in question is a proper quantum system, then L is an irreducible complete orthomodular AC lattice and there is a bijective correspondence (given by the support function) between the set P of pure states in S and the set $\text{At}(L)$ of atoms in L .*

23.3 Filters and the Projection Postulate

We assume now that the statistical duality satisfies the stronger Axioms 1' and 2'. Any subset of observables $\tilde{O} \subset O$ defines a subset $\tilde{L}$ of effects, and any effect is an isotony class of operations. Thus, for any $\tilde{O} \subset O$ there is also a subset $\tilde{O}$ of operations such that the resulting subset of effects $e \circ \tilde{O} = \tilde{L}$. This suggests that by singling out a sufficiently large and regular set of operations the fundamental result of Corollary 23.1 could be obtained also for $\tilde{L}$, without using Axioms 3 and 4. This is what we consider next.

Remark 23.6 With the structure specified by Theorem 23.3 and Propositions 23.1 and 23.2 one may construct for each $a \in L$ a map $\phi_a : S \rightarrow S$ whose restriction to pure states is uniquely defined by $s(\phi_a(\alpha)) = (s(\alpha) \vee a^\perp) \wedge a, \alpha \in P$, and which has the typical properties of a state transformation caused by an ideal first kind

measurement [44–46]. Due to the properties of the support function $s : S \rightarrow L$, the map ϕ_a fails to be linear, that is, it is not an operation. Apart from this, the characteristic properties of such a ϕ_a serve below as the defining properties of filters. $\triangleleft$

Filters

Filters are a special kind of operations reflecting certain ideality properties the yes-no (or *simple*) measurements may or may not possess. The properties of filters have been discussed extensively in the literature, see, e.g. [3, 13, 14, 16, 17, 20, 44–47]. This allows us to be brief in their introduction. Filters are defined through their action on pure states. The definition thus presumes that the set $P = \text{ex}(S)$ of pure states is not empty.

An operation $\phi \in \mathcal{O}$ is *pure* if

(P1) $\phi(\alpha) \in [0, 1] \cdot P$ for any pure state $\alpha \in P$,

and an operation ϕ is of the *first kind* if

(F1) $e(\phi(\alpha)) = 1$ implies $\phi(\alpha) = \alpha$ for any $\alpha \in P$,

(F2) $e(\phi^2(\alpha)) = e(\phi(\alpha))$ for any $\alpha \in P$.

To define the ideality of an operation we first assume that any pure state can be identified by an operation:

(S1) For any pure state $\alpha \in P$ there is a unique $\phi_\alpha \in \mathcal{O}$ such that $e(\phi_\alpha(\beta)) = 1$ implies $\beta = \alpha$ for any $\beta \in P$.

We then say that a pure operation ϕ is *ideal* if

(I1) $e(\phi(\alpha)) = e(\phi_{\alpha'}(\alpha))$ for any $\alpha \in P$ for which $\phi(\alpha) \neq 0$, where $\alpha' = e(\phi(\alpha))^{-1} \phi(\alpha)$ and $\phi_{\alpha'}$ as in (S1).

A pure, ideal, first kind operation is a *filter*. The set of filters will be denoted by $\mathcal{O}_f$. We comment briefly on the defining properties of filters.

The purity (P1) of an operation means that it takes a pure state to a pure state with a possible loss in the intensity. As pure states may be interpreted as maximal information states, a pure operation leaves the system in a maximal information state whenever it was in such a state.

With the ideality assumptions formalised by (P1), (F1), (F2) and (I1), one tries to characterise a kind of nontrivial operation which exerts a minimal influence on the state. According to (I1), an ideal ϕ maps any pure state α onto an eigenstate of ϕ closest to α , thus disturbing the system to a minimal extent. This is the minimal disturbance assumption discussed in Sect. 23.2.

Of the two first kind conditions (F1) and (F2), the first claims that if ϕ does not lead to a detectable effect when performed on the system in a pure state α then, provided that the operation is good enough, it does not alter the state of the system either. According to (F2), a repeated application of a good operation does not lead to a new effect.

As an immediate consequence of the defining properties of filters, we note that they are not only *weakly repeatable*, that is, $e(\phi^2(\alpha)) = e(\phi(\alpha))$ for any $\alpha \in \mathbf{P}$, but also *repeatable*, in the sense that $\phi^2(\alpha) = \phi(\alpha)$ for any $\alpha \in \mathbf{P}$. Moreover, filters satisfy the most common ideality requirement: if a good operation ϕ_1 is performed on the system in a pure state α which is an eigenstate of a good operation ϕ_2 (i.e. $e(\phi_2(\alpha)) = e(\alpha)$), and if ϕ_2 which commutes weakly with ϕ_1 (i.e. $\phi_1 \circ \phi_2$ and $\phi_2 \circ \phi_1$ are isotonic), then ϕ_1 leaves the system in a state which is still an eigenstate of ϕ_2 .

We say that the set $\mathcal{O}_f$ of filters is *sufficiently rich* if the operations that exist according to (S1) are filters and if, in addition, it fulfills the following condition:

(S2) For each filter $\phi \in \mathcal{O}_f$ there is another filter $\phi' \in \mathcal{O}_f$ such that $e \circ \phi' = (e \circ \phi)^\perp$.

Condition (S1), the identification of pure states, expresses the idea that any pure state α can be produced by a particular selection or filtering process ϕ_α , which under the conditions (F1) and (F2) assumes the form $\phi_\alpha(\beta) = e(\phi_\alpha(\beta))\alpha$ for any $\beta \in \mathbf{P}$. The second sufficiency condition (S2) stipulates that if an effect a can be obtained from a filter, that is $a = e \circ \phi$ for some $\phi \in \mathcal{O}_f$, then also its ‘negation’ $a^\perp = e - a$ can be produced by a pure ideal first kind operation.

Projection Postulate

The set $\tilde{\mathbf{L}}$ of *propositions* of the convex scheme $(S, \mathcal{O}, E)$ is now defined as the set of all decision effects $a \in \text{ex}(E)$ with nonempty certainly-yes-domain $a^1 = \{\alpha \in \mathbf{P} \mid a(\alpha) = 1\}$ together with the null effect 0,

$$\tilde{\mathbf{L}} = \{a \in \text{ex}(E) \mid a = 0 \text{ or } a^1 \neq \emptyset\}.$$

For a given system $(S, \mathcal{O}, E)$ the set $\mathcal{O}_f$ of filters may be empty and the set $\tilde{\mathbf{L}}$ of propositions may be trivial, $\tilde{\mathbf{L}} = \{0, e\}$. However, for any $\phi \in \mathcal{O}_f$, $\phi \neq 0$, the resulting effect $e \circ \phi$ has a nonempty certainly-yes-domain $(e \circ \phi)^1$. By Remark 23.6 it is also natural to expect that for any $a \in \tilde{\mathbf{L}}$, $a \neq 0$, there is a filter ϕ_a such that $e \circ \phi_a = a$. With the projection postulate we confirm this expectation, whilst adding the assumption of uniqueness. We now make a further assumption: the set of pure states is not only nonempty but is also *strongly ordering* on $\tilde{\mathbf{L}}$, that is, it is ordering and such that for any $f, g \in \tilde{\mathbf{L}}$, if $f^1 \neq \emptyset$, and $f^1 \subset g^1$, then $f \leq g$. The following axiom specifies further a statistical duality $(S, \mathcal{O}, p)$ that satisfies Axioms 1' and 2'.

Axiom 5 (*Projection Postulate*) The statistical duality $(S, \mathcal{O}, p)$ possesses the following properties: its set $\mathbf{P}$ of pure states is strongly ordering on $\tilde{\mathbf{L}}$, the set of filters $\mathcal{O}_f \subset \mathcal{O}$ is sufficiently rich, and there is a bijective mapping $\Phi : \tilde{\mathbf{L}} \rightarrow \mathcal{O}_f$ with the property that for every $a \in \tilde{\mathbf{L}}$ and $\alpha \in \mathbf{P}$, $a(\alpha) = e(\Phi(a)(\alpha))$.

The projection postulate captures features of a subset of measurements known to be realised to a good degree of accuracy in both classical and quantum physical theories.

Remark 23.7 It can be argued that these features, which thus are intrinsic to the fabric of the physical world as we know it, constitute necessary preconditions for a proposition system to refer to objects. A justification of the projection postulate along these lines has been given in the work of Stachow [48]. Furthermore, it was shown there that the atomicity and covering property of a logic are necessary conditions of a language of individual, isolated objects. It is an interesting but as yet unsolved problem to investigate whether these arguments, which were formalised in the framework of the so-called dialogical approach to quantum logic by Mittelstaedt and Stachow (e.g., [49]), can be carried over into the present approach. $\triangleleft$

The projection postulate guarantees the existence of a sufficiently rich collection of operations associated with the important class of pure, ideal, first-kind measurements, but it does not restrict the theory to such measurements only. Neither does it distinguish between classical and quantum descriptions. In any case, this postulate has strong structural implications on the order structure of the set $\tilde{\mathcal{L}}$ of propositions. They will be studied next.

Lemma 23.1 *For a statistical duality $(S, \mathcal{O}, p)$ satisfying the projection postulate, the set $\tilde{\mathcal{L}}$ of propositions is a nonempty partially ordered set with $a \mapsto a^\perp$ as orthocomplementation.*

Proof Since $\mathcal{P} \neq \emptyset$, the set $\mathcal{O}_f$ of filters is nonempty and thus also $\tilde{\mathcal{L}} \neq \emptyset$. The set $\text{ex}(\mathcal{E})$ of decision effects is closed under the map $a \mapsto a^\perp = e - a$. If $a \in \tilde{\mathcal{L}}$, with $a = e \circ \Phi(a)$, $\Phi(a) \in \mathcal{O}_f$, then by (S2) $a^\perp = e \circ \Phi(a)'$ for some $\Phi(a)' \in \mathcal{O}_f$, so that $a^\perp \in \tilde{\mathcal{L}}$. Clearly, $\Phi(a)' = \Phi(a^\perp)$. Let $b \in \tilde{\mathcal{L}}$ be such that $b \leq a$ and $b \leq a^\perp$, and assume that $b \neq 0$. Then for any $\alpha \in b^\perp$, $a(\alpha) = 1$ and $a^\perp(\alpha) = 1$ which is impossible. Thus $b = 0$, that is, $a \wedge_{\tilde{\mathcal{L}}} a^\perp = 0$. The remaining claims are obvious. $\square$

Lemma 23.2 *For any $a, b \in \tilde{\mathcal{L}}$, if $a \leq b^\perp$, then $a + b \in \tilde{\mathcal{L}}$. Similarly, for any triple (a, b, c) of mutually orthogonal elements of $\tilde{\mathcal{L}}$, $a + b$, $a + c$, $b + c$, $a + b + c \in \tilde{\mathcal{L}}$.*

Proof For $a \perp b$, $a + b \leq e$. Now $(a + b)^\perp \supset a^\perp \cup b^\perp \neq \emptyset$. Assume that $a + b = \frac{1}{2}(f + g)$ for some $f, g \in \mathcal{E}$. Then $(a + b)^\perp = f^\perp \cap g^\perp$ so that $a + b \leq f$ and $a + b \leq g$. Thus $a + b = a + b + \frac{1}{2}(f - (a + b) + g - (a + b))$, which implies that $f = g = a + b$, that is, $a + b \in \text{ex}(\mathcal{E})$. Hence $a + b \in \tilde{\mathcal{L}}$. Copying the argument for a triple of mutually orthogonal elements $a, b, c \in \tilde{\mathcal{L}}$ one immediately concludes also that $a + b + c \in \tilde{\mathcal{L}}$. $\square$

Corollary 23.2 *For a statistical duality $(S, \mathcal{O}, p)$ satisfying the projection postulate, the set $\tilde{\mathcal{L}}$ of propositions is orthomodular.*

Proof Let $a, b, c \in \tilde{\mathcal{L}}$ be a triple of mutually orthogonal elements. Then not only $a + b$, $a + c$ and $b + c$ but also $a + b + c \in \tilde{\mathcal{L}}$. This means that $\tilde{\mathcal{L}}$ is triangle-closed in the sense of [50]. By [50, Theorem 3.2] this is equivalent to $\tilde{\mathcal{L}} \subset \mathcal{E}$ being orthomodular, and, in particular, $a + b = a \vee_{\tilde{\mathcal{L}}} b$ for $a, b \in \tilde{\mathcal{L}}$, $a \leq b^\perp$. $\square$

Lemma 23.3 *For any $\alpha \in \mathbf{P}$, $e \circ \phi_\alpha \in \text{At}(\tilde{\mathbf{L}})$. Moreover, $\tilde{\mathbf{L}}$ is atomic, that is, any $a \in \tilde{\mathbf{L}}$, $a \neq 0$, contains an atom. The map $\mathbf{P} \ni \alpha \mapsto e \circ \phi_\alpha \in \text{At}(\tilde{\mathbf{L}})$ is a bijection, with $e \circ \phi_\alpha$ being the support of α .*

Proof Let $a \in \tilde{\mathbf{L}}$, $\alpha \in \mathbf{P}$, and assume that $a \leq \Phi^{-1}(\phi_\alpha)$. If $a \neq 0$ then for any $\beta \in a^1$, $\Phi^{-1}(\phi_\alpha)(\beta) = e(\phi_\alpha(\beta)) = 1$, so that by (S1) $\beta = \alpha$, that is $a^1 = \{\alpha\}$. Therefore $\Phi(a) = \phi_\alpha$, or equivalently, $a = \Phi^{-1}(\phi_\alpha)$, which entails that for any $\alpha \in \mathbf{P}$, $\Phi^{-1}(\phi_\alpha)$ is an atom. Clearly, for any $a \in \tilde{\mathbf{L}}$, one has $\Phi^{-1}(\phi_\alpha) = e \circ \phi_\alpha \leq a$ for all $\alpha \in a^1$. $\square$

Lemma 23.4 *The set $\mathbf{P}$ of pure states is sufficient for $\tilde{\mathbf{L}}$, and $\tilde{\mathbf{L}}$ has the Jauch–Piron property.*

Proof The sufficiency of $\mathbf{P}$ for $\tilde{\mathbf{L}}$ is obvious. Let $a, b \in \tilde{\mathbf{L}}$ be such that $a^1 \cap b^1 \neq \emptyset$. For any $\alpha \in a^1 \cap b^1$, $e \circ \phi_\alpha$ is contained both in a and in b and $e(\phi_\alpha \alpha) = 1$. $\square$

We observe that the range $\mathbf{E}(\mathcal{A})$ of an observable $\mathbf{E} \in \mathbf{O}$ is Boolean if it is contained in $\tilde{\mathbf{L}}$. Therefore we may again justify the separability assumption of $\tilde{\mathbf{L}}$ with the requirement that any Boolean subsystem of $\tilde{\mathbf{L}}$ could be realised as the range of an observable. With the separability assumption, $\tilde{\mathbf{L}}$ thus acquires the structure specified in Theorem 23.3. By virtue of Proposition 23.2, the ideality property (I1) of filters then gives the covering property. Hence we have the following.

Theorem 23.4 *If the operational description $(\mathbf{S}, \mathbf{O}, \mathbf{E})$ defined by Axioms 1' and 2' satisfies the projection postulate and $\tilde{\mathbf{L}} = \{a \in \text{ex}(\mathbf{E}) \mid a^1 \neq \emptyset \text{ or } a = 0\}$ is separable, then $\tilde{\mathbf{L}}$ forms a complete atomic orthomodular orthocomplemented lattice with the covering property. Moreover, the support function gives a bijective correspondence between the pure states in $\mathbf{S}$ and the atoms of $\tilde{\mathbf{L}}$.*

To obtain the restriction to an irreducible logic $\tilde{\mathbf{L}}$, it is most straightforward to require that any two pure states can be superposed into a new pure state. With the structures given by the projection postulate, we may immediately adopt Definition 23.3 to conclude that for a proper quantum system, the structures of Corollary 23.1 are again available.

It is to be emphasised, however, that even though the two sets of Axioms [1, 2, 3, 4] and [1', 2', 5] lead to the common structure of Corollary 23.1, the first approach starts with the weaker assumptions concerning the pair $(\mathbf{S}, \mathbf{E})$. It is therefore conceivable that there are pairs $(\mathbf{S}, \mathbf{L})$ with the structure of this corollary appearing as models for the first approach but not for the second approach.

23.4 Hilbert Space Coordinatisation

The Basic Problem

In the Hilbert space formulation of quantum mechanics the pair $(\mathbf{S}, \mathbf{E})$ is identified as the set $\mathbf{S}$ of positive trace-1 operators and effect operators on a complex separable Hilbert space together with the set $\mathbf{L}$ of (orthogonal) projections on it. Given the latter

set, a deep theorem due to Gleason [51] ensures that all the probability measures on $\mathbf{L}$ arise from the states through the familiar trace formula.⁴

Until now we have presented two sets of axioms for the structures $(\mathbf{S}, \mathbf{E})$ and $(\mathbf{S}, \mathbf{L})$ common to all probabilistic physical theories. It remains to show that the only realisations of these abstract schemes are those given by the Hilbert space description of a proper quantum system described above, possibly with the help of further assumptions. We present an outline of the solution to this problem, including a discussion of some critical points that are still open.

One traditional way of approaching this task—to be followed also here—has been to isolate first the proposition system $\mathbf{L}$ and to determine the models of this structure alone. Once this is achieved, the structure of $\mathbf{S}$ is added and $\mathbf{E}$ is determined. One might expect that this way of voluntarily neglecting a good part of the basic structures of the pairs $(\mathbf{S}, \mathbf{E})$ and $(\mathbf{S}, \mathbf{L})$ cannot be the most optimal approach. We return to this question later.

The Fundamental Representation Theorem

Let K be a division ring with an involutive antiautomorphism $\lambda \mapsto \lambda^*$ (such that $(\lambda + \mu)^* = \lambda^* + \mu^*$, $(\lambda\mu)^* = \mu^*\lambda^*$, $\lambda^{**} = \lambda$), and let V be a (left) vector space over K . A *Hermitian form* on V is a mapping $f : V \times V \rightarrow K$ with the following properties: for any $u, v, w \in V$ and $\lambda, \mu \in K$,

$$\begin{aligned} f(\lambda u + \mu v, w) &= \lambda f(u, w) + \mu f(v, w) \\ f(u, v)^* &= f(v, u) \\ f(v, v) &= 0 \text{ implies } v = 0. \end{aligned}$$

If V admits a Hermitian form f we say that V , or rather $(V, K, *, f)$, is a *Hermitian space*. A subspace $M \subset V$ of a Hermitian space is *f-closed* if $M = M^{\perp\perp}$, where

$$M^{\perp} = \{v \in V \mid f(v, x) = 0 \text{ for all } x \in M\}.$$

Let $L_f(V)$ denote the set of all f -closed subspaces of V . In addition to the trivial subspaces $\{0\}$ and V , any finite-dimensional subspace is f -closed. If V is infinite-dimensional, these subspaces do not exhaust the set $L_f(V)$. The subset inclusion $\subset$ together with the map $M \mapsto M^{\perp}$ gives $L_f(V)$ the structure of an irreducible complete orthocomplemented AC lattice. The converse result is a fundamental representation theorem of projective geometry, proved in detail, for instance, in [52, Theorem 34.5].

Theorem 23.5 *If $\mathbf{L}$ is an irreducible complete orthocomplemented AC lattice of length at least 4 (that is, the length of a maximal chain is ≥ 4), then there is a Hermitian space $(V, K, *, f)$ such that $\mathbf{L}$ is ortho-isomorphic to the lattice $L_f(V)$.*

⁴For a detailed discussion of this theorem, see, e.g. [10].

A Hermitian space $(V, K, *, f)$ is *orthomodular* if for any $M \in \mathbf{L}_f(V)$,

$$M + M^\perp = V.$$

A Hermitian space $(V, K, *, f)$ is known to be orthomodular if and only if the lattice $\mathbf{L}_f(V)$ is orthomodular, see e.g. [53, Theorem 2.8]. Thus we have the following corollary:

Corollary 23.3 *Assume that $\mathbf{L}$ is an irreducible complete orthocomplemented orthomodular AC lattice of length at least 4. Then there is an orthomodular space $(V, K, *, f)$ such that $\mathbf{L}$ is ortho-isomorphic to $\mathbf{L}_f(V)$, in short, $\mathbf{L} \simeq \mathbf{L}_f(V)$. In particular, all the finite-dimensional subspaces of V are in $\mathbf{L}_f(V)$ and the atoms of $\mathbf{L}_f(V)$ are the one-dimensional subspaces of V .*

The pure states $\alpha \in \mathbf{P}$ are in one-to-one onto correspondence with the atoms $[v] = \{\lambda v \mid \lambda \in K\} \in \mathbf{L}_f(V)$ and they are uniquely determined by their values on the atoms, that is, by the numbers $\alpha_{[v]}([u]) \in [0, 1]$, $[u] \in \text{At}(\mathbf{L}_f(V))$. It is to be stressed that this corollary does not yet give any information on the structure of the numbers $\alpha_{[v]}([u])$; in particular, it is not known if $\alpha_{[v]}([u])$ could be related to the K -number $f(u', v')$ for some $v' \in [v]$, $u' \in [u]$. If such a conclusion could be reached then K should be an extension of $\mathbb{R}$.

The well-known models for an orthomodular space $(V, K, *, f)$ are the Hilbert spaces $\mathcal{H}$ over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, the quaternions. In these models, the form f is the scalar product on V , and by Gleason's theorem, the probabilities $\alpha_{[v]}([u])$ are of the form $\alpha_{[v]}([u]) = |f(v', u')|^2$ for any $v' \in [v]$, $u' \in [u]$ with $f(v', v') = f(u', u') = 1$, provided that $\dim(\mathcal{H}) \geq 3$ (which is the case in Corollary 23.3). However, the Hilbert spaces do not exhaust the orthomodular spaces. In the finite dimensional case this is evident, as shown by a simple example.

Example 23.1 Any finite-dimensional Hermitian space $(V, K, *, f)$ is orthomodular and each subspace M of V is f -closed, see, e.g. [54]. Hence the lattice $\mathbf{L}_f(V)$ of f -closed subspaces coincides with the lattice $\mathbf{L}(V)$ of all subspaces of V , which is modular (and thus also orthomodular). It is obvious that the space $(V, K, *, f)$ need not be a Hilbert space. To see it, consider the finite-dimensional rational vector space $\mathbb{Q}^n$ with the natural form $f(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^n q_i p_i$. The form f is Hermitian so that $(\mathbb{Q}^n, \mathbb{Q}, \text{id}, f)$ is an orthomodular space. Clearly $\mathbb{Q}^n$ is not complete with respect to the distance defined by f . We return to this example in Sect. 23.4 where we study probability measures on $\mathbf{L}_f(\mathbb{Q}^n)$. $\triangleleft$

This example leaves open the infinite-dimensional case. In [55] Keller was able to construct an explicit example of an infinite-dimensional orthomodular space that is very far from being a Hilbert space. Further examples emerged later [56], and we now know that there are plenty of orthomodular spaces other than the Hilbert spaces. The problem then arises to characterise the Hilbert spaces among the orthomodular spaces. This is solved in the next subsection.

A Theorem of Solér

The following result characterises the Hilbert spaces among the orthomodular spaces.

Theorem 23.6 ([57]) *Let $(V, K, *, f)$ be an orthomodular space. The division ring K is either $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ and $(V, K, *, f)$ is the corresponding Hilbert space⁵ if and only if there is an infinite sequence of nonzero vectors e_i , $i \in \mathbb{N}$, such that $f(e_i, e_j) = 0$ for all $i \neq j$, with the property $f(e_i, e_i) = f(e_j, e_j)$ for all $i, j \in \mathbb{N}$.*

We emphasise that in this theorem V is required to be infinite-dimensional and orthomodular. The next two examples demonstrate that neither of these assumptions can be relaxed.

Example 23.2 The vectors $(1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ form an orthonormal basis in $(\mathbb{Q}^n, \mathbb{Q}, \text{id}, f)$ but the space, though orthomodular, is not a Hilbert space. $\triangleleft$

Example 23.3 Consider the infinite-dimensional vector space $V = \ell_2(\mathbb{Q})$ of the square summable sequences of rational numbers $\mathbf{q} = (q_1, q_2, q_3, \dots)$ with the Hermitian form $f(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{\infty} q_i p_i$. The lattice $\mathbf{L}_f(V)$ of f -closed subspaces is a complete, irreducible AC lattice of infinite length, but it is not orthomodular. The vectors $(1, 0, \dots, 0, \dots), \dots, (0, \dots, 0, 1, 0, \dots), \dots$ form an orthonormal basis in V , but V is not a Hilbert space. $\triangleleft$

For $\mathbf{L} \simeq \mathbf{L}_f(V)$ the existence of a sequence of mutually orthogonal vectors (e_i) in V follows from the assumption that $\mathbf{L}$ contains an infinite sequence of pairwise orthogonal atoms. Such an assumption is physically well motivated e.g. by the spectroscopic data or by the assumption that the quantum system can be localised in an Euclidean space. It is then worth stressing that, contrary to our intuition that comes from using complex numbers, it is the ‘norm’ requirement $f(e_i, e_i) = f(e_j, e_j)$ that is here highly non-trivial. Indeed, suppose that $f(e_i, e_i) = \lambda$ and $f(e_i, e_j) = 0$. We have to find an element $\mu \in K$ such that $\mu f(e_j, e_j) \mu^* = \lambda$; in this way $f(\mu e_j, \mu e_j) = f(e_i, e_i)$, see Sect. 23.4. This is a quadratic equation in K that cannot be solved in general. In $\mathbb{R}$ or $\mathbb{C}$ one would simply take the square root of the positive number $\lambda \lambda^*$ whereas this would not work in $\mathbb{Q}$, say. For instance, a one-dimensional subspace $[\mathbf{q}] = \{\lambda \mathbf{q} \mid \lambda \in \mathbb{Q}\}$ of $(\mathbb{Q}^n, \mathbb{Q}, \text{id}, f)$ contains a unit vector only if $\sqrt{\sum q_i^2}$ is rational.

Combining Corollary 23.3 with the theorem of Solér we get the following.

Theorem 23.7 *Assume that $\mathbf{L}$ is an irreducible complete orthocomplemented orthomodular AC lattice that contains an infinite sequence of orthogonal atoms. Then there is an orthomodular space $(V, K, *, f)$ such that $\mathbf{L}$ is orthoisomorphic to $\mathbf{L}_f(V)$, K is $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ and $(V, K, *, f)$ is the corresponding Hilbert space if and only if V contains an infinite sequence of mutually orthogonal vectors (v_i) with the property*

$$f(v_i, v_i) = f(v_j, v_j) \quad \text{for all } i, j \in \mathbb{N}. \quad (23.7)$$

⁵If $K = \mathbb{R}$ then $*$ is the identity. For $K = \mathbb{C}$ the map $*$ cannot be the identity and if it is continuous then it is the complex conjugation. For $K = \mathbb{H}$ the map is the quaternionic conjugation.

By assumption, there is an infinite sequence of orthogonal vectors. The essential question is which properties of (S, L) would imply that such a sequence could be chosen to have the ‘norm’ property (23.7).

Purely lattice theoretical conditions on L are known that are sufficient to ensure that $L \simeq L(\mathcal{H})$ for a Hilbert space $\mathcal{H}$. We can refer to the so-called *angle bisection property* [58] or the existence of *harmonic conjugate* pairs of atoms [54, 59]. They are of a geometric nature and, in the light of the present understanding, they seem to lack any physical interpretation. Therefore, they are not useful for the axiomatic scheme followed here.

The necessary and sufficient conditions for the conclusion $L \simeq L(\mathcal{H})$ of Theorem 23.7 are expressed in terms of (V, f) . One might expect that the assumptions of this theorem together with the full structure of the pair (S, L) , in particular, the bijection between P and $\text{At}(L)$, could already force L to be a Hilbertian lattice. We investigate some aspects of this question in Sect. 23.5, although to the best of our knowledge, this problem remains still largely open.

Two questions remain: what can be said if L is only of finite length, and how can the states be represented once we have represented L ?

As concerns the latter question we recall that if $(V, K, *, f)$ is a Hilbert space of dimension at least 3, then all the probability measures on $L = L_f(V)$ are described by Gleason’s theorem. According to this theorem, for any probability measure α on $L_f(V)$, there is a unique positive trace one operator $\varrho : V \rightarrow V$ such that, for any $M \in L_f(V)$, we have $\alpha(M) = \text{tr}[\varrho P_M]$, where P_M is the projection onto M . For $\dim(V) = 2$, the set of all probability measures on $L_f(V)$ is, however, much bigger than those defined by the states. But these additional probability measures are not supported by $L_f(V)$; for details, see [3, Section 25.2]. To the best of our knowledge, there is no empirical evidence which would require the use of such probability measures as states of a two-level quantum system.

The situation is very different when $(V, K, *, f)$ is not a Hilbert space. Very little is known of the probability measures on the lattices $L_f(V)$. Keller [60] gives examples of $L_f(V)$ for which one may construct a rich supply of probability measures $m : L_f(V) \rightarrow [0, 1]$, see also [56, Problem 7]. No classification theorem of the Gleason type is available for these examples, and it also seems that for them there is no one-to-one correspondence between pure probability measures and atoms of $L_f(V)$.

Finite-Dimensional Case: An Example

As already noted above, the structure of a quantum logic (S, L) satisfying $L \simeq L_f(V)$ with $\dim(V) < \infty$ may be substantially different from the infinite-dimensional case. To elucidate this further we continue Example 23.1 by determining the set of states S for the logic $L_f(\mathbb{Q}^n)$.

Consider the rational orthomodular space $\mathbb{Q}^n$ with the lattice $L(\mathbb{Q}^n) = L_f(\mathbb{Q}^n)$. For any $M \in L(\mathbb{Q}^n)$ one has $\mathbb{Q}^n = M + M^\perp$. Hence, for each $q \in \mathbb{Q}^n$ there is a unique decomposition $q = q_1 + q_2$, with $q_1 \in M$, $q_2 \in M^\perp$. This entails that the map $P_M : \mathbb{Q}^n \rightarrow \mathbb{Q}^n$ defined by $P_M q = q_1$ is linear, idempotent and Hermitian, that is, $f(q, P_M p) = f(P_M q, p)$ for all $q, p \in \mathbb{Q}^n$. For any atom $[v] \in L(\mathbb{Q}^n)$, one may thus define the map $\alpha_{[v]}$ via

$$\alpha_{[v]}(M) = \frac{f(v, P_M v)}{f(v, v)}; \quad (23.8)$$

this is a probability measure on $L(\mathbb{Q}^n)$ and its support is the defining atom, that is, $s(\alpha_{[v]}) = [v]$. Clearly, the mapping $\alpha_{[v]} \mapsto s(\alpha_{[v]})$ gives a one-to-one correspondence between the set of probability measures on $L(\mathbb{Q}^n)$ of the form $\alpha_{[v]}$ and the set of atoms of $L(\mathbb{Q}^n)$.

Let P_{at} be the set of states defined by the atoms of $L(\mathbb{Q}^n)$, that is, $\alpha \in P_{\text{at}}$ if $\alpha = \alpha_{[v]}$ for some $v \in \mathbb{Q}^n$, $v \neq 0$. Any σ -convex combination of states $(\alpha_{[v_i]})_{i \geq 1}$ with weights $(\lambda_i)_{i \geq 1}$ is again a state (probability measure) on $L(\mathbb{Q}^n)$. We let S_{at} denote the set of all such states. It has all the regularity properties of Sect. 23.2, including the strong ordering on $L(\mathbb{Q}^n)$. First of all, each $\alpha \in S_{\text{at}}$ has a support in $L(\mathbb{Q}^n)$; if $\alpha = \sum_i \lambda_i \alpha_{[v_i]}$, then $s(\alpha) = \vee \{[v_i] \mid \lambda_i \neq 0\}$. Moreover, if $s(\alpha) = [v]$ for some atom $[v]$, then $\alpha = \alpha_{[v]}$. Secondly, $\text{ex}(S_{\text{at}}) = P_{\text{at}}$, which also confirms that the restriction of the support projection to P_{at} defines a bijection between the sets $\text{ex}(S_{\text{at}})$ and $\text{At}(L(\mathbb{Q}^n))$.

Let S denote the set of all probability measures on $L(\mathbb{Q}^n)$. We demonstrate next that S_{at} is a proper subset of S .

To begin with, we note first that $L(\mathbb{Q}^n)$ can be naturally embedded in $L(\mathbb{R}^n)$. Indeed, for $M \in L(\mathbb{Q}^n)$, choose an orthogonal basis $e_1, \dots, e_k$, $k \leq n$, with $M = \text{lin}_{\mathbb{Q}}\{e_1, \dots, e_k\}$, and define $\tilde{M} = \text{lin}_{\mathbb{R}}\{e_1, \dots, e_k\}$. Then $L(\mathbb{Q}^n) \ni M \mapsto \tilde{M} \in L(\mathbb{R}^n)$ is an injective mapping.

Fix a nonzero vector $v \in \mathbb{R}^n$ such that at least one of its components is irrational. For $M \in L(\mathbb{Q}^n)$ define

$$\alpha_{[v]}(M) = \frac{\langle v \mid P_{\tilde{M}} v \rangle}{\langle v \mid v \rangle}, \quad (23.9)$$

where $\langle \cdot \mid \cdot \rangle$ denotes the natural inner product in $\mathbb{R}^n$. Clearly, $\alpha_{[v]}$ is a probability measure on $L(\mathbb{Q}^n)$. However, for any $q \in \mathbb{Q}^n$, $q \neq 0$, $\alpha_{[v]}([q]) \neq 0$, which shows that $\alpha_{[v]}$ has no support in $L(\mathbb{Q}^n)$. Hence $\alpha_{[v]}$ is not in S_{at} .

This example shows that the subspace lattice $L(\mathbb{Q}^n)$ of the non-Hilbertian orthomodular space $(\mathbb{Q}^n, \mathbb{Q}, \text{id}, f)$ admits a rich subset of states S_{at} that has all the listed regularity properties. Therefore, one could consider $(S_{\text{at}}, L(\mathbb{Q}^n))$ as a logic of a proper quantum system. In this case the logic admits also additional probability measures which cannot be considered as states of the quantum system since they are not supported in $L(\mathbb{Q}^n)$.

One may speculate whether $(V, K, *, f)$ can be forced to be a Hilbert space by requiring that the set of all pure probability measures is defined on $L_f(V)$ so as to be in one-to-one correspondence with the atoms of $L_f(V)$. Although this seems to be an appealing property, it remains a conjecture for now, or rather, a hope for the future.

23.5 The Role of Symmetries in the Representation Theorem

In his review [54] Holland formulated the axiom of *ample unitary group*, according to which for each pair of mutually orthogonal vectors $u, v \in V$, there is a bijective linear map $U : V \rightarrow V$ such that $U(v) = u$ and $f(Ux, Uy) = f(x, y)$ for each $x, y \in V$. Clearly, this assumption yields the crucial condition required by Solér's theorem. However, this is a very strong assumption, and, in any case, it is not a property given by the pair (S, L) . Rather than accepting this postulate we follow [61] to elucidate its possible physical content.

Implementing Symmetries as Operators on V

The idea of symmetry receives its natural mathematical representation as a transformation on the set of entities the symmetry refers to. The basic structures are now encoded in the sets L and S and in the duality between them. These sets possess various physically relevant structures which define the corresponding automorphism groups. From the outset, any of them could be used to formulate the notion of symmetry in quantum logic. In view of the theorem of Solér we shall consider only two of them: symmetries of the set $\text{At}(L)$ of atoms of L , and the symmetries of the logic L .⁶

Definition 23.5 (a) A mapping $\ell_o : \text{At}(L) \rightarrow \text{At}(L)$ is an *orthosymmetry* if it is bijective and for any $p, q \in \text{At}(L)$,

$$p \perp q \iff \ell_o(p) \perp \ell_o(q).$$

(b) A mapping $\ell : L \rightarrow L$ is a *symmetry* if it is bijective and it preserves the order and the orthocomplementation, that is, for any $a, b \in L$,

$$\begin{aligned} a \leq b &\iff \ell(a) \leq \ell(b), \\ \ell(a^\perp) &= \ell(a)^\perp. \end{aligned}$$

Let $\text{Aut}_o(\text{At}(L))$ and $\text{Aut}(L)$ denote the sets of orthosymmetries and symmetries on L , respectively. Both of these sets are groups with respect to the composition of mappings. Moreover, any symmetry ℓ , when restricted to $\text{At}(L)$, defines an orthosymmetry.

Assume now that the logic L allows a vector space coordinatisation of the form $L \simeq L_f(V)$ for an orthomodular space $(V, K, *, f)$. Any $\ell_o \in \text{Aut}_o(\text{At}(L))$ as well as $\ell \in \text{Aut}(L)$ defines the corresponding automorphism on the set $\mathbf{P}(V)$ of atoms of $L_f(V)$ and on the whole $L_f(V)$, respectively. We continue to denote them as ℓ_o and ℓ and call them orthosymmetry and symmetry, respectively.

Lemma 23.5 *Let $\ell_o \in \text{Aut}_o(\mathbf{P}(V))$. There is a unique symmetry $\hat{\ell}_o$ belonging to $\text{Aut}(L_f(V))$ such that $\hat{\ell}_o([v]) = \ell_o([v])$ for all $[v] \in \mathbf{P}(V)$. Moreover, the map $\text{Aut}_o(\mathbf{P}(V)) \ni \ell_o \mapsto \hat{\ell}_o \in \text{Aut}(L_f(V))$ is a group isomorphism.*

⁶Various definitions of the notion of symmetry in quantum mechanics are studied e.g. in [62, 63].

Proof Let $\ell_o \in \text{Aut}_o(\text{At}(\mathbf{L}_f(V)))$ and define for any (nonempty) subset $M \subset V$, $M \neq \{0\}$,

$$\hat{\ell}_o(M) = \{x \in \ell_o([v]) \mid v \in M, v \neq 0\}$$

and put $\hat{\ell}_o(\{0\}) = \{0\}$. Since $\text{Aut}_o(\text{At}(\mathbf{L}_f(V)))$ is a group we also have $\widehat{\ell_o^{-1}}$ defined in the same way. A direct computation shows that $\hat{\ell}_o(\widehat{\ell_o^{-1}}(M)) = KM$ and $\widehat{\ell_o^{-1}}(\hat{\ell}_o(M)) = KM$. Indeed, for any (nonempty) M , $M \neq \{0\}$, we have

$$\begin{aligned} \widehat{\ell_o^{-1}}(\hat{\ell}_o(M)) &= \{w \in \ell_o^{-1}([v]) \mid v \in \hat{\ell}_o(M)\} \\ &= \{w \in \ell_o^{-1}([v]) \mid v \in \ell_o([x]), x \in M, x \neq 0\} \\ &= \{w \in \ell_o^{-1}(\ell_o([x])) \mid x \in M, x \neq 0\} \\ &= \{w \in [x] \mid x \in M, x \neq 0\} = KM, \end{aligned}$$

and if $M = \{0\}$, then $\widehat{\ell_o^{-1}}(\hat{\ell}_o(\{0\})) = \{0\}$. Similarly, one gets the other set equality. If M is a subspace, then

$$\hat{\ell}_o(\widehat{\ell_o^{-1}}(M)) = M, \quad \widehat{\ell_o^{-1}}(\hat{\ell}_o(M)) = M.$$

Using the fact that for any two (nonzero) vectors $u, v \in V$ and for any $\ell_o \in \text{Aut}_o(\text{At}(\mathbf{L}_f(V)))$, $f(u, v) = 0$ is equivalent to $[u] \perp [v]$ and to $\ell_o([u]) \perp \ell_o([v])$ (meaning that $f(x, y) = 0$ for any $x \in \ell_o([u])$, $y \in \ell_o([v])$), one easily verifies that

$$\hat{\ell}_o(M)^\perp = \hat{\ell}_o(M^\perp)$$

for any (nonempty) set $M \subset V$.

Let now $M \in \mathbf{L}_f(V)$. Since $M = (M^\perp)^\perp$, we have $\hat{\ell}_o(M) = \hat{\ell}_o(M^\perp)^\perp$ so that by [10, Lemma 4.35] $\hat{\ell}_o(M)$ is an f -closed subspace of V , that is, $\hat{\ell}_o(M) \in \mathbf{L}_f(V)$. Hence, the map $\mathbf{L}_f(V) \ni M \mapsto \hat{\ell}_o(M) \in \mathbf{L}_f(V)$ is well defined. Clearly, it is a bijection, with the inverse $(\hat{\ell}_o)^{-1} = \widehat{\ell_o^{-1}}$, it preserves the orthocomplementation and, by construction, also the order. Hence, for any $\ell_o \in \text{Aut}_o(\text{At}(\mathbf{L}_f(V)))$, $\hat{\ell}_o \in \text{Aut}(\mathbf{L}_f(V))$.

Next we note that for all nonzero $v \in V$ one has $\hat{\ell}_o([v]) = \ell_o([v])$; this shows that $\hat{\ell}_o$ extends the map ℓ_o . Let $\ell \in \text{Aut}(\mathbf{L}_f(V))$, and assume that this is another extension of ℓ_o . Since the lattice $\mathbf{L}_f(V)$ is atomistic, we now have

$$\begin{aligned} \ell(M) &= \ell(\vee\{[v] \mid [v] \subset M\}) = \vee\{\ell([v]) \mid [v] \subset M\} \\ &= \vee\{\ell_o([v]) \mid [v] \subset M\} = \vee\{\hat{\ell}_o([v]) \mid [v] \subset M\} \\ &= \hat{\ell}_o(\vee\{[v] \mid [v] \subset M\}) = \hat{\ell}_o(M) \end{aligned}$$

for any $M \in \mathbf{L}_f(V)$, showing that $\ell = \hat{\ell}_o$, that is, the extension is unique. The map $\text{Aut}_o(\text{At}(\mathbf{L}_f(V))) \ni \ell_o \mapsto \hat{\ell}_o \in \text{Aut}(\mathbf{L}_f(V))$ is thus well defined. Its injectivity and

surjectivity are obvious and it also preserves the group structure: $\widehat{\ell_o^{-1}} = (\widehat{\ell_o})^{-1}$ and $\widehat{\ell_o \circ \ell'_o} = \widehat{\ell_o} \circ \widehat{\ell'_o}$ for all $\ell_o, \ell'_o \in \text{Aut}_o(\text{At}(\mathbf{L}_f(V)))$. $\square$

Let $\mathbf{L}(V)$ be the (complete, irreducible, modular, AC) lattice of all subspaces of V , and let $\text{Aut}(\mathbf{L}(V))$ be the group of order isomorphisms on $\mathbf{L}(V)$. Further, let $\mathbf{F}(V) = \{L \in \mathbf{L}(V) \mid \dim(L) < \infty\}$, and recall that $\mathbf{P}(V) \subset \mathbf{F}(V) \subset \mathbf{L}_f(V)$. Note also that any $M \in \mathbf{L}(V)$ can be expressed as $M = \cup\{L \in \mathbf{F}(V) \mid L \subset M\} = \vee\{L \in \mathbf{F}(V) \mid L \subset M\}$, and $\ell(L) \in \mathbf{F}(V)$ for any $L \in \mathbf{F}(V)$, $\ell \in \text{Aut}(\mathbf{L}_f(V))$.

Lemma 23.6 *For any $\ell \in \text{Aut}(\mathbf{L}_f(V))$,*

$$\Phi_\ell(M) = \cup\{\ell(L) \mid L \subset M, L \in \mathbf{F}(V)\}$$

defines an order-preserving bijection $\Phi_\ell : \mathbf{L}(V) \rightarrow \mathbf{L}(V)$ which extends the map ℓ .

Proof This is obtained by an adaptation of the proof of [64, Lemma 1]. We show first that $\Phi_\ell(M) \in \mathbf{L}(V)$ for any $M \in \mathbf{L}(V)$. Indeed, if $x \in \Phi_\ell(M)$, then $x \in \ell(L)$ for some $L \in \mathbf{F}(V)$, $L \subset M$, and thus $\lambda x \in \ell(L) \subset \Phi_\ell(M)$ for any $\lambda \in K$. Moreover, if $y \in \Phi_\ell(M)$, then $y \in \ell(H)$ for some $H \in \mathbf{F}(V)$, $H \subset M$, so that

$$\begin{aligned} x + y &\in \ell(L) + \ell(H) = \ell(L) \vee \ell(H) \\ &= \ell(L \vee H) = \ell(L + H) \subset \Phi_\ell(M), \end{aligned}$$

since the subspaces involved are all finite-dimensional. Hence $\Phi_\ell(M) \in \mathbf{L}(V)$ for any $M \in \mathbf{L}(V)$.

To prove that the map Φ_ℓ has an inverse, we use the following observation:

$$\{H = \ell(L) \mid L \in \mathbf{F}(V), L \subset M\} = \{H \in \mathbf{F}(V) \mid H \subset \Phi_\ell(M)\}.$$

Since $\ell \in \text{Aut}(\mathbf{L}_f(V))$ is a group, we also have

$$\Phi_{\ell^{-1}}(M) = \cup\{\ell^{-1}(L) \mid L \subset M, L \in \mathbf{F}(V)\},$$

and using the above observation one quickly confirms that

$$\Phi_{\ell^{-1}}(\Phi_\ell(M)) = \Phi_\ell(\Phi_{\ell^{-1}}(M))$$

for any $M \in \mathbf{L}(V)$. Hence, for any $\ell \in \text{Aut}(\mathbf{L}_f(V))$ the map $\Phi_\ell : \mathbf{L}(V) \rightarrow \mathbf{L}(V)$ is a bijection, with the inverse $\Phi_{\ell^{-1}}$. By definition the map Φ_ℓ preserves the order, that is, $\Phi_\ell(M) \subset \Phi_\ell(N)$ if and only if $M \subset N$ for any $M, N \in \mathbf{L}(V)$. Hence, $\Phi_\ell \in \text{Aut}(\mathbf{L}(V))$ for any $\ell \in \text{Aut}(\mathbf{L}_f(V))$.

We show next that Φ_ℓ extends ℓ . Let $M \in \mathbf{L}_f(V) \subset \mathbf{L}(V)$. We note that $\Phi_\ell(M) = \cup\{\ell(L) \mid L \in \mathbf{F}(V), L \subset M\}$ and $\ell(M) = \cup\{H \mid H \in \mathbf{F}(V), H \subset \ell(M)\}$; furthermore, $H \subset \ell(M)$ if and only if $H = \ell(F)$, $F \in \mathbf{F}(V)$, $F \subset M$. Therefore we observe that $\Phi_\ell(M) = \ell(M)$. Since any $M \in \mathbf{L}(V)$ can be expressed as $M = \vee\{L \mid L \in \mathbf{F}(V)\}$,

$L \subset M\}$, one readily verifies that Φ_ℓ is the only order isomorphism of $L(V)$ which extends $\ell \in \text{Aut}(L_f(V))$. $\square$

Let $S : V \rightarrow V$ be a bijective map which is g -linear for some isomorphism g of K (that is, S is bijective and additive on V and $S(\lambda v) = g(\lambda)Sv$ for all $v \in V$ and $\lambda \in K$). Such an S induces an order isomorphism on $L(V)$ by $\Phi_S(M) = \{Sv \mid v \in M\}$. Moreover, if S' is another bijective h -linear map $V \rightarrow V$ inducing the same order isomorphism, that is, $\Phi_S = \Phi_{S'}$, then there is a $\lambda \in K$ such that $Sv = \lambda S'v$ for any $v \in V$ [65, Proposition III.1.2, Corollary III.1.2]. The first fundamental theorem of projective geometry [65, p. 44] gives the converse result: if $\dim(V) \geq 4$, then for any $\Phi \in \text{Aut}(L(V))$ there is an isomorphism $g : K \rightarrow K$ and a bijective g -linear map $S : V \rightarrow V$ inducing Φ .

Let $\ell_o \in \text{Aut}_o(P(V))$, and let $S_\ell : V \rightarrow V$ thus be a bijective g_ℓ -linear map which induces the extension $\Phi_\ell \in \text{Aut}(L(V))$ associated with the extension $\ell \in \text{Aut}(L_f(V))$ of ℓ_o . Thus, for any $v \in V$, $v \neq 0$,

$$S_\ell(Kv) = \Phi_\ell(Kv) = \ell(Kv) = \ell_o(Kv).$$

Since ℓ_o preserves the orthogonality of atoms, the one-dimensional subspaces $S_\ell(Ku)$ and $S_\ell(Kv)$ are orthogonal if and only if Ku and Kv are orthogonal, that is, $f(Ku, Kv) = 0$ if and only if $f(S_\ell(Ku), S_\ell(Kv)) = 0$.

A direct computation shows that the map

$$(u, v) \mapsto g_\ell^{-1}(f(S_\ell u, S_\ell v)) = \tilde{f}(u, v)$$

is a Hermitian form on $(V, K, *)$ such that $\tilde{f}(u, v) = 0$ if and only if $f(u, v) = 0$ for all $u, v \in V$. By virtue of the infinite-dimensional version of the Birkhoff-von Neumann theorem [52], there is a nonzero $\varrho_\ell \in K$ such that $\tilde{f}(u, v) = \varrho_\ell f(u, v)$ for all $u, v \in V$. Moreover, ϱ_ℓ is a symmetric element of K , and since $\lambda \mapsto \lambda^*$ is an antiautomorphism of K one also has $\lambda \varrho_\ell = \varrho_\ell \lambda$ for all $\lambda \in K$, that is, $\varrho_\ell \in \text{Cent}(K)$, the centre of K .

Corollary 23.4 *For any $\ell_o \in \text{Aut}_o(P(V))$ there is an isomorphism $g_\ell : K \rightarrow K$ and a bijective orthogonality preserving g_ℓ -linear map $S_\ell : V \rightarrow V$ such that*

$$\ell_o(Kv) = S_\ell(Kv) \tag{23.10}$$

for any $v \in V$, $v \neq 0$. Moreover, there is a $\varrho_\ell \in \text{Cent}(K)$, $\varrho_\ell \neq 0$, $\varrho_\ell = \varrho_\ell^*$, such that

$$f(S_\ell u, S_\ell v) = g_\ell(\varrho_\ell) g_\ell(f(u, v)) \tag{23.11}$$

for all $v, u \in V$.

This corollary is a precursor of the theorem of Wigner according to which the ‘transition probability preserving bijections on the set of pure states’ are implemented by unitary or antiunitary operators acting on the underlying Hilbert space $\mathcal{H}$ of the

standard logic $L = L(\mathcal{H})$ [10]. In that frame, the orthosymmetries are exactly the transition probability-zero preserving bijections on the pure states. If $\dim(\mathcal{H}) \geq 3$ then this group coincides with the group of transition probability preserving bijections on the set of pure states [62, Corollary 4]. Now the length of L is at least 4 so that $\dim(V) \geq 3$.

Symmetries and the Solér Conditions

We now study the role of symmetry in providing a partial justification of the assumptions of Solér's theorem. Clearly, the result is obtained if $L \simeq L_f(V)$ has the following property: Given any two mutually orthogonal atoms $[x], [y] \in L_f(V)$, there are nonzero vectors $x' \in [x]$ and $y' \in [y]$ such that

$$f(x', x') = f(y', y'). \quad (23.12)$$

Before investigating the conditions the theorem of Solér imposes on the set of symmetries, we recall that a proper quantum object is an *elementary* quantum object with respect to a group G of (for instance, space-time) transformations if there is a group homomorphism $\sigma : G \rightarrow \text{Aut}_o(\mathbf{P}(V))$ and if for any pure state (atom) $[v] \in \mathbf{P}(V)$, the set $\{\sigma_g([v]) \mid g \in G\}$ of pure states (atoms) is complete in the sense of superpositions, that is, any other pure state (atom) $[u] \in \mathbf{P}(V)$ can be expressed as a superposition of some of the pure states (atoms) $\sigma_g([v])$, $g \in G$. Even though this does not solve our problem, it shows that for an elementary quantum object the set of symmetries $\text{Aut}_o(\mathbf{P}(V))$ is rather large and the notion of superposition has a role in it. The next lemma binds the above condition (23.12) more tightly to the issue at hand.

Lemma 23.7 *Let $[x], [y]$ be any two mutually orthogonal atoms in $L_f(V)$. If there are nonzero vectors $x' \in [x]$ and $y' \in [y]$ such that $f(x', x') = f(y', y')$ then there is an $\ell_o \in \text{Aut}_o(\mathbf{P}(V))$ which swaps $[x]$ and $[y]$, that is, $\ell_o([x]) = [y]$ and $\ell_o([y]) = [x]$. Moreover, there is a $[v] \leq [x] \vee [y]$ such that $\ell_o([v]) = [v]$.*

Proof Let $M = [x] \vee [y] = [x] \oplus [y]$. Clearly, $[x] = [x']$, $[y] = [y']$. Any $u \in M$ can be written uniquely as $u = \alpha x' + \beta y'$, $\alpha, \beta \in K$. Fix $\lambda \in \text{Cent}(K)$, $\lambda \neq 0$, and define

$$U_M(u) = U_M(\alpha x' + \beta y') = \lambda(\alpha y' + \beta x').$$

The map U_M is a linear bijection on M , and for any $u, v \in M$, $\lambda f(u, v)\lambda^* = f(U_M u, U_M v)$. Let $\bar{v} = x' + y'$ and observe that $[\bar{v}]$ is fixed by U_M . Since M is f -closed, $V = M + M^\perp$, so that any $w \in V$ can uniquely be decomposed as $w = w_1 + w_2$, with $w_1 \in M$, $w_2 \in M^\perp$. We define a canonical extension of U_M to the whole V by $Uw = U(w_1 + w_2) = U_M w_1 + \lambda w_2$. Then U is a bijective linear map on V . Moreover, $f(Uw, Uv) = \lambda f(w, v)\lambda^*$ for all $w, v \in V$, and for each $u \in M$, $Uu = U_M u$. Hence, in particular,

$$\Phi_U([x]) = [y], \quad \Phi_U([y]) = [x], \quad \Phi_U([\bar{v}]) = [\bar{v}]. \quad \square$$

This lemma shows that condition (23.12) implies the existence of a special symmetry of $\mathbf{L}_f(V)$ that interchanges the two orthogonal atoms $[x]$ and $[y]$ and has a superposition of them as a fixed point.

To get the opposite implication, and thus come to the final conclusion, we add the following assumptions concerning the group $\text{Aut}_o(\mathbf{P}(V))$ and the form f .

- (A) The symmetry group is *abundant* in the following sense: for any pair of mutually orthogonal atoms $[x], [y] \in \mathbf{P}(V)$ there is a symmetry $\ell_o \in \text{Aut}_o(\mathbf{P}(V))$ that swaps $[x]$ and $[y]$, that is, $\ell_o([x]) = [y]$ and $\ell_o([y]) = [x]$, and has some of their superpositions as a fixed point, that is, $\ell_o([v]) = [v]$ for some $[v] \leq [x] \vee [y]$;
- (R) The form f is *regular* in the following sense: for each $v \in V$, $f(v, v) \in \text{Cent}(K)$ and $g(f(v, v)) = f(v, v)$ for any automorphism g of K .

Lemma 23.8 *Let $[x], [y]$ be any two mutually orthogonal atoms in $\mathbf{L}_f(V)$. If the group $\text{Aut}_o(\mathbf{P}(V))$ is abundant and the form f is regular then there are nonzero vectors $x' \in [x]$ and $y' \in [y]$ such that $f(x', x') = f(y', y')$.*

Proof Let $\ell_o \in \text{Aut}_o(\mathbf{P}(V))$ be an orthosymmetry swapping the atoms $[x]$ and $[y]$ and having a $[v] \leq [x] \vee [y]$ as a fixed point. Let $S_\ell, g_\ell, \varrho_\ell$ constitute a realisation of ℓ_o as given in Corollary 23.4. Applying Eq. (23.11) first to the vector v and its transform $S_\ell v = \lambda v, \lambda \in K$, one gets $g_\ell(\varrho_\ell) = \lambda \lambda^*$. Applying then the same equation to x and $S_\ell x = \alpha x, \alpha \in K$, then yields

$$f(\alpha y, \alpha y) = g_\ell(\varrho_\ell)g_\ell(f(x, x)) = \lambda f(x, x)\lambda^* = f(\lambda x, \lambda x). \quad \square$$

We summarise the results of this section in the form of a theorem.

Theorem 23.8 *Assume that the logic (S, L) of the statistical duality $(S, \mathbf{O}, p)$ has the structure of Corollary 23.1 and that L has an abundant set of orthosymmetries. If there is an infinite sequence of mutually orthogonal atoms in L , and if the form f of the coordinatisation $(V, K, *, f)$ of the logic is regular, then V is a Hilbert space over $\mathbb{R}, \mathbb{C}$, or $\mathbb{H}$, and L is (ortho-order) isomorphic to the lattice of closed subspaces of the Hilbert space V .*

With this theorem, the statistical duality $(S, \mathbf{O}, p)$ of a proper quantum system is completely resolved: the states $\alpha \in S$ of the system are identified with positive trace one operators ϱ of an infinite-dimensional Hilbert space $\mathcal{H}$, the observables $(E, \Omega, \mathcal{A}) \in \mathbf{O}$ are expressed as semispectral measures, also called normalised positive operator measures, taking values in the set of bounded operators on $\mathcal{H}$, and the numbers $p(\alpha, E, X)$ are determined to be given by the ‘Born rule’ $p(\alpha, E, X) = \text{tr}[\varrho E(X)]$. The pure states are the one-dimensional projections and the L -valued observables are the spectral measures.

We are left with the question whether the regularity of the form f , the requirement (R), can be stated as a property of the logic (S, L) of the duality $(S, \mathbf{O}, p)$. Another open question is the choice of the number field left open by Theorem 23.6. This question will be discussed briefly in the final section.

23.6 The Case of the Complex Field

It is well known that the complex Hilbert space $\mathcal{H}$ is in many respects simpler than the real or quaternionic Hilbert spaces. We recall only the polarisation identity (valid in the complex case) and Stone's theorem, which is of fundamental importance. But is the choice $\mathbb{C}$ only a mathematical convenience? Some of the differences between the three cases are elucidated in [3, Chap. 22]. Without aiming at a systematic answer to this question, we wish to conclude this chapter with two physically motivated arguments (not discussed in [3]) distinguishing the complex case.

The first argument is due to [66]. Let $\mathbb{K}$ stand for $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and consider for each $r \in \mathbb{R}$, $r \neq 0$, a simple observable defined by $\{r\} \mapsto M$ and $\{0\} \mapsto M^\perp$. If $f : \mathbb{R} \rightarrow \mathbb{K}$ is any measurable function, one may integrate it with respect to this observable to get the operator $f(r, M) = f(r)P_M + f(0)P_{M^\perp}$. One may now pose the question whether there is an f such that for each (r, M) the operator $f(r, M)$ constitutes a symmetry transformation, that is, an automorphism on $L(\mathcal{H})$. Clearly, if $\mathbb{K} = \mathbb{C}$, the function $x \mapsto e^{ix}$ defines a unitary operator $e^{ir}P_M + P_{M^\perp}$ for each (r, M) , and the family of the associated automorphisms $U_M^r : N \mapsto U_M^r(N)$ with the following properties:

- (i) $U_M^r = \text{id}$ if and only if either $M \in \{\{0\}, \mathcal{H}\}$ or $r \in 2\pi\mathbb{Z}$;
- (ii) if $U_M^r \neq \text{id}$, then $U_M^r(N) = N$ if and only if M and N are compatible⁷;
- (iii) if $M \subset N^\perp$, then $U_M^r \circ U_N^s = U_{M \vee N}^s$;
- (iv) $U_M^r \circ U_M^s = U_M^{r+s}$ for all $r, s \in \mathbb{R}$ and $M \in L(\mathcal{H})$.

In [66] it is shown that it is only the case $\mathbb{K} = \mathbb{C}$ which admits a family of automorphisms U_M^r satisfying (i)–(iv).

The other argument favouring $\mathbb{C}$ over $\mathbb{R}$ concerns the derivability of Heisenberg–Kennard–Robertson-type preparation uncertainty relations. This argument is based on the properties of quadratic functionals and it can be formulated in the frame of the so-called sum logics (L, S) [67].

Let V be a real vector space. A function $f : V \rightarrow \mathbb{R}$ is a *quadratic functional* if it is non-negative and has the property

$$f(x+y) + f(x-y) = 2f(x) + 2f(y) \quad (23.13)$$

for all $x, y \in V$. Such functionals f satisfy the inequality

$$f(x) \cdot f(y) \geq \frac{1}{4} |f(x+y) - f(x) - f(y)|^2 \quad (23.14)$$

for all $x, y \in V$.

The logic (L, S) is a *sum logic* if S is sufficient and for any two bounded (that is, compactly supported) real observables $A, B \in \mathcal{O}$ there is a unique observable $C \in \mathcal{O}$ such that for any state $\alpha \in S$,

⁷See Exercise 10 for the definition of compatibility.

$$\text{Exp}(\mathbf{C}, \alpha) = \text{Exp}(\mathbf{A}, \alpha) + \text{Exp}(\mathbf{B}, \alpha),$$

where, for instance, $\text{Exp}(\mathbf{A}, \alpha)$ denotes the first moment, that is the expectation $\int_{\mathbb{R}} x d(\alpha \circ \mathbf{A})(x)$ of the probability measure $\alpha \circ \mathbf{A}$.

Under the conditions of Theorem 23.8, $(\mathbf{L}, \mathbf{S})$ is a sum logic. We denote $\mathbf{C} = \mathbf{A} + \mathbf{B}$. Since for any $\lambda \in \mathbb{R}$ we may define $\lambda \mathbf{A}$ as the observable with the property $\text{Exp}(\lambda \mathbf{A}, \alpha) = \lambda \text{Exp}(\mathbf{A}, \alpha)$, we conclude that the set $\mathbf{O}_b$ of bounded observables is a real vector space. This logic is also a *quadratic logic* in the sense that there are functions $f : \mathbf{O}_b \times \mathbf{S} \rightarrow \mathbb{R}$ such that $f(\cdot, \alpha)$ is a quadratic functional for any α . For instance, the second moment $f_1(\mathbf{A}, \alpha) = \text{Exp}(\mathbf{A}^2, \alpha)$ as well as the variance $f_2(\mathbf{A}, \alpha) = \text{Var}(\mathbf{A}, \alpha)$ are such functionals and $f_1(\mathbf{A} - \text{Exp}(\mathbf{A}, \alpha), \alpha) = f_2(\mathbf{A}, \alpha)$ for all $\mathbf{A} \in \mathbf{O}_b$ and $\alpha \in \mathbf{S}$. Thus, for instance,

$$\text{Var}(\mathbf{A}, \alpha) \cdot \text{Var}(\mathbf{B}, \alpha) \geq \frac{1}{4} |\text{Var}(\mathbf{A} + \mathbf{B}, \alpha) - \text{Var}(\mathbf{A}, \alpha) - \text{Var}(\mathbf{B}, \alpha)|^2$$

for all $\mathbf{A}, \mathbf{B} \in \mathbf{O}_b$ and $\alpha \in \mathbf{S}$.

To obtain a Heisenberg–Kennard–Robertson-type inequality, we extend the real vector space $\mathbf{O}_b$ to the complex one $\hat{\mathbf{O}}_b = \{\mathbf{A} + i\mathbf{B} \mid \mathbf{A}, \mathbf{B} \in \mathbf{O}_b\}$, where $i = \sqrt{-1}$. Let Φ be a bilinear $\mathbf{O}_b$ -valued mapping on the Cartesian product $\mathbf{O}_b \times \mathbf{O}_b$ and assume that it has the property that $\Phi(\mathbf{A}, b) = \Phi(a, \mathbf{B}) = 0$ for any constant observables $a, b \in \mathbb{R}$, that is, observables with the supports $\{a\}$ and $\{b\}$. Define the following function

$$\hat{f}_1(\mathbf{A} + i\mathbf{B}, \alpha) = f_1(\mathbf{A}, \alpha) + f_1(\mathbf{B}, \alpha) + f_0(\Phi(\mathbf{A}, \mathbf{B}), \alpha),$$

where f_0 is the expectation functional on $\mathbf{O}_b \times \mathbf{S}$, that is, $f_0(\mathbf{A}, \alpha) = \text{Exp}(\mathbf{A}, \alpha)$ for any $\mathbf{A} \in \mathbf{O}_b$, $\alpha \in \mathbf{S}$. Clearly, $\hat{f}_1$ is real valued on function on $\hat{\mathbf{O}}_b \times \mathbf{S}$ and its restriction to $\mathbf{O}_b \times \mathbf{S}$ equals f_1 . By the linearity of f_0 and by the bilinearity of Φ one gets

$$\hat{f}_1(\mathbf{A} + i\mathbf{B}, \alpha) + \hat{f}_1(\mathbf{A} - i\mathbf{B}, \alpha) = 2f_1(\mathbf{A}, \alpha) + 2f_1(\mathbf{B}, \alpha)$$

for all $\mathbf{A}, \mathbf{B} \in \mathbf{O}_b$, $\alpha \in \mathbf{S}$. Assuming that $\hat{f}_1$ is also non-negative, we see that $\hat{f}_1(\cdot, \alpha)$ is a quadratic functional on $\hat{\mathbf{O}}_b$ for any $\alpha \in \mathbf{S}$. According to the complex version of the result (23.14) one gets

$$\hat{f}_1(\mathbf{A}, \alpha) \cdot \hat{f}_1(i\mathbf{B}, \alpha) \geq \frac{1}{4} |\hat{f}_1(\mathbf{A} + i\mathbf{B}, \alpha) - \hat{f}_1(\mathbf{A}, \alpha) - \hat{f}_1(i\mathbf{B}, \alpha)|^2$$

for all $\mathbf{A}, \mathbf{B} \in \mathbf{O}_b$, $\alpha \in \mathbf{S}$. Noting that $\hat{f}_1(\mathbf{A}, \alpha) = f_1(\mathbf{A}, \alpha)$ and $\hat{f}_1(i\mathbf{B}, \alpha) = f_1(\mathbf{B}, \alpha)$, we may rewrite this inequality as

$$f_1(\mathbf{A}, \alpha) \cdot f_1(\mathbf{B}, \alpha) \geq \frac{1}{4} |\text{Exp}(\Phi(\mathbf{A}, \mathbf{B}), \alpha)|^2.$$

Since Φ has the usual scaling invariance property (of a commutator),

$$\text{Exp}(\Phi(\mathbf{A} - \text{Exp}(\mathbf{A}, \alpha), \mathbf{B} - \text{Exp}(\mathbf{B}, \alpha)), \alpha) = \text{Exp}(\Phi(\mathbf{A}, \mathbf{B}), \alpha),$$

one obtains the familiar inequality

$$\text{Var}(\mathbf{A}, \alpha) \cdot \text{Var}(\mathbf{B}, \alpha) \geq \frac{1}{4} |\text{Exp}(\Phi(\mathbf{A}, \mathbf{B}), \alpha)|^2. \quad (23.15)$$

To conclude, if one wants to derive the Heisenberg–Kennard–Robertson-type inequality (23.15) from an axiom like Eq. (23.13), this can be done only after extending the real vector space of observables to a complex one. This is a further justification for the choice of the complex field for quantum mechanics on a Hilbert space.

23.7 Exercises

1. Show that the function $[0, 1] \times \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$ defined by (23.1) has the following properties:

$$\begin{aligned} \langle \lambda, \alpha, \beta \rangle &= \langle 1 - \lambda, \beta, \alpha \rangle \text{ for all } \lambda, \alpha, \beta, \\ \langle \lambda, \alpha, \alpha \rangle &= \alpha \text{ for all } \lambda, \alpha, \\ \langle 0, \alpha, \beta \rangle &= \beta \text{ for all } \alpha, \beta, \\ \text{if } \langle \lambda, \alpha, \beta \rangle &= \langle \lambda, \alpha, \gamma \rangle \text{ for some } \lambda \neq 1, \alpha \in \mathcal{S}, \text{ then } \beta = \gamma, \\ \langle \lambda, \alpha, \langle \mu, \beta, \gamma \rangle \rangle &= \langle \lambda + (1 - \lambda)\mu, \langle \lambda(\lambda + (1 - \lambda)\mu)^{-1}, \alpha, \beta \rangle, \gamma \rangle, \\ \text{for all } \lambda, \mu, \text{ with } \lambda + (1 - \lambda)\mu &\neq 0, \text{ and } \alpha, \beta, \gamma. \end{aligned}$$

If a nonempty set $\mathcal{S}$ is equipped with a function $[0, 1] \times \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$ with the above five properties, then $\mathcal{S}$ is called a *convex (pre-)structure*. If $\mathcal{S}$ is a convex structure, then there is a real vector space U and a bijective affine function $f : \mathcal{S} \rightarrow U$ such that $f(\mathcal{S}) \subset U$ is a convex set [26, Theorem 2].

2. Let $\mathcal{T}_s(\mathcal{H})$ be the real vector space of selfadjoint trace class operators on $\mathcal{H}$ and let $\mathcal{T}_s(\mathcal{H})^+$ be its subset of positive operators. Show that $\mathcal{T}_s(\mathcal{H})^+$ is a proper cone generating $\mathcal{T}_s(\mathcal{H})$.
3. Let tr denote the trace functional on $\mathcal{T}_s(\mathcal{H})$. Recall that it is a strictly positive linear functional on $\mathcal{T}_s(\mathcal{H})$. Let $\mathcal{S}(\mathcal{H})$ denote the set of positive trace one operators on $\mathcal{H}$. Show that $\mathcal{S}(\mathcal{H})$ is a base for the cone $\mathcal{T}_s(\mathcal{H})^+$.
4. Show that $\text{conv}(\mathcal{S} \cup -\mathcal{S})$ is a convex, absorbing, and balanced set.
5. Show that the Minkowski functional of $\text{conv}(\mathcal{S} \cup -\mathcal{S})$ is a seminorm [30, Theorem II.1.4].
6. Show that any $f \in \mathcal{E}$ has a unique extension to a positive continuous linear functional on V bounded by e .

7. Show that any map $\phi : V^+ \rightarrow V^+$ with the properties (23.5) and (23.6) has a unique extension to a positive linear contracting mapping of V into V .
8. Show that the map $\mathcal{A} \ni X \mapsto M_E(X) \in L$ preserves both the order (i.e. $X \subset Y \implies M_E(X) \leq M_E(Y)$) and the complementation (i.e. $M_E(X') = M_E(X)^\perp$).
9. Show that the range $E(\mathcal{A})$ of an observable $E \in \tilde{\mathcal{O}}$ is a Boolean sub- σ -algebra of L .
10. Any two elements $a, b \in L$ are *compatible* if there exist three mutually orthogonal elements $a_1, b_1, c \in L$ such that $a = a_1 \vee c$ and $b = b_1 \vee c$. Show that $a, b \in L$ are compatible if and only if there is an observable $E \in \tilde{\mathcal{O}}$ such that $a = E(X)$ and $b = E(Y)$ for some X and Y .
11. Show that the centre of L , $\text{Cent}(L)$, is a Boolean sub-algebra of L .
12. Show that any two projection operators on a (complex separable) Hilbert space are compatible (with respect to the natural ortho-order structure of the projection lattice) if and only if the projections commute.
13. Show that the covering property of an atomic lattice L is equivalent to the property: for any $a \in L$, $p \in \text{At}(L)$, the element $(a \vee p) \wedge a^\perp$ is either an atom or 0.
14. Show that the logic of a proper quantum system (either in the sense of superpositions or complementarity) is irreducible.
15. Show that the filter ϕ_α associated with the pure state α has the form $\phi_\alpha(\beta) = e(\phi_\alpha(\beta))\alpha$ for any $\beta \in P$.
16. Show that for any operation $\phi \in \mathcal{O}$ there exists an operation $\phi' \in \mathcal{O}$ such that $e \circ \phi' = (e \circ \phi)^\perp$.
17. Show that any filter is repeatable, $\phi^2(\alpha) = \phi(\alpha)$ for any $\alpha \in P$.
18. Show that the set $\{\phi_P \mid P \in \mathcal{P}(\mathcal{H})\}$ of Lüders operations defines a sufficiently rich family of filters in the Hilbert space frame.
19. Show that in the Hilbert space frame the set of propositions of Axiom 5 coincides with the set $\mathcal{P}(\mathcal{H})$ of projection operators and that the projection postulate is a theorem in this case.
20. Show that for any triple (a, b, c) of mutually orthogonal elements of L , $a + b + c \in L$.
21. Work out the details of the proof of Corollary 23.2.
22. Show that the formula (23.8) defines a probability measure on the logic $L(\mathbb{Q})$.
23. Show that the support of the state (23.8) is its defining atom.
24. Show that $\text{ex}(S_{\text{at}}) = P_{\text{at}}$.
25. Use the matrix representation of the projection $P_{[v]}$ (with respect to the canonical basis), define the positive trace one matrix $\varrho_v = P_{[v]}/\text{tr}[P_{[v]}]$ and write the probability (23.8) in trace form $\alpha_{[v]}(M) = \text{tr}[\varrho_v P_M]$. Hint: note that $f(v, v) = \text{tr}[P_{[v]}]$. Show that, in general, for an $\alpha \in S_{\text{at}} \subset S$ there is no $\mathbb{Q}$ -linear operator $\varrho : \mathbb{Q}^n \rightarrow \mathbb{Q}^n$ for which the probabilities $\alpha(M)$ could be expressed as $\text{tr}[\varrho P_M]$.
26. Prove Lemma 23.5.
27. Prove Lemma 23.6.
28. Prove Lemma 23.7.
29. Prove Lemma 23.8.

30. With the notations of Sect. 23.6 show that in the case $\mathbb{K} = \mathbb{C}$ the unitary operators U_M^r share the properties (i)–(iv).
31. Show that (23.13) implies (23.14).
32. Show that under the conditions of Theorem 23.8, (L, S) is a sum logic.
33. Show that the functionals $f_1(A, \alpha) = \text{Exp}(A^2, \alpha)$ and $f_2(A, \alpha) = \text{Var}(A, \alpha)$ are quadratic.
34. Fill in the details to obtain the inequality (23.15).

References

1. Birkhoff, G., von Neumann, J.: The logic of quantum mechanics. *Ann. Math.* **37**(4), 823–843 (1936)
2. Mackey, G.W.: *The Mathematical Foundations of Quantum Mechanics* W.A. Benjamin, Inc., New York-Amsterdam (1963)
3. Beltrametti, E., Cassinelli, G.: *The Logic of Quantum Mechanics*. *Encyclopedia of Mathematics and its Applications*, Vol. 15. Addison-Wesley Publishing Co., Massachusetts (1981)
4. Gudder, S.P.: *Stochastic Methods in Quantum Mechanics*. North-Holland, New York (1979). North-Holland Series in Probability and Applied Mathematics
5. Jauch, J.M.: *Foundations of Quantum Mechanics*. Addison-Wesley Publishing Co., Massachusetts London-Don Mills, (1968)
6. Mittelstaedt, P.: *Quantum Logic*. *Synthese Library*, vol. 126. D. Reidel Publishing Co., Dordrecht (1978)
7. Piron, C.: *Foundations of Quantum Physics*. W.A. Benjamin, Inc., Massachusetts-London-Amsterdam (1976)
8. Pták, P., Pulmannová, S.: *Orthomodular Structures as Quantum Logics*. *Fundamental Theories of Physics*, vol. 44. Kluwer Academic Publishers Group, Dordrecht (1991). Translated from the 1989 Slovak original by the authors
9. Rédei, M.: *Quantum Logic in Algebraic Approach*. *Fundamental Theories of Physics*, vol. 91. Kluwer Academic Publishers Group, Dordrecht (1998)
10. Varadarajan, V.S.: *Geometry of Quantum Theory*, 2nd edn. Springer, New York (1985)
11. Ludwig, G.: Attempt of an axiomatic foundation of quantum mechanics and more general theories. II. *Comm. Math. Phys.* **4**, 331–348 (1967)
12. Mielnik, B.: Theory of filters. *Comm. Math. Phys.* **15**, 1–46 (1969)
13. Mielnik, B.: Generalized quantum mechanics. *Comm. Math. Phys.* **37**, 221–256 (1974)
14. Davies, E.B., Lewis, J.T.: An operational approach to quantum probability. *Comm. Math. Phys.* **17**, 239–260 (1970)
15. Edwards, C.M.: The operational approach to algebraic quantum theory. I. *Comm. Math. Phys.* **16**, 207–230 (1970)
16. Edwards, C.M.: Classes of operations in quantum theory. *Comm. Math. Phys.* **20**, 26–56 (1971)
17. Edwards, C.M.: The theory of pure operations. *Comm. Math. Phys.* **24**, 260–288 (1972)
18. Ludwig, G.: *Foundations of Quantum Mechanics*. I. *Texts and Monographs in Physics*. Springer, New York (1983). Translated from the German by Carl A. Hein
19. Hartkämper, A., Neumann, H.: (eds.) *Foundations of Quantum Mechanics and Ordered Linear Spaces*. *Lecture Notes in Physics*, Vol. 29. Springer, Berlin (1974)
20. Davies, E.B.: *Quantum Theory of Open Systems*. Academic Press London, New York (1976)
21. Foulis, D.J., Randall, C.H.: The empirical logic approach to the physical sciences. In: *Foundations of Quantum Mechanics and Ordered Linear Spaces*. *Lecture Notes in Phys.*, Vol. 29, pp. 230–249. Springer, Berlin (1974)
22. Foulis, D., Randall, C.: Empirical logic and quantum mechanics. *Synthese* **29**(1–4), 81–111 (1974)

23. Chiribella, G., D'Ariano, G.M., Perinotti, P.: Informational derivation of quantum theory. *Phys. Rev. A* **84**, 012311 (2011)
24. Clifton, R., Bub, J., Halvorson, H.: Characterizing quantum theory in terms of information-theoretic constraints. *Found. Phys.* **33**(11), 1561–1591 (2003)
25. Cassinelli, G., Lahti, P.: An Axiomatic Basis for Quantum Mechanics. [arXiv:1508.03709](https://arxiv.org/abs/1508.03709) (2015)
26. Stone, M.H.: Postulates for the barycentric calculus. *Ann. Mat. Pure Appl.* **29**, 25–30 (1949)
27. Gudder, S.: Convex structures and operational quantum mechanics. *Comm. Math. Phys.* **29**, 249–264 (1973)
28. Cornette, W.M., Gudder, S.P.: The mixture of quantum states. *J. Math. Phys.* **15**, 842–850 (1974)
29. Ellis, A.J.: The duality of partially ordered normed linear spaces. *J. Lond. Math. Soc.* **39**, 730–744 (1964)
30. Schaefer, H.H.: *Topological Vector Spaces*, vol. 3. Springer, New York (1971). Third printing corrected, Graduate Texts in Mathematics
31. Edwards, C.M., Gerzon, M.A.: Monotone convergence in partially ordered vector spaces. *Ann. Inst. H. Poincaré Sect. A (N.S.)* **12**, 323–328 (1970)
32. Ellis, A.J.: Linear operators in partially ordered normed vector spaces. *J. Lond. Math. Soc.* **41**, 323–332 (1966)
33. Alfsen, E.M.: *Compact Convex Sets and Boundary Integrals*, vol. 57. Springer, New York (1971). *Ergebnisse der Mathematik und ihrer Grenzgebiete Band*
34. Schaefer, H.H.: Orderings of vector spaces. In: Hartkämper, A., Neumann, H. (eds.) *Foundations of Quantum Mechanics and Ordered Linear Spaces. Lecture Notes in Phys.*, pp. 4–10, Vol. 29. Springer, Berlin (1974)
35. Mączyński, M.: The orthogonality postulate in axiomatic quantum mechanics. *Int. J. Theor. Phys.* **8**, 353–360 (1973)
36. Zierler, N.: Axioms for non-relativistic quantum mechanics. *Pacific J. Math.* **11**, 1151–1169 (1961)
37. Jauch, J., Piron, C.: Can hidden variables be excluded in quantum mechanics? *Helv. Phys. Acta* **36**, 827–837 (1963)
38. Bugajska, K., Bugajski, S.: The projection postulate in quantum logic. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **21**, 873–877 (1973)
39. Dirac, P.: *The Principles of Quantum Mechanics*, 4th edn. Oxford, Clarendon Press (1981)
40. Stone, M.: The theory of representations of boolean algebras. *Trans. Am. Math. Soc.* **40**, 37–111 (1936)
41. Loomis, L.: On the representation of σ -complete boolean algebras. *Bull. Amer. Math. Soc.* **35**, 757–760 (1947)
42. Sikorski, R.: On the inducing of homomorphisms by mappings. *Fund. Math.* **36**, 7–22 (1949)
43. Bohr, N.: Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **48**, 696–702 (1935)
44. Pool, J.C.T.: Baer*-semigroups and the logic of quantum mechanics. *Comm. Math. Phys.* **9**, 118–141 (1968)
45. Pool, J.C.T.: Semimodularity and the logic of quantum mechanics. *Comm. Math. Phys.* **9**, 212–228 (1968)
46. Cassinelli, G., Beltrametti, E.: Ideal, first-kind measurements in a proposition-state structure. *Comm. Math. Phys.* **40**, 7–13 (1975)
47. Lahti, P., Bugajski, S.: Fundamental principles of quantum theory. II. From a convexity scheme to the DHB theory. *Int. J. Theor. Phys.* **24**(11), 1051–1080 (1985)
48. Stachow, E.-W.: Structures of quantum language for individual systems. In: Mittelstaedt, P., Stachow, E.-W. (eds.) *Recent Developments in Quantum Logic: Proceedings of the International Symposium on Quantum Logic. Cologne, Germany, June 13–16, 1984*, pp. 129–145. Bibliographisches Institut - Wissenschaftsverlag, Mannheim (1985)
49. P. Mittelstaedt and E.-W. Stachow, editors. *Recent Developments in Quantum Logic: Proceedings of the International Symposium on Quantum Logic, Cologne, Germany, June 13–16, 1984*, Mannheim, 1985. Bibliographisches Institut - Wissenschaftsverlag

50. Lahti, P., Mączyński, M.: Orthomodularity and quadratic transformations in probabilistic theories of physics. *J. Math. Phys.* **33**(12), 4133–4138 (1992)
51. Gleason, A.M.: Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.* **6**, 885–893 (1957)
52. Maeda, F., Maeda, S.: *Theory of Symmetric Lattices*. Springer, New York (1970). *Die Grundlehren der mathematischen Wissenschaften*, Band 173
53. Piziak, R.: Orthomodular lattices and quadratic spaces: a survey. *Rocky Mt. J. Math.* **21**, 951–992 (1991)
54. Holland Jr., S.S.: Orthomodularity in infinite dimensions; a theorem of M. Solér. *Bull. Amer. Math. Soc. (N.S.)* **32**(2), 205–234 (1995)
55. Keller, H.A.: Ein nicht-klassischer Hilbertscher Raum. *Math. Z.* **172**(1), 41–49 (1980)
56. Gross, U.-M.K.H.: On a class of orthomodular quadratic spaces. *L'Enseignement Mathématique* **31**, 187–212 (1985)
57. Solér, M.P.: Characterization of Hilbert spaces by orthomodular spaces. *Comm. Algebra* **23**(1), 219–243 (1995)
58. Morash, R.: Angle bisection and orthoautomorphisms in hilbert lattices. *Can. J. Math.* **25**, 261–272 (1973)
59. Wilbur, W.: On characterizing the standard quantum logic. *Trans. Am. Math. Soc.* **233**, 265–281 (1977)
60. Keller, H.: Measures on non-classical hilbertian spaces. *Notas mathematicas*, Universidad Catoliga Santiago, Chile **16**, 49–71 (1984)
61. Cassinelli, G., Lahti, P.: A theorem of Solér, the theory of symmetry, and quantum mechanics. *Int. J. Geom. Methods Mod. Phys.* **9**(7), 1260005 (2012)
62. Cassinelli, G., De Vito, E., Lahti, P., Levrero, A.: *The Theory of Symmetry Actions in Quantum Mechanics - With An Application to the Galilei Group*. *Lecture Notes in Physics*, vol. 654. Springer, Berlin (2004)
63. Molnár, L.: *Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces*. Volume 1895 of *Lecture Notes in Mathematics*. Springer, New York (2007)
64. Fillmore, P.A., Longstaff, W.E.: On isomorphisms of lattices of closed subspaces. *Canad. J. Math.* **36**(5), 820–829 (1984)
65. Baer, R.: *Linear Algebra and Projective Geometry*. Academic Press Inc., New York (1952)
66. Pulmannová, S.: Axiomatization of quantum logics. *Int. J. Theoret. Phys.* **35**(11), 2309–2319 (1996)
67. Lahti, P., Mączyński, M.: Heisenberg inequality and the complex field in quantum mechanics. *J. Math. Phys.* **28**(8), 1764–1769 (1987)

Index

Symbols

C^* -algebra, 132
 $L(f, E) = E[f]$, 81, 104
 $L(x^k, E)$, 175
 $L^2(\Omega, \mathcal{A}, \mu) = L^2(\mu)$, 68
 $\mathcal{A}$ -partition, 69
 $\Phi(f) = \int_{\Omega} f dE$, 85
 α -spread, 278
 $\mathcal{C}(\mathcal{H})$, 37
 $E(\mathcal{H})$, 196
 $\mathcal{F}(\mathcal{H})$, 37
 $\mathcal{D}(f, E)$, 104
 $\mathcal{E}$ -equivalent, 406
 $\mathcal{H}$, 16
 $\mathcal{H} \otimes \mathcal{K}$, 33
 $\mathcal{HS}(\mathcal{H})$, 45
 $\int_{\Omega} f d\mu = \int_{\Omega} f(\omega) d\mu(\omega)$, 66
 $\mathcal{L}(\mathcal{H})$, 17
 $\mathcal{L}_s(\mathcal{H})$, 19
 $\mathcal{L}_s(\mathcal{H})^+$, 19
 $\mathbb{T}$, 110
 $\mathcal{A}$ -measurable function, 65
 $\mathcal{A}$ -measurable set, 64
 $\mathfrak{M}$, 154
 $O(A)$, 204
 $O(\Omega, \mathcal{A}, \mathcal{H})$, 202
 $O(\Omega, \mathcal{B}(\Omega), \mathcal{H})$, 202
 $\mu * \nu$, 169
 $\mathcal{O}(\mathcal{H})$, 209
 $\mathcal{P}(\mathcal{H})$, 22
 $P_1(\mathcal{H})$, 193
 E -compatible, 155
 $E[k]$, 175
 σ -additive, 65
 σ -algebra, 63
 σ -ring, 74
 σ -weak topology, 127

$\mathfrak{I}$, 154
 $\mathfrak{I}^*$, 154
 $\mathcal{T}(\mathcal{H})$, 47
 $\widetilde{\mathcal{D}}(f, E)$, 105
 n -positive, 137
 $p \star E$, 169
 ℓ^2 , 15
 $\text{Exp}(E_{\varrho})$, 203
 $\text{Var}(E_{\varrho})$, 203
 $*$ -algebra, 131
 $*$ -ideal, 38
 $*$ -representation, 133
 $*$ -subalgebra, 131

A

Absolute value of an operator, 25
 Additive, 71
 Adjoint, 18, 138
 Algebra, 130
 Algebra (of sets), 64
 Almost everywhere, 67
 Anti-Cauchy inequality, 322
 Approximation, 289
 finite error bar width, 295
 finite noise, 302
 unbiased, 289
 Associate observable, 155

B

Banach $*$ -algebra, 131
 Banach algebra, 131
 Beam splitter, 438
 Bilinear, 134
 Bimeasure, 76
 Biobservable, 234, 262

Biorthogonal decomposition, 59
 Bochner integral, 66
 Boolean algebra, 64
 Borel σ -algebra, 63
 Borel set, 63
 Bounded above, 20
 Bounded bilinear map, 139
 Bounded linear map, 17
 Bounded sesquilinear form, 18
 Bra, 39

C

Calibration error, 289, 293
 Carathéodory–Hahn extension theorem, 73
 Cauchy net, 147
 Cauchy–Schwarz inequality, 13
 Cayley transform, 112
 Channel, 156
 Closure of an operator, 103
 Coherent state, 435
 Commutative algebra, 131
 Commuting positive operator measures, 77
 Commuting spectral measures, 77
 Compact operator, 37
 Complete measure space, 67
 Complete orthonormal system, 30
 Completely positive, 137, 154
 Completion, 67, 135
 Complex measure, 69
 Compound system
 composition rules, 214
 Continuous spectrum, 40, 117
 Convergence in measure, 66
 Convergence of a net, 20
 Convolution, 169
 Countably generated σ -algebra, 144
 Counting measure, 97
 Coupling, 290
 Covariant time observable, 394
 Covering property (of a lattice), 512
 Cyclic representation, 133
 Cyclic vector, 133

D

Dawson’s integral, 414
 Decomposable operator, 180
 Degeneracy, 40
 Densely defined, 101
 Density operator, 192
 Direct integral, 179
 Direct sum, 31

Direct sum of representations, 133
 Directed set, 20
 Disturbance
 minimal, 512
 Dominated convergence theorem, 71
 Dual, 17
 Duality, 128
 Dual map, 128

E

Effect, 506
 decision, 506
 sharp, 506
 Eigenspace, 40
 Eigenvalue, 40, 117
 Eigenvector, 40, 117
 Entanglement, 219
 EPR, 470
 Error
 random, 289
 systematic, 289
 value deviation, 299
 Error analysis, 288
 Error bar width, 295
 bias-free, 297
 Error measure
 state-independent, 288
 state-specific, 288
 Error-disturbance relation, 275
 Essential representation, 133
 Essential subspace, 133
 Essential supremum, 98
 Essentially bounded, 98
 Experimental function, 504
 Experimental statement, 503

F

Filter, 515
 repeatable, 516
 weakly repeatable, 516
 Final projection, 26
 Finite rank operator, 37
 Form
 Hermitian, 519
 quadratic, 530
 Fourier coefficient, 30
 Fréchet–Riesz representation theorem, 17

G

Generalised eigenvalue, 183
 Generalised eigenvector, 183

Generalised sequence, 20

Graph, 101

Greatest lower bound, 20

H

Heisenberg pair, 347

Hilbert codimension, 31

Hilbert dimension, 31

Hilbert–Schmidt operator, 45

Hilbert space, 16

Hilbert space completion, 135

Hilbert sum, 31

Hilbert tensor product, 33

Hilbert tensor product of operators, 56

Homomorphism, 131

I

Ideal, 131

left, 131

right, 131

two-sided, 131

Identity (element), 131

Instrument, 154

Incompatibility, 262

degree of, 273

measure of, 335

Indeterminacy, 8

Infimum, 20

Informationally complete, 406

Informationally equivalent, 406

Initial projection, 26

Inner product, 13

Inner product space, 13

Instrument, 506

completely positive, 155

Lüders, 239

von Neumann, 239

Integrable, 70

Integrable function, 66

Integrable simple function, 66

Integral, 66, 70, 80, 81, 104

Intensity, 502

Interference term, 213

Interpretation

minimal, 2, 7

Invariant subspace, 133

Involution, 131

Isometric, 18, 111

Isomorphism, 131

J

Joint observable, 262

K

Kantorovich duality, 292

Kernel, 131

Ket, 39

Kraus decomposition, 150

L

Lattice, 23

atomic, 511

distributive, 64

Least upper bound, 20

Lebesgue extension, 67

Lebesgue integral, 66

Lebesgue–Stieltjes measure, 83

Limit of a net, 20

Localisation

approximate, 352

Logic, 509

centre of a, 513

irreducible, 513

of a statistical duality, 510

quadratic, 531

sum, 530

Lower bound, 20

M

Matrix

Kolmogorov decomposition, 178

positive semidefinite, 178

Measurable function, 65

Measurable space, 64

Measure, 65

bi-, 76

determinate, 356

positive operator, 71

regular complex, 85

regular positive Borel, 78

regular positive operator, 85

semispectral, 71

spectral, 71

Toeplitz, 168

Measure separating set, 143

Measure space, 65

Measurement

calibration condition, 228

complete, 240

d-ideal, 249

first kind, 247

- joint, 234
- Lüders, 239
- minimal, 237
- mixed, 235
- nondegenerate, 250
- p-ideal, 249
- pointer function, 227
- probability reproducibility, 226
- reading scale, 227
- repeatable, 247
- scheme, 226
- sequential, 233
- value reproducible, 247
- von Neumann, 239
- Measurement dilation, 161
- Measurement problem, 493
- Minimal, 141, 145, 146
- Monotone class, 74
- Monotone class theorem, 74
- Multiplication, 130
- Multiplicity, 40
- Mutually orthogonal, 27

N

- Naimark representation, 146
- Negligible, 67
- Net, 20
- Noise
 - global measurement, 302
 - measurement, 301
 - overall, 304
- Noise operator, 204
- Noncommutativity
 - degree of, 304
 - total, 347
- Nondegenerate eigenvalue, 44
- Nondegenerate representation, 133
- Norm, 14
- Normal, 18, 133
- Normal linear map, 128
- Normal operator
 - spectral measure, 96
- Normalised, 71
- Normalised positive linear map, 86
- Normed *-algebra, 131
- Normed algebra, 131
- Null set, 67

O

- Objectification problem, 493
- Observable

- coexistent, 265
- compatible, 263
- complementary, 514
- discrete, 202
- functions, 262
- informationally complete, 202
- rank-1 refinement, 240
- sharp, 202
- smearings, 262
- trivial, 203
- unsharp, 202
- Operation, 153, 506
 - dual, 153
 - extremal, 506
 - first kind, 515
 - isotonic, 506
 - Lüders, 211
 - pure, 515
 - repeatable, 516
 - weakly repeatable, 516
- Operator
 - adjoint, 102
 - antiunitary, 210
 - bounded (linear), 17
 - closable, 103
 - closed, 102
 - effect, 196
 - essentially selfadjoint, 103
 - projection, 197
 - selfadjoint, 102
 - symmetric, 102
- Operator bimeasure, 234
- Operator (in $\mathcal{H}$), 101
- Orthocomplementation, 24
- Orthogonal
 - pairwise, 504
- Orthogonal complement, 16
- Orthogonal projection, 17
- Orthogonal sequence
 - of experimental functions, 504
- Orthogonal set or family, 27
- Orthogonality postulate, 499, 508
- Orthonormal basis, 30
- Orthonormal set or family, 27
- Orthosymmetry, 524
- Overall width, 279

P

- Parallelogram law, 15
- Parseval identities, 30
- Partial isometry, 26
- Partial trace, 58

Pauli matrices, 221
 Pauli problem, 408
 Phase shifter, 437
 Photon detector
 ideal, 436
 unsharp, 436
 Point measure, 144
 Point spectrum, 40, 117
 Pointer observable
 mixture condition, 491
 value-definiteness, 491
 Polar decomposition, 27, 59
 Polarisation identity, 15
 Positive bilinear map, 139
 Positive linear map, 85, 129
 Positive operator, 19
 Positive operator bimeasure, 77
 Positive operator measure (POM), 71
 Positive sesquilinear form, 14
 POVM, 71
 Predual, 128, 132
 Probabilistic model, 507
 Probability bimeasure, 234
 Probability measure, 65
 Probability space, 65
 Product σ -algebra, 75
 Projection, 17, 22
 Projection postulate, 499, 516
 Projection theorem, 16
 Projection valued measure, 71
 Proper quantum system, 513
 Property (C), 76
 Property (D), 76
 Proposition
 experimental, 503
 Purification, 159
 Pythagorean theorem, 15

R

Radon transform, 409
 Representations, 133
 Residual spectrum, 40
 Resolution width, 306
 Resolvent set, 40, 116
 Riesz-Markov-Kakutani representation theorem, 85
 Ring, 64

S

Schmidt decomposition, 59
 Schrödinger pair, 347
 Selfadjoint, 18, 133

Selfadjoint operator
 spectral measure, 115, 116
 Semiring, 64
 Semispectral function, 83
 Semispectral measure, 71
 absolutely continuous, 178
 convolution, 171
 covariant, 171
 invariant, 171
 Kolmogorov decomposition, 179
 minimal diagonal Naimark dilation, 179
 minimal diagonalisation, 183
 multiplicity of an outcome, 183
 phase space, 174
 rank- r , 179
 smearing, 169
 Separable, 30
 Separately σ -additive, 76
 Sequence, summable, 267
 Sesquilinear, 14
 Simple function, 66
 Simple tensor, 33
 Simplex, 7
 Smearing, 169
 Space
 $\perp$ -closed subspace, 519
 Hardy sub-, 168
 Hermitian, 519
 Hilbert, 16
 inner product, 13
 Lindelöf, 73
 orthomodular, 520
 Spectral function, 83
 Spectral measure, 71
 canonical, 163
 Spectral representation, 42, 91, 95, 96, 113, 116
 Spectrum, 40, 93, 116
 Square root, 24, 118
 Standard model, 243
 generalised, 243
 State, 192
 mixture, 193
 pure, 193
 separable, 219
 superposition, 194
 vector, 194
 State automorphism, 209
 Statistical duality, 6, 500
 Stinespring representation, 145
 Stinespring type representation, 141
 Stone's theorem, 120
 Strong convergence, 20

Strongly continuous one-parameter unitary group, 120
 Strong (operator) topology, 127
 Subalgebra, 131
 Subrepresentation, 133
 Sufficiency, 510
 Sum of a summable family, 27
 Summable, 27
 Superposition (of pure states), 513
 Support, 115
 of a function, 85
 of a PO measure, 73
 of a state, 510
 projection, 200
 Supremum, 20
 Symmetry (of a logic), 524

T

Tensor product, 33
 Tensor product of linear maps, 135
 Tensor product of vector spaces, 134
 Test
 direct, 288
 indirect, 288
 Theorem
 Gleason's, 199
 Wigner's, 210
 Total variation, 69
 Trace, 48
 Trace class, 47
 Transpose, 128

U

Ultraweak, 127

Uncertainty
 region, 287
 Uncertainty principle, 8, 275
 Uncertainty relation
 measurement, 275
 preparation, 275
 universal, 299
 Unit (element), 131
 Unital, 131, 141, 145, 146
 Unitarily equivalent representations, 133
 Unitary, 18, 133
 Unitary operator
 spectral measure, 96
 Upper bound, 20

V

Value, 276
 definite, 276
 indeterminate, 276
 Von Neumann algebra, 132

W

WAY, 477
 Weak convergence, 20
 Weak (operator) topology, 127
 Weak value, 244
 Weyl pair, 347
 Wigner transform, 409

Y

Yanase condition, 479