Quantum Theory of Open Systems

E. B. DAVIES

St. John's College, Oxford

1976

ACADEMIC PRESS

London · New York · San Francisco

A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Copyright © 1976 by
ACADEMIC PRESS INC. (LONDON) LTD.

All Rights Reserved

No part of this book may be reproduced in any form by photostat, microfilm, or any other means, without written permission from the publishers

Library of Congress Catalog Card Number: 76-016963
ISBN: 0-12-206150-0

FILMSET BY COMPOSITION HOUSE,
SALISBURY, ENGLAND

PRINTED IN GREAT BRITAIN BY
WHITSTABLE LITHO LTD., KENT

Preface

In this book the mathematical foundations of quantum theory are developed in a way which is well adapted to the analysis of open quantum mechanical systems. The relatively recent development of non-equilibrium quantum statistical mechanics as a rigorous mathematical discipline, and continuing research into measurement theory, have led to the realisation that there is considerable advantage in laying the foundations of the subject in a more general and flexible manner than was done by von Neumann in the early days of the subject.

Undoubtedly the new approach owes a great deal to the extraordinary fertility of the field of quantum optics. The experimental investigation of the statistics of coherent photon beams by the measurement of multiparticle correlation functions fits very unnaturally into the framework of orthocomplemented lattices and projection-valued measures. Indeed the measurement of waiting times for the arrival of photons at counters runs completely counter to the philosophy that "ideal" measurements can be considered as taking place at an instant of time.

The novelty of the new approach lies in the much heavier emphasis on mixed states than is usual, and in the full acceptance of irreversible time evolution. This means that there is a partial abandonment of the Hamiltonian formalism, a necessary sacrifice if one is to describe systems which interact with the external world. For the most part the influence of the external world is incorporated into the dynamical equations in a phenomenological manner. In chapters 9 and 10, however, it is shown that the phenomenological evolution equations can be derived from more fundamental Hamiltonian equations for the system plus world.

Although much of the material in the book could have been developed in an axiomatic form I have not done so for two reasons. The first is that many physicists have a profound suspicion of axiomatics and I did not wish them to be prejudiced against the subject from the beginning. The second reason is that any axiomatic approach to quantum theory must deal explicitly with the very difficult philosophical problems of the theory of measurement, which I wished to avoid as far as possible, preferring to concentrate on mathematical model-building. I must here declare my prejudice that much of the discussion of axiomatic measurement theory has been based on too narrow a view of what physicists can measure.

I have also decided to develop the theory at the Hilbert space level. Although there are substantial advantages in using C^*-algebras when studying the dynamics of infinite systems, I hope that this choice will make the material available to a wider audience. The technically more sophisticated reader should be able to generalise the results himself, with the aid of the few starred sections, which require some knowledge of the theory of C^*-algebras.

The text of the book assumes a knowledge of functional analysis at the graduate level together with a general familiarity with the formalism of quantum theory. Much of the material, however, retains its interest even for finite-dimensional Hilbert spaces, and with that restriction would probably be accessible to a beginning graduate student.

In conclusion I should like to thank the large number of people who have contributed to my understanding of the subject, and in particular my teachers and colleagues D. A. Edwards, M. Kac, J. T. Lewis and G. W. Mackey, whose encouragement and interest I have greatly valued. I have also much appreciated Mrs. P. Hawtin's rapid and accurate typing of the manuscript.

January 1976 E.B.D.

Contents

Preface v
Author's Note x

1 Mathematical Background 1

1.1 Bounded operators 1
1.2 Convexity and order 2
1.3 Trace class operators 2
1.4 States 4
1.5 Duality 4
1.6 Convergence 5
1.7 Operator algebras* 6
1.8 Unbounded operators 7
1.9 One parameter semigroups 8
1.10 Tensor products 10
1.11 Measure theory 11
Notes 14

2 Operational Ideas 15

2.1 The repeated Stern–Gerlach experiment 15
2.2 Operations and effects 17
2.3 Pure operations 21
2.4 Irreversible dynamics 26
2.5 A more general formulation* 30
Notes 34

3 Continuous Measurements 35

3.1 Observables 35
3.2 Simultaneous measurement 38
3.3 Approximate position measurements 39
3.4 Observables on phase space 41
3.5 The Wigner distribution 44
Notes 46

4 Operation-Valued Measures 49

4.1 Operations and measures 49
4.2 Composition of instruments 51
4.3 Strongly repeatable instruments 54
4.4 Conditional expectations* 59
4.5 Covariant instruments 60
4.6 Position measurements 66
Notes 68

5 Statistical Properties of Quantum Fields 69

5.1 Measurements in quantum optics 69
5.2 Definition of quantum stochastic processes 70
5.3 Classification of quantum stochastic processes 73
5.4 Measurements on quantum fields 79
5.5 Generators of dynamical semigroups 82
Notes 84

6 The Generalised Wigner–Weisskopf Atom 85

6.1 Beam-foil spectroscopy 85
6.2 Irreducible and simple processes 86
6.3 Recurrence and transience 90
6.4 Transition processes 98
Notes 101

7 Contraction Semigroups on Hilbert Space 103

7.1 Introduction 103
7.2 Dilations of contraction semigroups 105
7.3 Spectrum of the Hamiltonian 108
7.4 Absorption of a particle within a region 112
Notes 116

8 Quantum and Classical Fields 117

8.1 Fock space 117
8.2 Classical states 120
8.3 Evolution of quantum fields with fluctuations 124
8.4 The Weyl algebra* 127
8.5 Generalised coherent states 129
Notes 134

9 Hamiltonian Formalism 135

9.1 Fundamental considerations 135
9.2 Complete positivity* 136
9.3 Projection-valued measures 141
9.4 Finite-dimensional dynamical semigroups 143
Notes 147

10 General Theory of Open Systems 149

10.1 Systems interacting with reservoirs 149
10.2 A second example 150
10.3 Stochastic evolution equations 151
10.4 Derivation of the master equation 153
10.5 Solutions of the master equation 154
Notes 156

Appendix 158

References 161

Index 169

Author's Note

Some explanation of my conventions may be helpful. A reference to Eq. (A.B) in the text means the equation of that number in the same chapter, while Eq. (A.B.C) means Eq. (B.C) of chapter A. The same convention applies to lemmas and theorems, whose numbers are not bracketed. In the definition of a new term any word enclosed in brackets is to be understood as applying to any later use of the term.

Most of the mathematical material necessary for an understanding of the book is collected into chapter 1, which also serves to fix the notation used. A few technical results of a more advanced nature are, however, collected into an appendix at the end. The starred sections require a more sophisticated mathematical approach and may be omitted without loss of continuity. Each chapter ends with notes and references to the literature, which in several cases is so extensive that I am painfully aware of having hardly scratched the surface.

1

Mathematical Background

1.1 Bounded operators

We shall use the symbol $\mathscr{H}$ to denote a complex Hilbert space with an inner product $\langle \phi, \psi \rangle$ which is complex linear in ϕ and conjugate linear in ψ. If S is any set in $\mathscr{H}$ then $\text{lin}(S)$ is the smallest linear subspace containing S and $\overline{\text{lin}}(S)$ is the smallest closed linear subspace containing S. The set $\mathscr{L}(\mathscr{H})$ of all bounded linear operators on $\mathscr{H}$ is an algebra with an involution * (adjoint operation) satisfying

$$(AB)^* = B^*A^*, \qquad A^{**} = A \tag{1.1}$$

and

$$\|A^*A\| = \|A\|^2 = \|A^*\|^2 \tag{1.2}$$

for all, $A, B \in \mathscr{L}(\mathscr{H})$. An operator $A \in \mathscr{L}(\mathscr{H})$ is called compact if $\{A\phi_n\}$ has a norm convergent subsequence for every norm bounded sequence $\{\phi_n\}$ in $\mathscr{H}$. The set $\mathscr{C}(\mathscr{H})$ of all compact operators on $\mathscr{H}$ is a norm closed *-subalgebra and also an ideal in $\mathscr{L}(\mathscr{H})$. The subalgebra $\mathscr{C}(\mathscr{H})$ coincides with $\mathscr{L}(\mathscr{H})$ if and only if $\mathscr{H}$ is finite-dimensional. An operator on $\mathscr{H}$ is said to be degenerate or of finite rank if its range is a finite-dimensional subspace of $\mathscr{H}$. The set of all degenerate operators is norm dense in $\mathscr{C}(\mathscr{H})$.

An operator A in $\mathscr{L}(\mathscr{H})$ is said to be self-adjoint if $A = A^*$, unitary if $AA^* = A^*A = 1$, and to be a projection if $A^2 = A = A^*$. Every bounded operator A has a unique decomposition $A = B + iC$ where B and C are self-adjoint. The spectrum $\text{sp}(A)$ of A is the set of complex numbers z such that $(A - z1)$ does not have an inverse in the algebra $\mathscr{L}(\mathscr{H})$. It is a non-empty closed bounded subset of the complex plane $\mathbb{C}$. If A is self-adjoint the spectrum is contained in the real line $\mathbb{R}$. If A is compact the spectrum consists of a sequence of points in the complex plane which converges to the origin, or else is a finite set. The support of A is defined as the orthogonal complement $L^\perp$ of its null space L. If A is self-adjoint its support equals the closure of its range.

If A is self-adjoint we define $A \geqslant 0$ to mean $\text{sp}(A) \subseteq [0, \infty)$ or equivalently $\langle A\phi, \phi \rangle \geqslant 0$ for all $\phi \in \mathscr{H}$. We define $A \geqslant B$ to mean $(A - B) \geqslant 0$. Every

$A \geqslant 0$ has a unique positive square root which we denote by $A^{1/2}$. If $A \in \mathscr{L}(\mathscr{H})$ we put $|A| = (A^*A)^{1/2}$. If $A = A^*$ we define the positive and negative parts A^+ and A^- of A to be

$$A^+ = \frac{|A| + A}{2}, \qquad A^- = \frac{|A| - A}{2}. \tag{1.3}$$

1.2 Convexity and order

If S is a set which is (partially) ordered by a relation $\leqslant$ which is reflexive, transitive and antisymmetric, we say that S is a lattice if any two elements x and y of S have a least upper bound $x \vee y$ and a greatest lower bound $x \wedge y$ in S. We say that S is a complete lattice if every subset of S has a least upper bound and a greatest lower bound in S.

A convex set in a vector space V is defined as a set K such that if $\phi_1, \phi_2 \in K$ and $0 < \lambda < 1$ then

$$\lambda\phi_1 + (1 - \lambda)\phi_2 \in K. \tag{2.1}$$

An extreme point of K is a point $\phi \in K$ such that if $\phi_1, \phi_2 \in K$ and $0 < \lambda < 1$ and

$$\phi = \lambda\phi_1 + (1 - \lambda)\phi_2 \tag{2.2}$$

then $\phi = \phi_1 = \phi_2$. A cone V^+ in V is a non-empty set such that if ϕ_1, $\phi_2 \in V^+$ and $\lambda_1, \lambda_2 \geqslant 0$ then

$$\lambda_1\phi_1 + \lambda_2\phi_2 \in V^+ \tag{2.3}$$

and if both $\pm\phi$ lie in V^+ then $\phi = 0$. A cone induces an ordering on V by the definition $\phi_1 \geqslant \phi_2$ if $(\phi_1 - \phi_2) \in V^+$. A pure element of V^+ is a $\phi \in V^+$ such that if $0 \leqslant \psi \leqslant \phi$ then $\psi = \lambda\phi$ for some $0 \leqslant \lambda \leqslant 1$. Equivalently $\phi \in V^+$ is pure if and only if when $\phi_1, \phi_2 \in V^+$ and $\phi = \phi_1 + \phi_2$ there exist λ_1 and λ_2 in $\mathbb{R}$ such that $\phi_1 = \lambda_1\phi$ and $\phi_2 = \lambda_2\phi$.

We define a positive linear functional on an ordered vector space V to be a linear functional ϕ such that if $x \in V^+$ then $\phi(x) \geqslant 0$. More generally we define a positive linear map $T: V_1 \to V_2$ between two ordered vector spaces to be a linear map such that if $x \in V_1^+$ then $T(x) \in V_2^+$. We denote by $\mathscr{L}^+(V_1, V_2)$ the set of all positive linear maps. It is a cone provided $V_1 = V_1^+ - V_1^+$.

We denote by $\mathbb{R}^+$ the set of non-negative real numbers, by $\mathbb{Z}$ the set of integers and by $\mathbb{Z}^+$ the set of integers $\geqslant 0$.

1.3 Trace class operators

If A is a compact operator on $\mathscr{H}$ then $|A|$ is also compact, with non-negative eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ which we arrange in decreasing order and repeat according to their multiplicity. We say A is of trace class if it has finite trace norm, where

$$\|A\|_{tr} = \sum_{n=1}^{\infty} \lambda_n. \tag{3.1}$$

The set $\mathscr{T}(\mathscr{H})$ of all trace class operators on $\mathscr{H}$ is a linear space and $\|\cdot\|_{tr}$ is indeed a norm on $\mathscr{T}(\mathscr{H})$ with respect to which it is a Banach space. Moreover $\|A\| \leqslant \|A\|_{tr}$ for all $A \in \mathscr{T}(\mathscr{H})$. If $\{\phi_n\}$ is an orthonormal basis in $\mathscr{H}$ and $A \in \mathscr{T}(\mathscr{H})$ the sum

$$\operatorname{tr}[A] = \sum_n \langle A\phi_n, \phi_n \rangle \tag{3.2}$$

is absolutely convergent and independent of the orthonormal basis chosen. It defines a linear functional $\operatorname{tr}: \mathscr{T}(\mathscr{H}) \to \mathbb{C}$ which is norm bounded, satisfying

$$|\operatorname{tr}[A]| \leqslant \|A\|_{tr} \tag{3.3}$$

for all $A \in \mathscr{T}(\mathscr{H})$. If A is compact and self-adjoint with eigenvalues $\{\lambda_n\}$ repeated according to multiplicity then

$$\|A\|_{tr} = \sum_n |\lambda_n| \tag{3.4}$$

and if this sum is finite then

$$\operatorname{tr}[A] = \sum_n \lambda_n. \tag{3.5}$$

The space of trace class operators contains all degenerate operators and is therefore dense in $\mathscr{C}(\mathscr{H})$. We denote by $\mathscr{T}_s(\mathscr{H})$ the Banach space of self-adjoint trace class operators with the trace norm and by $\mathscr{T}_s(\mathscr{H})^+$ the cone of non-negative trace class operators.

If $A \in \mathscr{T}(\mathscr{H})$ then A^* and $|A|$ lie in $\mathscr{T}(\mathscr{H})$. If $A \in \mathscr{T}_s(\mathscr{H})^+$ then

$$\operatorname{tr}[A] = \|A\|_{tr}. \tag{3.6}$$

If $A \in \mathscr{T}_s(\mathscr{H})$ then A^+ and A^- lie in $\mathscr{T}_s(\mathscr{H})^+$ and

$$\|A\|_{tr} = \|A^+\|_{tr} + \|A^-\|_{tr}. \tag{3.7}$$

It is easily deducible from Eqs (3.2) and (3.6) that if A_n is an increasing sequence of operators in $\mathscr{T}_s(\mathscr{H})^+$, and if $A \in \mathscr{T}_s(\mathscr{H})^+$ satisfies

$$\lim_{n\to\infty} \langle A_n\phi, \phi \rangle = \langle A\phi, \phi \rangle \tag{3.8}$$

for all $\phi \in \mathscr{H}$, then

$$\lim_{n\to\infty} \|A - A_n\|_{tr} < \infty. \tag{3.9}$$

1.4 States

The state space V of a Hilbert space $\mathscr{H}$ is defined as the Banach space $\mathscr{T}_s(\mathscr{H})$ with the trace norm $\|\cdot\|_{\text{tr}}$ and the cone $V^+ = \mathscr{T}_s(\mathscr{H})^+$. The states are defined as the non-negative trace class operators of trace one, elsewhere called mixed states or density matrices. The states form a convex set whose extreme points, the pure states, are the operators of the form

$$|\psi\rangle\langle\psi| : \phi \to \langle\phi, \psi\rangle\psi \tag{4.1}$$

where $\psi \in \mathscr{H}$ and $\|\psi\| = 1$. Two unit vectors ψ_1 and ψ_2 define the same pure state if and only if $\psi_2 = e^{i\theta}\psi_1$ for some $\theta \in \mathbb{R}$. By the spectral theorem every state has a representation

$$\rho = \sum_{n=1}^{\infty} \lambda_n |\psi_n\rangle\langle\psi_n| \tag{4.2}$$

where $\|\psi_n\| = 1$, $\lambda_n \geqslant 0$ for all n and

$$\sum_{n=1}^{\infty} \lambda_n = 1. \tag{4.3}$$

We shall occasionally use the word "state" to denote a unit vector in Hilbert space when established usage demands this.

1.5 Duality

If $X \in \mathscr{T}(\mathscr{H})$ and $Y \in \mathscr{L}(\mathscr{H})$ then XY and YX lie in $\mathscr{T}(\mathscr{H})$ with

$$\|XY\|_{\text{tr}} \leqslant \|X\|_{\text{tr}}\|Y\| \tag{5.1}$$

$$\|YX\|_{\text{tr}} \leqslant \|Y\|\,\|X\|_{\text{tr}} \tag{5.2}$$

and

$$\text{tr}[XY] = \text{tr}[YX]. \tag{5.3}$$

Thus every $Y \in \mathscr{L}(\mathscr{H})$ defines a bounded linear functional ϕ_Y on $\mathscr{T}(\mathscr{H})$ by

$$\phi_Y(X) = \text{tr}[XY]. \tag{5.4}$$

The map $\phi: \mathscr{L}(\mathscr{H}) \to \mathscr{T}(\mathscr{H})^*$ is a linear isometric isomorphism of the two spaces, by which we identify $\mathscr{L}(\mathscr{H})$ with $\mathscr{T}(\mathscr{H})^*$. In the same way the formula

$$\psi_X(Y) = \text{tr}[YX] \tag{5.5}$$

where $X \in \mathscr{T}(\mathscr{H})$ and $Y \in \mathscr{C}(\mathscr{H})$, defines a linear isometric isomorphism ψ of $\mathscr{T}(\mathscr{H})$ onto $\mathscr{C}(\mathscr{H})^*$ which we use to identify these two spaces.

The above duality relations are also valid between $\mathscr{C}_s(\mathscr{H})$, $\mathscr{T}_s(\mathscr{H})$ and $\mathscr{L}_s(\mathscr{H})$. Moreover the orderings of $\mathscr{T}_s(\mathscr{H})$ and $\mathscr{L}_s(\mathscr{H})$ are compatible in the following sense. $X \in \mathscr{T}_s(\mathscr{H})^+$ if and only if $\mathrm{tr}[XY] \geqslant 0$ for all $Y \in \mathscr{L}_s(\mathscr{H})^+$. Conversely $Y \in \mathscr{L}_s(\mathscr{H})^+$ if and only if $\mathrm{tr}[XY] \geqslant 0$ for all $X \in \mathscr{T}_s(\mathscr{H})^+$.

1.6 Convergence

We shall use several different notions of convergence in $\mathscr{L}(\mathscr{H})$. We say A_n converges in norm to A, $A_n \xrightarrow{n} A$, if

$$\lim_{n\to\infty} \|A_n - A\| = 0. \tag{6.1}$$

We say A_n converges strongly to A, $A_n \xrightarrow{s} A$, if

$$\lim_{n\to\infty} \|A_n\phi - A\phi\| = 0 \tag{6.2}$$

for all $\phi \in \mathscr{H}$. We say A_n converges weakly to A, $A_n \xrightarrow{w} A$, if

$$\lim_{n\to\infty} \langle A_n\phi, \psi\rangle = \langle A\phi, \psi\rangle \tag{6.3}$$

for all ϕ, $\psi \in \mathscr{H}$. Norm convergence implies strong convergence, which implies weak convergence. We say A_n converges ultraweakly to A, $A_n \xrightarrow{uw} A$, if

$$\lim_{n\to\infty} \mathrm{tr}[A_n\rho] = \mathrm{tr}[A\rho] \tag{6.4}$$

for all $\rho \in \mathscr{T}(\mathscr{H})$. Ultraweak convergence implies weak convergence, but on a norm-bounded set of operators the two notions coincide. If $\{A_n\}$ is a norm-bounded increasing sequence of self-adjoint operators on $\mathscr{H}$, then A_n converges weakly, strongly and ultraweakly to an operator A which is also the least upper bound of the sequence in the sense that $A \geqslant A_n$ for all n and if $B \geqslant A_n$ for all n then $B \geqslant A$. In this situation we write $A_n \uparrow A$.

The above definitions for sequences apply without modification to generalised sequences (nets), where the index n runs over a general directed set. Generalised sequences are not usually necessary except when considering non-separable Hilbert spaces, which are rarely of physical interest. Because of this we shall not often refer explicitly to generalised sequences, being confident that the reader who is really interested in these questions will know when they are needed.

The convergence of products of sequences of operators has to be treated carefully. If $A_m \xrightarrow{s} A$ and $B_m \xrightarrow{s} B$ then $A_mB_m \xrightarrow{s} AB$. If $A_m \xrightarrow{w} A$ and $B_m \xrightarrow{s} B$ then $A_mB_m \xrightarrow{w} AB$, but the same is not true if the order of multiplication is reversed. If $A_m \xrightarrow{w} A$ then $A_mB \xrightarrow{w} AB$ and $BA_m \xrightarrow{w} BA$ for all B. If $A_m \xrightarrow{w} A$

then $A_m^* \xrightarrow{w} A^*$ but if $A_m \xrightarrow{s} A$ one can only conclude that $A_m^* \xrightarrow{w} A^*$. If $A \in \mathcal{T}(\mathcal{H})$ and $B_m \xrightarrow{s} B$ then Eq. (4.2) may be used to prove that $B_m A$ converges to BA in trace norm. It follows that if $A_m \to A$ in trace norm and $B_m \xrightarrow{s} B$ then $B_m A_m \to BA$ in trace norm.

The above results also hold for generalised sequences if one assumes that the operators are uniformly bounded in norm, a property which is automatic for sequences by the uniform boundedness theorem.

We say that a positive linear functional $\phi: \mathcal{L}_s(\mathcal{H}) \to \mathbb{R}$ is normal if $A_n \uparrow A$ implies that $\phi(A_n)$ converges to $\phi(A)$.

Lemma 6.1 *A positive linear functional ϕ on $\mathcal{L}_s(\mathcal{H})$ is normal if and only if there exists $\sigma \in \mathcal{T}_s(\mathcal{H})^+$ such that*

$$\phi(A) = \mathrm{tr}[A\sigma] \tag{6.5}$$

for all $A \in \mathcal{L}_s(\mathcal{H})$.

Proof Suppose first that ϕ is given by Eq. (6.5). If $A_n \uparrow A$ then $A_n \xrightarrow{s} A$ so $A_n\sigma$ converges in trace norm to $A\sigma$ and $\phi(A_n) \to \phi(A)$. Therefore ϕ is normal.

Conversely suppose ϕ is positive and normal. If $A \in \mathcal{L}_s(\mathcal{H})$ then

$$-\|A\|1 \leqslant A \leqslant \|A\|1 \tag{6.6}$$

so

$$-\|A\|\phi(1) \leqslant \phi(A) \leqslant \|A\|\phi(1) \tag{6.7}$$

and the restriction of ϕ to $\mathcal{C}_s(\mathcal{H})$ is a bounded positive linear functional. Therefore there exists $\sigma \in \mathcal{T}_s(\mathcal{H})^+$ such that Eq. (6.5) holds for all $A \in \mathcal{C}_s(\mathcal{H})$. If P_n are projections of finite rank such that $P_n \uparrow 1$ and $A \in \mathcal{L}_s(\mathcal{H})^+$ then $A^{1/2}P_nA^{1/2} \uparrow A$ so

$$\begin{aligned}\phi(A) &= \lim_{n\to\infty} \phi(A^{1/2}P_nA^{1/2}) \\ &= \lim_{n\to\infty} \mathrm{tr}[A^{1/2}P_nA^{1/2}\sigma] \\ &= \mathrm{tr}[A\sigma].\end{aligned} \tag{6.8}$$

Since Eq. (6.5) holds for all $A \in \mathcal{L}_s(\mathcal{H})^+$ it also holds for all $A \in \mathcal{L}_s(\mathcal{H})$.

QED

1.7 Operator algebras*

We mention the general theory of C^*-algebras only in a few starred sections which, except in chapter 9, may be omitted without loss of continuity. We shall in any case need only a few introductory ideas from the theory.

A C^*-algebra with identity is a Banach space $\mathscr{A}$ which is also an algebra with identity and an involution satisfying

$$\|AB\| \leqslant \|A\|\,\|B\| \tag{7.1}$$

and

$$\|A^*A\| = \|A\|^2 = \|A^*\|^2 \tag{7.2}$$

for all $A, B \in \mathscr{A}$. The spectrum $\mathrm{sp}(A)$ of an element $A \in \mathscr{A}$ is the set of $z \in \mathbb{C}$ such that $(A - z1)$ does not have an inverse in $\mathscr{A}$. An element $A \in \mathscr{A}$ can be put in the form $A = B^*B$ for some $B \in \mathscr{A}$ if and only if $A = A^*$ and $\mathrm{sp}(A) \subseteq \mathbb{R}^+$. The set of all such elements forms a norm closed cone $\mathscr{A}^+$ in $\mathscr{A}$.

If $\mathscr{H}$ is a Hilbert space then $\mathscr{L}(\mathscr{H})$ is a C^*-algebra. A $*$-subalgebra $\mathscr{A}$ of $\mathscr{L}(\mathscr{H})$ containing the identity operator and closed in the weak operator topology is called a von Neumann algebra.

If $\mathscr{A}$ is a C^*-algebra with identity, a representation of $\mathscr{A}$ on a Hilbert space $\mathscr{H}$ is defined as a homomorphism $\pi: \mathscr{A} \to \mathscr{L}(\mathscr{H})$ such that $\pi(A^*) = \pi(A)^*$ for all $A \in \mathscr{A}$ and $\pi(1) = 1$. If π is a representation then $\|\pi(A)\| \leqslant \|A\|$ for all $A \in \mathscr{A}$, and π is a positive linear map. A state ϕ on $\mathscr{A}$ is a linear functional $\phi: \mathscr{A} \to \mathbb{C}$ such that $\phi(X^*) = \overline{\phi(X)}$ for all $X \in \mathscr{A}$, $\phi(X^*X) \geqslant 0$ for all $X \in \mathscr{A}$ and $\phi(1) = 1$. The Gelfand–Naimark–Segal (GNS) theorem states that if ϕ is a state on a C^*-algebra $\mathscr{A}$ with identity then there is a representation π of $\mathscr{A}$ on a Hilbert space $\mathscr{H}$ and a unit vector $\psi \in \mathscr{H}$ such that

$$\phi(X) = \langle \pi(X)\psi, \psi \rangle \tag{7.3}$$

for all $X \in \mathscr{A}$. Moreover the representation is unique up to isomorphism if it is cyclic in the sense that

$$\{\pi(X)\psi : X \in \mathscr{A}\} \tag{7.4}$$

is a dense linear subspace of $\mathscr{H}$.

1.8 Unbounded operators

A (densely defined) unbounded operator A on a Banach space $\mathscr{B}$ is a linear map $A: \mathscr{D} \to \mathscr{B}$, where $\mathscr{D} = \mathrm{dom}(A)$ is a dense linear subspace of $\mathscr{B}$. A is said to be closed if whenever $x_m \in \mathscr{D}$, $x_m \xrightarrow{n} x$ and $Ax_m \xrightarrow{n} y$ it follows that $x \in \mathscr{D}$ and $Ax = y$. An operator is said to be closable if it has a closed extension, in which case its closure $\bar{A}$ is the smallest closed extension. Then $x \in \mathrm{dom}(\bar{A})$ if and only if there is a sequence $x_m \in \mathscr{D}$ such that $x_m \xrightarrow{n} x$ and Ax_m converges, in which case $\bar{A}x = \lim_{m\to\infty} Ax_m$.

If A is an unbounded operator on $\mathscr{B}$ we say $z \in \mathbb{C}$ does not lie in the spectrum $\mathrm{sp}(A)$ of A if $(A - z1)$ is one-one with range equal to $\mathscr{B}$, and the inverse operator $(A - z1)^{-1}$ is bounded. If A is closed the last condition is unnecessary by the closed graph theorem. If $z \notin \mathrm{sp}(A)$ and $\|(A - z1)^{-1}\| = c$ and $|w - z| < c$ then $w \notin \mathrm{sp}(A)$ and

$$(A - w1)^{-1} = \sum_{n=0}^{\infty} (w - z)^n (A - z1)^{-n-1} \tag{8.1}$$

the series being norm convergent. Therefore $\mathrm{sp}(A)$ is a closed set.

An unbounded operator A on a Hilbert space $\mathscr{H}$ has an adjoint A^* with

$$\begin{aligned}\mathrm{dom}(A^*) = \{y \in \mathscr{H} : x \to \langle Ax, y\rangle \text{ is a norm bounded}\\ \text{linear functional on } \mathrm{dom}(A)\}.\end{aligned} \tag{8.2}$$

If $y \in \mathrm{dom}(A^*)$ then A^*y is characterised by

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \tag{8.3}$$

valid for all $x \in \mathrm{dom}(A)$. A^* is densely defined if and only if A is closable.

The operator A is called symmetric if

$$\langle Ax, y\rangle = \langle x, Ay\rangle \tag{8.4}$$

for all $x, y \in \mathrm{dom}(A)$. Every symmetric operator is closable and its closure is still symmetric. A is called self-adjoint if $A = A^*$, and essentially self-adjoint if $\bar{A}$ is self-adjoint. If A is symmetric we say that $A \geqslant 0$ if $\langle A\phi, \phi\rangle \geqslant 0$ for all $\phi \in \mathrm{dom}(A)$. If A is self-adjoint then $\mathrm{sp}(A) \subseteq \mathbb{R}$, and $\mathrm{sp}(A) \subseteq \mathbb{R}^+$ if and only if $A \geqslant 0$.

We assume that the reader knows the spectral theorem and that an operator H on $\mathscr{H}$ is self-adjoint if and only if iH is the generator of a strongly continuous one parameter unitary group on $\mathscr{H}$.

1.9 One parameter semigroups

If $\mathscr{B}$ is a Banach space a one-parameter semigroup on $\mathscr{B}$ is a family $\{T_t\}$ of bounded linear operators on $\mathscr{B}$ parametrised by $t \in \mathbb{R}^+$ and satisfying $T_0 = 1$ and $T_s T_t = T_{s+t}$ for all $s, t \geqslant 0$. The semigroup is called strongly continuous if

$$\lim_{t \to 0} \|T_t f - f\| = 0 \tag{9.1}$$

for all $f \in \mathscr{B}$, in which case the function $t \to T_t f$ is continuous for all $t \geqslant 0$ and all $f \in \mathscr{B}$. The infinitesimal generator Z of the semigroup is defined by

$$Zf = \lim_{t \to 0} t^{-1}\{T_t f - f\} \tag{9.2}$$

the domain $\mathrm{dom}(Z)$ of Z being the set of all $f \in \mathscr{B}$ for which the limit exists. If T_t is a strongly continuous one-parameter semigroup on $\mathscr{B}$ then Z is densely defined and is a closed linear operator. If $f \in \mathrm{dom}(Z)$ then $f(t) = T_t f$ lies in $\mathrm{dom}(Z)$ for all $t \geqslant 0$ and

$$\frac{df(t)}{dt} = Zf(t) \tag{9.3}$$

for all $t \geqslant 0$. The semigroup is determined by its generator through the formula

$$T_t f = \lim_{n \to \infty} (1 - tn^{-1}Z)^{-n} f \tag{9.4}$$

for all $t \geqslant 0$ and $f \in \mathscr{B}$, the inverse being a bounded operator for all sufficiently large n. We write $T_t = e^{Zt}$ although e^{Zt} is just a formal expression except when Z is bounded, in which case

$$e^{Zt} = \sum_{n=0}^{\infty} \frac{Z^n t^n}{n!}. \tag{9.5}$$

Theorem 9.1 *If Z is the generator of a strongly continuous one-parameter semigroup S_t on a Banach space $\mathscr{B}$ and A is a bounded operator then $(Z + A)$ is the generator of a strongly continuous one-parameter semigroup T_t on $\mathscr{B}$ related to S_t by*

$$T_t = S_t + \int_{s=0}^{t} T_{t-s} A S_s \, ds \tag{9.6}$$

for all $t \geqslant 0$, where the integral is strongly convergent.

Proof Equation (9.6) uniquely determines T_t, subject to its being uniformly norm bounded on finite intervals in $\mathbb{R}^+$, for by iteration we obtain the norm-convergent expansion

$$\begin{aligned} T_t = S_t &+ \int_{s=0}^{t} S_{t-s} A S_s \, ds \\ &+ \int_{s=0}^{t} \int_{u=0}^{s} S_{t-s} A S_{s-u} A S_u \, du \, ds \\ &+ \int_{s=0}^{t} \int_{u=0}^{s} \int_{v=0}^{u} S_{t-s} A S_{s-u} A S_{u-v} A S_v \, dv \, du \, ds + \dots \end{aligned} \tag{9.7}$$

from which it follows that T_t is a semigroup. Moreover if $f \in \mathscr{B}$ then by Eq. (9.7)

$$T_t f - f = (S_t f - f) + O(t) \tag{9.8}$$

as $t \to 0$, so T_t is strongly continuous at zero. Also

$$t^{-1}(T_t f - f) = t^{-1}(S_t f - f) + t^{-1} \int_0^t S_{t-s} A S_s f \, ds + O(t) \tag{9.9}$$

as $t \to 0$ and

$$\lim_{t \to 0} t^{-1} \int_0^t S_{t-s} A S_s f \, ds = Af \tag{9.10}$$

for all $f \in \mathscr{B}$. Therefore the domains of the infinitesimal generators of T_t and S_t are equal, and that of T_t is $(Z + A)$. Thus T_t is a strongly continuous semigroup with the stated generator, and by Eq. (9.4) the only one. QED

1.10 Tensor products

If $\mathscr{H}_1$ and $\mathscr{H}_2$ are two vector spaces their algebraic tensor product $\mathscr{H}_1 \odot \mathscr{H}_2$ has the following characteristic property. For every bilinear map B from $\mathscr{H}_1 \times \mathscr{H}_2$ into a vector space $\mathscr{H}$ there exists a unique linear map $B': \mathscr{H}_1 \odot \mathscr{H}_2 \to \mathscr{H}$ such that

$$B(f, g) = B'(f \otimes g) \tag{10.1}$$

for all $f \in \mathscr{H}_1$ and $g \in \mathscr{H}_2$. If $\mathscr{H}_1$ and $\mathscr{H}_2$ are Hilbert spaces then the formula

$$\left\langle \sum_{i=1}^{m} \alpha_i f_i \otimes g_i, \sum_{j=1}^{n} \beta_j f'_j \otimes g'_j \right\rangle = \sum_{i,j} \alpha_i \bar{\beta}_j \langle f_i, f'_j \rangle \langle g_i, g'_j \rangle \tag{10.2}$$

defines an inner product on $\mathscr{H}_1 \odot \mathscr{H}_2$. The completion of $\mathscr{H}_1 \odot \mathscr{H}_2$ for the associated norm is called the (Hilbert space) tensor product of $\mathscr{H}_1$ and $\mathscr{H}_2$ and is denoted $\mathscr{H}_1 \otimes \mathscr{H}_2$. If $\{f_i\}$ and $\{g_j\}$ are orthonormal bases of $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively then $\{f_i \otimes g_j\}$ is an orthonormal basis of $\mathscr{H}_1 \otimes \mathscr{H}_2$. The Hilbert space tensor product has the following characteristic property with respect to bilinear maps. If B is bilinear from $\mathscr{H}_1 \times \mathscr{H}_2$ into the Hilbert space $\mathscr{H}$ and

$$\langle B(f, g), B(f', g') \rangle = \langle f, f' \rangle \langle g, g' \rangle \tag{10.3}$$

for all $f, f' \in \mathscr{H}_1$, and $g, g' \in \mathscr{H}_2$, then there exists a unique isometric linear map $B': \mathscr{H}_1 \otimes \mathscr{H}_2 \to \mathscr{H}$ such that

$$B(f, g) = B'(f \otimes g) \tag{10.4}$$

for all $f \in \mathscr{H}_1$, and $g \in \mathscr{H}_2$. If A_1 and A_2 are bounded operators on the Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively there is a unique bounded operator, denoted by $A_1 \otimes A_2$, on $\mathscr{H}_1 \otimes \mathscr{H}_2$ such that

$$(A_1 \otimes A_2)(f \otimes g) = (A_1 f) \otimes (A_2 g) \tag{10.5}$$

for all $f \in \mathscr{H}_1$ and $g \in \mathscr{H}_2$. The map $A \to A \otimes 1$ is an isometric embedding of $\mathscr{L}(\mathscr{H}_1)$ as a subalgebra of $\mathscr{L}(\mathscr{H}_1 \otimes \mathscr{H}_2)$.

If $l^2(N, \mathscr{H})$ is the Hilbert space of square summable $\mathscr{H}$-valued functions on the set N we may identify $l^2(N) \otimes \mathscr{H}$ with $l^2(N, \mathscr{H})$ as follows. The map

$$B: l^2(N) \times \mathscr{H} \to l^2(N, \mathscr{H}) \tag{10.6}$$

defined by

$$\{B(a, \psi)\}(n) = a(n)\psi \tag{10.7}$$

is bilinear and satisfies Eq. (10.3). Therefore it may be extended to a unitary map

$$B': l^2(N) \otimes \mathscr{H} \to l^2(N, \mathscr{H}) \tag{10.8}$$

such that

$$\left\{B'\left(\sum_{r=1}^{R} a_r \otimes \psi_r\right)\right\}(n) = \sum_{r=1}^{R} a_r(n)\psi_r. \tag{10.9}$$

One may similarly identify $L^2(X) \otimes \mathscr{H}$ with the space of square integrable $\mathscr{H}$-valued functions on X.

1.11 Measure theory

A σ-field $\mathscr{F}$ on a set X is a non-empty collection of subsets of X closed under complements, countable unions and countable intersections. Every σ-field contains the empty set $\varnothing$ and the whole space X as members. If X is any topological space then $C(X)$ denotes the space of all complex-valued continuous functions on X and $C_{\mathbb{R}}(X)$ the space of all real-valued continuous functions. The Baire σ-field of X is the smallest σ-field with respect to which all continuous functions are measurable, while the Borel σ-field is the smallest σ-field containing all open sets. In general the Baire σ-field is contained in the Borel σ-field. If X is a compact Hausdorff space there is a natural one-one correspondence between

(*i*) the positive linear functionals on $C_{\mathbb{R}}(X)$; (11.1)

(*ii*) the Baire measures on X; (11.2)

(*iii*) the Borel measures on X which are regular in the sense that for all Borel sets E

$$\mu(E) = \sup\{\mu(K): K \subseteq E \text{ and } K \text{ is compact}\}. \tag{11.3}$$

If X is a compact Hausdorff space the following three conditions are equivalent:

(*i*) X is second countable (there is a countable base to the topology of X); (11.4) (*ii*) X is metrisable; (11.5)

(*iii*) $C_{\mathbb{R}}(X)$ is separable (contains a countable dense set). (11.6)

If these conditions are satisfied then the Baire and Borel σ-fields coincide, and every Borel measure on X is regular. The support of a measure μ on X is then defined as the complement of the largest open set U in X such that $\mu(U) = 0$.

In quantum-mechanical measurement theory it is almost always the case that a measurable quantity takes its values in a locally compact Hausdorff space X which is second countable, for example an open or closed set in $\mathbb{R}^n$. The one-point compactification of X is a compact metrisable space $\bar{X}$ obtained by adjoining one point, called ∞, to X. There is a one-one correspondence between the (finite) Borel measures on X and the Borel measures μ on $\bar{X}$ such that $\mu(\infty) = 0$. Because of this there is little loss of generality in supposing that every measurable quantity takes its values in a compact metrisable space.

From the mathematical point of view almost every theorem in the book can be extended from compact metrisable spaces to compact Hausdorff spaces by replacing the word "Borel" by "regular Borel" or "Baire". However, it is not known whether several of the critical theorems can be extended to an arbitrary set X with a σ-field $\mathscr{F}$, and some, for example the following one, cannot be.

Theorem 11.1 *Let $\{\Omega_n\}_{n=1}^{\infty}$ be locally compact second countable topological spaces and let μ_n be measures on $\prod_{r=1}^{n} \Omega_r$ which are compatible in the sense that if $n > m$*

$$\mu_m(E) = \mu_n\left(E \times \prod_{r=m+1}^{n} \Omega_r\right) \tag{11.7}$$

for all Borel sets E in $\prod_{r=1}^{m} \Omega_r$. Then there is a measure μ on $\prod_{n=1}^{\infty} \Omega_n$ which is compatible with all the measures μ_n.

Proof If $\bar{\Omega}_n$ is the one-point compactification of Ω_n then $\prod_{n=1}^{\infty} \Omega_n$ is a Borel set in $\prod_{n=1}^{\infty} \bar{\Omega}_n$. It is therefore sufficient to consider the case where Ω_n are all compact metrisable spaces. The measures correspond to positive linear functionals ϕ_n on the algebras $\mathscr{A}_n = C_{\mathbb{R}}(\prod_{s=1}^{n} \Omega_s)$. We can regard $\mathscr{A}_n$ as a subalgebra of $\mathscr{A} = C_{\mathbb{R}}(\prod_{r=1}^{\infty} \Omega_r)$ by identifying functions on $\prod_{s=1}^{n} \Omega_r$ with functions on $\prod_{r=1}^{\infty} \Omega_r$ which do not depend on any coordinates greater than n. Then $\mathscr{A}_n \subseteq \mathscr{A}_{n+1}$ and the algebra $\mathscr{B} = \bigcup_{n=1}^{\infty} \mathscr{A}_n$ is dense in $\mathscr{A}$ by the Stone Weierstrass theorem. ϕ_{n+1} is an extension of ϕ_n and the common extension to ψ on $\mathscr{B}$ is a positive linear functional with norm

$$\|\psi\| = \mu_n\left(\prod_{s=1}^{n} \Omega_r\right) \tag{11.8}$$

for all n. ψ can therefore be uniquely extended to a positive bounded linear functional ϕ on $\mathscr{A}$, and this functional corresponds to a measure μ on $\prod_{r=1}^{\infty} \Omega_r$ which has the required properties. QED

If μ is a measure on a compact metrisable space Ω we say that μ is discrete if there is a countable set $E \subseteq \Omega$ such that $\mu(\Omega \setminus E) = 0$. We say μ is continuous if for every set $E \subseteq \Omega$ such that $\mu(E) \neq 0$ there exists a set F with $F \subseteq E$ and

$$0 < \mu(F) < \mu(E). \tag{11.9}$$

It is a consequence of the compactness of Ω that every measure μ on Ω can be uniquely written in the form

$$\mu = \mu_c + \mu_d \tag{11.10}$$

where μ_c is continuous and μ_d is discrete.

We shall occasionally have to deal with measures on a Hilbert space $\mathscr{H}$, which will always be taken to mean countably additive measures on the Borel σ-field of the Hilbert space $\mathscr{H}$. If $\mathscr{H}_1$ is a finite-dimensional subspace of $\mathscr{H}$ and E is a Borel set in $\mathscr{H}_1$, then a set of the form

$$E + \mathscr{H}_1^{\perp} = \{\alpha + \psi : \alpha \in E \quad \text{and} \quad \psi \in \mathscr{H}_1^{\perp}\} \tag{11.11}$$

is called a cylinder set in $\mathscr{H}$.

Theorem 11.2 *If $\mathscr{H}$ is a separable Hilbert space then the Borel σ-field $\mathscr{F}$ of $\mathscr{H}$ is the smallest class of sets containing all cylinder sets and closed under monotone increasing and decreasing countable limits.*

Proof If $\mathscr{G}$ is the σ-field generated by the cylinder sets and $\mathscr{M}$ is the smallest class of sets containing the cylinder sets and closed under increasing and decreasing sequences then $\mathscr{M} \subseteq \mathscr{G}$. Since the class of cylinder sets is a field, $\mathscr{M}$ is a σ-field so $\mathscr{M} = \mathscr{G}$, and we only have to prove that $\mathscr{G} = \mathscr{F}$.

Let $\phi \in \mathscr{H}$ and let $\{\phi_n\}_{n=1}^{\infty}$ be a countable dense subset of the unit ball of $\mathscr{H}$. Then

$$\{\psi \in \mathscr{H} : \|\psi - \phi\| < \varepsilon\} = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \{\psi \in \mathscr{H} : |\langle \phi - \psi, \phi_n \rangle| \leqslant \varepsilon - m^{-1}\} \tag{11.12}$$

and so lies in $\mathscr{G}$. Since $\mathscr{H}$ is separable every open set is the union of a countable number of open balls, and therefore lies in $\mathscr{G}$. But $\mathscr{F}$ is generated by the open sets in $\mathscr{H}$, and therefore coincides with $\mathscr{G}$. QED

Notes

Most of this material is standard so I refer to text-books instead of original sources. (Bold numerals refer to section numbers.)

1 Dixmier (1969b); Dunford and Schwartz (1963); Reed and Simon (1972).
2 Alfsen (1971); Dunford and Schwartz (1958).
3,4,5 Schatten (1960); Dunford and Schwartz (1963).
6,7 Dixmier (1969a); Dixmier (1969b); Sakai (1971); Kelley (1955).
8 Dunford and Schwartz (1963); Kato (1966); Reed and Simon (1972).
9 Dunford and Schwartz (1958); Kato (1966); Hille and Phillips (1957).
10 Schatten (1960); Dixmier (1969b); Reed and Simon (1972).
11 Halmos (1950); Dunford and Schwartz (1958).

2

Operational Ideas

2.1 The repeated Stern–Gerlach experiment

Let us consider the conventional quantum-mechanical description of the measurement of an observable represented by a self-adjoint operator A with discrete spectrum on a Hilbert space $\mathscr{H}$. If A has eigenvalues λ_i with corresponding spectral projections P_i and the state of the system at the instant before the measurement is given by the trace-class operator ρ on $\mathscr{H}$, then the measurement yields the value λ_i with the probability $\mathrm{tr}[\rho P_i]$. These numbers are all non-negative and their sum is one if ρ is normalised. It is conventionally assumed that if the value λ_i is observed then the state at the instant after the measurement is

$$\rho_i = \frac{P_i \rho P_i}{\mathrm{tr}[\rho P_i]}. \tag{1.1}$$

Now suppose that the first measurement is immediately followed by a second, associated with another observable B with eigenvalues μ_j and spectral projections Q_j. Then the probability of obtaining the value μ_j at the second measurement is $\mathrm{tr}[\rho_i Q_j]$ conditional upon having obtained λ_i at the first. The probability of obtaining λ_i at the first measurement and μ_j at the second is therefore

$$\begin{aligned} p_{ij} &= \mathrm{tr}[\rho P_i] \cdot \mathrm{tr}[\rho_i Q_j] \\ &= \mathrm{tr}[P_i \rho P_i Q_j] \\ &= \mathrm{tr}[\rho P_i Q_j P_i]. \end{aligned} \tag{1.2}$$

Note that $p_{ij} \geqslant 0$ for each i, and that

$$\sum_{i,j} p_{ij} = 1 \tag{1.3}$$

so that we have a probability distribution on $\mathbb{Z} \times \mathbb{Z}$ ($\mathbb{Z}$ denoting the set of integers) which represents the result of the composite measurement even when A and B do not commute.

For an arbitrary subset E of $\mathbb{Z} \times \mathbb{Z}$ we now define the operator $R(E)$ on $\mathscr{H}$ by

$$R(E) = \sum \{P_i Q_j P_i \colon (i, j) \in E\}. \tag{1.4}$$

Then $R(\cdot)$ has the following properties, characteristic of a positive-operator-valued measure.

(*i*) $R(E) \geqslant 0$ for all $E \subseteq \mathbb{Z} \times \mathbb{Z}$; (1.5)

(*ii*) $R(\emptyset) = 0$ and $R(\mathbb{Z} \times \mathbb{Z}) = 1$; (1.6)

(*iii*) if $\{E_n\}$ is a countable collection of disjoint sets then

$$R\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} R(E_n) \tag{1.7}$$

the sum converging in the strong operator topology. Moreover for any initial state ρ the probability $\text{prob}\{\rho, E\}$ of the result of the composite measurement being in the set E is

$$\text{prob}\{\rho, E\} = \text{tr}[\rho R(E)]. \tag{1.8}$$

Thus it is possible to define a joint distribution for two non-commuting discrete observables by introducing a positive-operator-valued measure, in a way which is a natural consequence of conventional ideas in quantum theory. In general this positive-operator-valued measure depends on which of the two measurements is made first.

Lemma 1.1 *The operators $R(E)$ are independent of which measurement is made first for all $E \subseteq \mathbb{Z} \times \mathbb{Z}$, if and only if P_i and Q_i commute for all i and j.*

Proof It is clear that if P_i and Q_j commute for all i, j then

$$P_i Q_j P_i = Q_j P_i Q_j \tag{1.9}$$

for all i, j. Conversely if Eq. (1.9) holds then

$$P_i = \sum_k P_i Q_k P_i = \sum_k Q_k P_i Q_k \tag{1.10}$$

so

$$\begin{aligned} Q_j P_i &= \sum_k Q_j Q_k P_i Q_k \\ &= Q_j P_i Q_j \\ &= \sum_k Q_k P_i Q_k Q_j \\ &= P_i Q_j. \qquad \text{QED} \end{aligned} \tag{1.11}$$

We now return to the general case, where the projections P_i and Q_j need not commute. With each pair (i, j) we have an associated change of state

$$\rho \to \rho_i \to \frac{Q_j \rho_i Q_j}{\text{tr}[Q_j \rho_i]} \tag{1.12}$$

or

$$\rho \to \frac{Q_j P_i \rho P_i Q_j}{\mathrm{tr}[Q_j P_i \rho P_i Q_j]}. \tag{1.13}$$

It is convenient to ignore the normalisation factors and for each $E \subseteq \mathbb{Z} \times \mathbb{Z}$ to define the change of state induced by the composite measurement conditional upon the result lying in the set E to be

$$S_E(\rho) = \sum \{Q_j P_i \rho P_i Q_j : (i, j) \in E\}. \tag{1.14}$$

Clearly S_E is a positive linear map on the state space V. The family $\{S_E\}$ of state change maps contains all the information about the measurement. In particular

$$\begin{aligned} \mathrm{tr}[S_E(\rho)] &= \sum \{\mathrm{tr}[Q_j P_i \rho P_i Q_j] : (i, j) \in E\} \\ &= \sum \{\mathrm{tr}[\rho P_i Q_j P_i] : (i, j) \in E\} \\ &= \mathrm{tr}[\rho R(E)] \\ &= \mathrm{prob}\{\rho, E\}. \end{aligned} \tag{1.15}$$

One can extend the above considerations to the successive measurement of any finite sequence of discrete observables, but it is clearly more fruitful to place the problem in a more general context.

2.2 Operations and effects

If V_1 and V_2 are the state spaces of complex Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ an operation on V_1 is defined as a positive linear map $T: V_1 \to V_1$ which also satisfies

$$0 \leqslant \mathrm{tr}[T\rho] \leqslant \mathrm{tr}[\rho] \tag{2.1}$$

for all $\rho \in V_1^+$. The following lemma is relevant.

Lemma 2.1 *Every positive linear map $T: V_1 \to V_2$ is norm bounded and*

$$\|T\| = \sup\{\mathrm{tr}[T\rho] : \rho \in V_1^+ \quad \text{and} \quad \mathrm{tr}[\rho] = 1\}. \tag{2.2}$$

Also every positive linear functional on V_1 is bounded.

Proof The above equality is a simple consequence of Eqs (1.3.6) and (1.3.7), so we have only to prove that the right-hand side is finite. If this is not so then for every $n \in \mathbb{Z}^+$ there exists $\rho_n \in V_1^+$ with $\mathrm{tr}[\rho_n] = 1$ and $\mathrm{tr}[T(\rho_n)] \geqslant 4^n$. If $\rho = \sum_{n=1}^{\infty} 2^{-n}\rho_n$ then $\rho \in V_1^+$ and $0 \leqslant 2^{-n}\rho_n \leqslant \rho$. Therefore

$$0 \leqslant 2^{-n} T(\rho_n) \leqslant T(\rho) \tag{2.3}$$

so

$$0 \leqslant 2^n \leqslant \mathrm{tr}[T(\rho)]. \tag{2.4}$$

But this cannot be true for all $n \in \mathbb{Z}^+$. The proof of the second part of the lemma is similar. QED

We now define a normal positive linear map $T: \mathscr{L}_s(\mathscr{H}_2) \to \mathscr{L}_s(\mathscr{H}_1)$ to be a positive linear map such that if $A_n \uparrow A$ then $T(A_n) \uparrow T(A)$.

Lemma 2.2 *If $T: V_1 \to V_2$ is a positive linear map then the adjoint $T^*: \mathscr{L}_s(\mathscr{H}_2) \to \mathscr{L}_s(\mathscr{H}_1)$ is a normal positive linear map. Moreover*

$$0 \leqslant T^*(1) \leqslant 1 \tag{2.5}$$

if and only if

$$0 \leqslant \mathrm{tr}[T(\rho)] \leqslant \mathrm{tr}[\rho] \tag{2.6}$$

for all $\rho \in V_1^+$. Every normal positive linear map $S: \mathscr{L}_s(\mathscr{H}_2) \to \mathscr{L}_s(\mathscr{H}_1)$ is the adjoint of a unique positive linear map $T: V_1 \to V_2$.

Proof If $A \in \mathscr{L}_s(\mathscr{H}_2)^+$ and $\rho \in V_1^+$ then

$$\mathrm{tr}[T^*(A)\rho] = \mathrm{tr}[AT(\rho)] \geqslant 0 \tag{2.7}$$

so $T^*(A) \geqslant 0$ by chapter 1 section 5. If $A_n \uparrow A$ then A_n converges to A in the ultraweak operator topology so

$$\mathrm{tr}[T^*(A_n)\rho] = \mathrm{tr}[A_n T(\rho)] \to \mathrm{tr}[AT(\rho)] = \mathrm{tr}[T^*(A)\rho]. \tag{2.8}$$

Therefore $T^*(A_n)$ converges to $T^*(A)$ in the ultraweak operator topology and T^* is normal. The equivalence of Eqs (2.5) and (2.6) is a consequence of

$$\mathrm{tr}[(1 - T^*(1))\rho] = \mathrm{tr}[\rho] - \mathrm{tr}[T(\rho)] \tag{2.9}$$

valid for all $\rho \in V_1^+$.

Now let $S: \mathscr{L}_s(\mathscr{H}_2) \to \mathscr{L}_s(\mathscr{H}_1)$ be a normal positive linear map. For each $\rho \in V_1^+$ the map $\phi: A \to \mathrm{tr}[S(A)\rho]$ is positive and normal so by Lemma 1.6.1 there exists a unique $\sigma \in V_2^+$ such that $\phi(A) = \mathrm{tr}[A\sigma]$ for all $A \in \mathscr{L}_s(\mathscr{H}_2)$. If we write $\sigma = T(\rho)$ then

$$\mathrm{tr}[S(A)\rho] = \mathrm{tr}[AT(\rho)] \tag{2.10}$$

for all $A \in \mathscr{L}_s(\mathscr{H}_2)$ and $\rho \in V_1^+$. It is now easy to see that T extends uniquely to a positive linear map from V_1 to V_2 and that $S = T^*$. QED

We take an operation to describe the change of state associated with a measurement which passes only a proportion of the ensemble tested. The probability of transmission of a state ρ by an operation S is taken to be $\mathrm{tr}[S(\rho)]$ while the output state conditional upon transmission is taken to be

$$\rho' = \frac{S(\rho)}{\mathrm{tr}[S(\rho)]}. \tag{2.11}$$

Associated with the operation S is its effect, defined as the operator $A = S^*(1)$. Since

$$\mathrm{tr}[S(\rho)] = \mathrm{tr}[\rho A] \tag{2.12}$$

for all $\rho \in V$ the effect A determines the probability of transmission, but not the form of the transmitted state.

The set $\mathscr{E}$ of all effects consists of all bounded operators A on $\mathscr{H}$ such that $0 \leqslant A \leqslant 1$. It is partially ordered by the operator partial ordering and has a least element 0 and a greatest element 1. The map $A \to A^{\perp} \equiv 1 - A$ is an orthocomplementation in the sense that $A^{\perp\perp} = A$ for all $A \in \mathscr{E}$ and if $A \leqslant B$ then $B^{\perp} \leqslant A^{\perp}$.

Within $\mathscr{E}$ the set $\mathscr{Q}$ of all projections is often called the set of questions. $\mathscr{Q}$ inherits the partial ordering of $\mathscr{E}$ and is closed under the orthocomplementation. $\mathscr{Q}$ is moreover a lattice, the greatest lower bound $P \wedge Q$ of two projections being the projection onto the intersection of their ranges, and the least upper bound $P \vee Q$ being the projection onto the closure of the sum of their ranges. The set $\mathscr{E}$ is not a lattice but on the other hand is a convex set in $\mathscr{L}(\mathscr{H})$.

Lemma 2.3 *The projections on $\mathscr{H}$ may be characterised as those effects which are extreme points of $\mathscr{E}$.*

Proof Suppose $A_1, A_2 \in \mathscr{E}$, $0 < \lambda < 1$ and

$$\lambda A_1 + (1 - \lambda)A_2 = P \tag{2.13}$$

for some projection P. Then if $P\phi = 0$

$$\begin{aligned} 0 &\leqslant \lambda\langle A_1\phi, \phi\rangle \\ &\leqslant \lambda\langle A_1\phi, \phi\rangle + (1 - \lambda)\langle A_2\phi, \phi\rangle \\ &= \langle P\phi, \phi\rangle = 0. \end{aligned} \tag{2.14}$$

By the spectral theorem since $A_1 \geqslant 0$ and $\langle A_1\phi, \phi\rangle = 0$ it follows that $A_1\phi = 0$. Now since

$$\lambda(1 - A_1) + (1 - \lambda)(1 - A_2) = 1 - P \tag{2.15}$$

it follows similarly that if $P\psi = \psi$ then $(1 - A_1)\psi = 0$. Therefore A_1 coincides with P both on the range and on the null-space of P, and hence everywhere. Similarly $A_2 = P$, so P is an extreme point.

If $0 \leqslant A \leqslant 1$ and A is not a projection then by the spectral theorem there exists a $c \in \mathrm{sp}(A)$ with $0 < c < 1$. Let f be a continuous function on $[0, 1]$ such that

$$0 \leqslant t \pm f(t) \leqslant 1 \tag{2.16}$$

for all $0 \leqslant t \leqslant 1$ and $f(c) \neq 0$. Then by the spectral theorem if $A_1 = A + f(A)$ and $A_2 = A - f(A)$ it may be seen that $0 \leqslant A_1 \leqslant 1, 0 \leqslant A_2 \leqslant 1$, and neither of A_1 and A_2 is equal to A. On the other hand $\frac{1}{2}A_1 + \frac{1}{2}A_2 = A$, so A is not an extreme point of $\mathscr{E}$. QED

The above characterisation of the projections as the indecomposible effects is one of the fundamental reasons for their importance in quantum theory. There is, however, a sense in which projections cannot be distinguished from other effects in an infinite-dimensional Hilbert space.

Lemma 2.4 *If $\mathscr{H}$ is a separable infinite-dimensional Hilbert space then the set $\mathscr{Q}$ of projections is dense in $\mathscr{E}$ for the weak operator topology.*

Proof By the spectral theorem every $B \in \mathscr{E}$ may be approximated arbitrarily closely in norm by a sum

$$A = \sum_{r=1}^{R} \lambda_r P_r \tag{2.17}$$

where $P_1, \ldots, P_R$ are orthogonal projections with sum equal to one and $0 \leqslant \lambda_r \leqslant 1$. One of these projections, say P_1, has infinite-dimensional range so we can write

$$P_1 \mathscr{H} \simeq l^2\{\mathbb{Z}^+ \times (1, \ldots, R)\}. \tag{2.18}$$

Now let $Q_{s,n}$ be the projection on $\mathscr{H}$ with range in $P_1\mathscr{H}$ such that

$$(Q_{s,n}\psi)(m, r) = \begin{cases} \psi(m, r) & \text{if } m \geqslant n \text{ and } r = s \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

and define

$$P_{r,n} = \begin{cases} P_1 - \sum_{s=2}^{R} Q_{s,n} & \text{if } r = 1 \\ P_r + Q_{r,n} & \text{if } r \geqslant 2. \end{cases} \tag{2.20}$$

Then

$$\lim_{n \to \infty} \sum_{r=1}^{R} \lambda_r P_{r,n} = \sum_{r=1}^{R} \lambda_r P_r \tag{2.21}$$

in the weak operator topology, and each of the projections $P_{r,n}$ has infinite-dimensional range. It is therefore sufficient to show that the operator A of Eq. (2.17) is a weak limit of projections if each of the projections $P_1, \ldots, P_R$ has infinite-dimensional range.

Since any two infinite-dimensional separable Hilbert spaces are isomorphic we can put

$$\mathscr{H} \simeq \sum_{r=1}^{R} \oplus L^2(0,1) \tag{2.22}$$

and

$$(A\psi)(r,x) = \lambda_r \psi(r,x) \tag{2.23}$$

for all $\psi \in \mathscr{H}$, $l \leqslant r \leqslant R$ and $x \in [0,1]$. If $n \in \mathbb{Z}^+$ we define the projection Q_n on $\mathscr{H}$ by

$$(Q_n\psi)(r,x) = \begin{cases} \psi(r,x) & \text{if } s2^{-n} \leqslant x < (s+\lambda_r)2^{-n} \\ 0 & \text{if } (s+\lambda_r)2^{-n} \leqslant x < (s+1)2^{-n} \end{cases} \tag{2.24}$$

where s is any integer. If ψ is continuous then it is easy to see that

$$\begin{aligned} \lim_{n\to\infty} \langle Q_n\psi, \psi\rangle &= \sum_{r=1}^{R} \lambda_r \int_0^1 |\psi(r,x)|^2\, dx \\ &= \sum_{r=1}^{R} \lambda_r \langle P_r\psi, \psi\rangle. \end{aligned} \tag{2.25}$$

Since such ψ are dense, $Q_n \xrightarrow{w} A$ as required. QED

2.3 Pure operations

If V_1 and V_2 are the state spaces of Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively we say that a positive linear map $T\colon V_1 \to V_2$ is pure if $T(\rho)$ is a pure element of V_2^+ whenever ρ is a pure element of V_1^+. This terminology is unfortunate in that the pure positive linear maps are not exactly the same as the pure elements of the cone of positive linear maps from V_1 into V_2, but it is now generally accepted. We identify T with its unique complex linear extension

$$T\colon \mathscr{T}(\mathscr{H}_1) \to \mathscr{T}(\mathscr{H}_2).$$

We define a conjugate linear map $B\colon \mathscr{H}_1 \to \mathscr{H}_2$ to be a map such that

$$B(\alpha_1\phi_1 + \alpha_2\phi_2) = \bar{\alpha}_1 B(\phi_1) + \bar{\alpha}_2 B(\phi_2) \tag{3.1}$$

for all $\alpha_1\alpha_2 \in \mathbb{C}$ and $\phi_1, \phi_2 \in \mathscr{H}_1$. We define an antiunitary map from $\mathscr{H}_1$ to $\mathscr{H}_2$ to be conjugate linear isometric map of $\mathscr{H}_1$ onto $\mathscr{H}_2$.

Theorem 3.1 *Every pure positive linear map $T\colon\mathscr{T}(\mathscr{H}_1)\to\mathscr{T}(\mathscr{H}_2)$ is of one of the following three forms:*

(i)
$$T(\rho) = B\rho B^* \tag{3.2}$$

where $B: \mathscr{H}_1 \to \mathscr{H}_2$ *is bounded and linear;*

(*ii*) $$T(\rho) = B\rho^* B^* \tag{3.3}$$

where $B: \mathscr{H}_1 \to \mathscr{H}_2$ *is bounded and conjugate linear;*

(*iii*) $$T(\rho) = \mathrm{tr}[\rho B] |\psi\rangle\langle\psi| \tag{3.4}$$

where $B \in \mathscr{L}(\mathscr{H}_1)^+$ *and* $\psi \in \mathscr{H}_2$.

In cases (*i*) *and* (*ii*) *the operator B is uniquely determined up to a constant of absolute value one.*

Proof An operator T of type (*iii*) will be called degenerate. We prove the theorem in cases of gradually increasing generality.

Case 1 dim $\mathscr{H}_1$ = dim $\mathscr{H}_2 = 2$ and T is trace-preserving. We identify $\mathscr{H}_1$ and $\mathscr{H}_2$ with $\mathbb{C}^2$ and let $\sigma_1, \sigma_2, \sigma_3$ denote the Pauli matrices. Given $x \in \mathbb{R}^3$ the operator

$$\rho = \frac{1 + \sum_{r=1}^{3} x_r \sigma_r}{2} \tag{3.5}$$

is positive and of trace one if and only if x lies in the unit ball of $\mathbb{R}^3$. A positive linear trace-preserving map T therefore induces an affine map $S: \mathbb{R}^3 \to \mathbb{R}^3$ which maps the unit ball into itself. T is pure if and only if S maps the surface of the unit ball into itself. By inspection this can happen in only three ways. If S maps every point of the unit ball onto one point of its surface then T is degenerate. Otherwise S is an orthogonal rotation of determinant ± 1, in which case the spin $\frac{1}{2}$ theory shows that T is of the type (*i*) or (*ii*) with B unitary or antiunitary.

We conclude that T is either degenerate or isometric. In the latter case if

$$T(|u_r\rangle\langle u_r|) = |v_r\rangle\langle v_r| \tag{3.6}$$

for $r = 1, 2$ then

$$T(|u_1\rangle\langle u_2|) = \begin{cases} & e^{i\theta}|v_1\rangle\langle v_2| \\ \text{or} & \\ & e^{i\theta}|v_2\rangle\langle v_1|. \end{cases} \tag{3.7}$$

Case 2 dim $\mathscr{H}_1 = n$ and T is trace-preserving. Let u_1, u_2 be unit vectors in $\mathscr{H}_1$ and

$$T(|u_r\rangle\langle u_r|) = |v_r\rangle\langle v_r|. \tag{3.8}$$

If $\phi \in \mathscr{M}_1 \equiv \mathrm{lin}\{u_1, u_2\}$ then

$$0 \leqslant |\phi\rangle\langle\phi| \leqslant c|u_1\rangle\langle u_1| + c|u_2\rangle\langle u_2| \tag{3.9}$$

for some $c \in \mathbb{R}^+$. If

$$T(|\phi\rangle\langle\phi|) = |\psi\rangle\langle\psi| \tag{3.10}$$

then by the positivity of T

$$0 \leqslant |\psi\rangle\langle\psi| \leqslant c|v_1\rangle\langle v_1| + c|v_2\rangle\langle v_2| \tag{3.11}$$

so $\psi \in \mathcal{M}_2 \equiv \text{lin}\{v_1, v_2\}$. Therefore T restricts to a pure positive linear map $T_0: \mathcal{T}(\mathcal{M}_1) \to \mathcal{T}(\mathcal{M}_2)$, which is either degenerate or isometric. Letting u_1, u_2 vary continuously over the unit sphere in $\mathcal{H}_1$ we see that T is either degenerate or isometric on the set of pure states. We continue with the latter case.

Let $u_1, \ldots, u_n$ be an orthonormal basis of $\mathcal{H}_1$ and let

$$T(|u_r\rangle\langle u_r|) = |v_r\rangle\langle v_r| \tag{3.12}$$

for $r = 1, \ldots, n$. Since T is isometric on the set of pure states, $v_1, \ldots, v_n$ is an orthonormal set in $\mathcal{H}_2$. Moreover

$$T(|u_r\rangle\langle u_s|) = \begin{cases} & z_{rs}|v_r\rangle\langle v_s| \\ \text{or} & \\ & z_{rs}|v_s\rangle\langle v_r| \end{cases} \tag{3.13}$$

for all r, s. If

$$T(|u_1\rangle\langle u_2|) = z_{12}|v_1\rangle\langle v_2| \tag{3.14}$$

and

$$T(|u_1\rangle\langle u_3|) = z_{13}|v_3\rangle\langle v_1| \tag{3.15}$$

then

$$T(|u_1\rangle\langle u_2 + u_3|) = z_{12}|v_1\rangle\langle v_2| + z_{13}|v_3\rangle\langle v_1| \tag{3.16}$$

which is of rank two, contradicting Eq. (3.7). Repetition of this argument proves that in Eq. (3.13) the same choice must be made for all r, s.

There are therefore two cases, of which we treat only one. If

$$T(|u_r\rangle\langle u_s|) = z_{rs}|v_s\rangle\langle v_r| \tag{3.17}$$

for all r, s and $u = u_1 + \cdots + u_n$ then

$$T(|u\rangle\langle u|) = \sum_{r,s} z_{rs}|v_s\rangle\langle v_r| \tag{3.18}$$

is a pure state so there exist $a_1, \ldots, a_n$ such that $z_{rs} = \bar{a}_r a_s$. Since $|z_{rs}| = 1$ for all r, s, $|a_r| = 1$ for all r. If B is the antiunitary operator such that $Bu_r = a_r v_r$ then

$$T(|u_r\rangle\langle u_s|) = B|u_s\rangle\langle u_r|B^* \tag{3.19}$$

so T is of type (*ii*). The uniqueness of B up to a constant is evident.

Case 3 dim $\mathcal{H}_1 = n$. We define the positive operator A on $\mathcal{H}_1$ by

$$\text{tr}[A\rho] = \text{tr}[T(\rho)] \tag{3.20}$$

and let P be the projection onto the orthogonal complement of the null-space of A. We then claim that

$$T(\rho) = T(P\rho P) \tag{3.21}$$

for all ρ. If $\psi_0 = P\psi_0$ and $P\psi_1 = 0$ then

$$\mathrm{tr}[T(|\psi_1\rangle\langle\psi_1|)] = \langle A\psi_1, \psi_1\rangle = 0 \tag{3.22}$$

so $T(|\psi_1\rangle\langle\psi_1|) = 0$. If $\psi \in \mathrm{lin}\{\psi_0, \psi_1\}$ then

$$0 \leqslant |\psi\rangle\langle\psi| \leqslant c|\psi_0\rangle\langle\psi_0| + c|\psi_1\rangle\langle\psi_1| \tag{3.23}$$

for some $c \in \mathbb{R}^+$ so

$$0 \leqslant T(|\psi\rangle\langle\psi|) \leqslant cT(|\psi_0\rangle\langle\psi_0|). \tag{3.24}$$

Since the right-hand side is a pure state

$$T(|\psi\rangle\langle\psi|) = c_\psi T(|\psi_0\rangle\langle\psi_0|) \tag{3.25}$$

for some c_ψ. Now c_ψ is a positive quadratic form on $\mathrm{lin}\{\psi_0, \psi_1\}$ which vanishes on ψ_1 and equals one on ψ_0. Therefore

$$c_\psi = |\langle\psi, \psi_0\rangle|^2 \cdot \|\psi_0\|^{-4} \tag{3.26}$$

which establishes Eq. (3.21) for all pure states, and hence all states. By restriction to the subspace $P\mathscr{H}_1$ we see that it is sufficient to prove case 3 when the operator A is invertible.

In this case the map S defined by

$$S(\rho) = A^{-1/2}T(\rho)A^{-1/2} \tag{3.27}$$

is positive, pure and trace-preserving and so of one of the three given types. Since

$$T(\rho) = A^{1/2}S(\rho)A^{1/2} \tag{3.28}$$

T is also of one of the given types. The uniqueness of B may also be deduced from the corresponding result in case 2.

Case 4 The general case. Let $\mathscr{H}^n$ be an increasing sequence of finite-dimensional subspaces with union dense in $\mathscr{H}_1$ and consider the restrictions $T_n: \mathscr{T}(\mathscr{H}^n) \to \mathscr{T}(\mathscr{H}_2)$ of T. If T_n are all degenerate then the norm boundedness of T implies that it also is degenerate. If this is not so then T_n are of types (*i*) or (*ii*) for all large enough n, and the operators B_n are compatible up to a constant, and uniformly bounded in norm. There is therefore a unique bounded operator B on $\mathscr{H}_1$ which extends all of them, up to a constant. Equation (3.2) or (3.3) is valid for all states with support in some $\mathscr{H}^n$, and hence by continuity for all states on $\mathscr{H}_1$. QED

Corollary 3.2 *Let $T: V \to V$ be a positive linear map with positive inverse such that*

$$\mathrm{tr}[T(\rho)] = \mathrm{tr}[\rho] \tag{3.29}$$

for all $\rho \in V$. Then there exists a unitary or antiunitary map U on $\mathscr{H}$ such that

$$T(\rho) = U\rho U^* \tag{3.30}$$

for all $\rho \in V$.

Proof Since T and T^{-1} are positive, they are pure maps. A map of type (*iii*) cannot be invertible, except in the trivial case where $\mathscr{H}$ is one-dimensional, because its range has dimension one. Maps of type (*i*) or (*ii*) are trace-preserving if and only if U is isometric, and invertible if and only if U is invertible. QED

Corollary 3.3 *Let $T: \mathscr{L}_s(\mathscr{H}) \to \mathscr{L}_s(\mathscr{H})$ be a positive linear map with positive inverse and $T(1) = 1$. Then there is a unitary or antiunitary map U on $\mathscr{H}$ such that*

$$T(A) = U^*AU \tag{3.31}$$

for all $A \in \mathscr{L}_s(\mathscr{H})$. There is a unique complex linear extension of T to $\mathscr{L}(\mathscr{H})$ which is either an algebra automorphism or an algebra antiautomorphism.

Proof If A_n is a monotonically increasing (generalised) sequence in $\mathscr{L}_s(\mathscr{H})$ which is norm bounded it converges strongly to its least upper bound A. Since T and T^{-1} are positive the sequence $T(A_n)$ has least upper bound and therefore limit $T(A)$. Thus T is normal and by Lemma 2.2 $T = S^*$, where S is a positive linear map on V with positive inverse. If $\rho \in V$ then

$$\begin{aligned}\mathrm{tr}[S(\rho)] &= \mathrm{tr}[\rho S^*(1)]\\ &= \mathrm{tr}[\rho]\end{aligned} \tag{3.32}$$

so there exists a unitary or antiunitary U on $\mathscr{H}$ such that

$$S(\rho) = U\rho U^* \tag{3.33}$$

for all $\rho \in V$. If $A \in \mathscr{L}_s(\mathscr{H})$ then

$$\mathrm{tr}[T(A)\rho] = \mathrm{tr}[AS(\rho)] = \mathrm{tr}[AU\rho U^*] = \mathrm{tr}[U^*AU\rho] \tag{3.34}$$

for all $\rho \in \mathscr{T}_s(\mathscr{H})$ so

$$T(A) = U^*AU. \tag{3.35}$$

If U is unitary the equation

$$T^{\sim}(A) = U^*AU \tag{3.36}$$

defines an extension of T to a complex linear automorphism of $\mathscr{L}(\mathscr{H})$. If U is antiunitary the equation

$$T^{\sim}(A) = U^*A^*U \tag{3.37}$$

defines an extension of T to a complex linear anti-automorphism of $\mathscr{L}(\mathscr{H})$.

QED

The above corollary makes the following theorem plausible, but we omit its surprisingly difficult proof, which involves some theorems about group cocycles.

Theorem 3.4 *Let $T_t: V \to V$ be a strongly continuous one-parameter group of positive linear maps on the state space V of $\mathscr{H}$ such that*

$$\mathrm{tr}[T_t(\rho)] = \mathrm{tr}[\rho] \tag{3.38}$$

for all $\rho \in V$. Then there exists an (unbounded) self-adjoint operator H on $\mathscr{H}$ such that

$$T_t(\rho) = e^{-iHt}\rho\, e^{iHt} \tag{3.39}$$

for all $t \in \mathbb{R}$.

2.4 Irreversible dynamics

It is generally considered that a quantum-mechanical system which is closed, or isolated from the external world, has a Hamiltonian evolution. If $\mathscr{H}$ is the Hilbert space of the system this is expressed by the existence of a self-adjoint operator H, called the Hamiltonian, such that the state ρ_t at time t is computed from the state at time zero according to the law

$$\rho_t = e^{-itH}\rho\, e^{itH}. \tag{4.1}$$

Open quantum-mechanical systems, however, frequently exhibit irreversible behaviour, for example a tendency for their entropy to increase with time. Such behaviour is a consequence of their interaction with the external world, which exerts an influence on the systems of a statistical or thermal nature.

According to this view, irreversible evolution is of a less fundamental nature than Hamiltonian evolution. If one were to quantise the external world then there would be a Hamiltonian for the total system which would include terms describing the interaction between the system and the external world. However, such a description would generally be very complicated because the external world should have an infinite number of degrees of freedom, even if the system itself is quite simple.

The possibility remains that one might be able to describe the evolution of the system, and the external world's influence on it, by phenomenological equations which are not of Hamiltonian type. If the influence of the external world on the system is of a statistical nature then we might expect to be able to calculate the state of the system at all times $t \geqslant t_0$ if it is known for all $t \leqslant t_0$. The simplest case to treat is the Markovian one, where the state for times $t \geqslant t_0$ depends only on the state at time t_0.

Given a Hilbert space $\mathscr{H}$ with state space V we define a dynamical semigroup to be a one-parameter family of linear operators $T_t : V \to V$ defined for all $t \geqslant 0$ and satisfying

(*i*) If $\rho \geqslant 0$ then $T_t\rho \geqslant 0$ for all $t \geqslant 0$; (4.2)

(*ii*) $\mathrm{tr}[T_t\rho] = \mathrm{tr}[\rho]$ for all $\rho \in V$ and all $t \geqslant 0$; (4.3)

(*iii*) $T_s T_t \rho = T_{t+s}\rho$ for all $\rho \in V$ and all $s, t \geqslant 0$; (4.4)

(*iv*) $\lim\limits_{t \to 0} \|T_t\rho - \rho\|_{\mathrm{tr}} = 0$ for all $\rho \in V$. (4.5)

The first two conditions correspond to ideas of conservation of probability. Condition (*iii*) expresses both the Markovicity condition and the stationarity (independence of time) of the evolution law.

One particular case of this is the Hamiltonian evolution law. If

$$T_t(\rho) = e^{-iHt}\rho\, e^{iHt} \tag{4.6}$$

where the Hamiltonian H is a bounded linear operator then the infinitesimal generator Z of T_t is given by

$$Z(\rho) = -i[H, \rho] \tag{4.7}$$

which lies in V if $H = H^* \in \mathscr{L}(\mathscr{H})$ and $\rho \in V$. The same result holds formally if H is unbounded but one has to specify the relevant domains (see chapter 5 section 5).

In the following chapters we shall write down the generators of various dynamical semigroups and study the time evolution of the systems. We shall also examine how dynamical semigroups can be obtained by "reduction of the wave-packet" from the Hamiltonian dynamics of a larger system.

We start by giving a simple example known as the Wigner–Weisskopf atom, which has Hilbert space $\mathscr{H} = \mathbb{C}^2$ so that everything can be explicitly computed. $\mathscr{T}(\mathscr{H})$ or $\mathscr{L}(\mathscr{H})$ is the space of all complex 2×2 matrices. The state space is the set of matrices

$$\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : a = \bar{a}, \quad c = \bar{b}, \quad d = \bar{d}\right\} \tag{4.8}$$

and has real dimension 4. The states are the self-adjoint matrices with eigenvalues λ_0 and λ_1 satisfying

$$\lambda_0 \geqslant 0, \qquad \lambda_1 \geqslant 0, \qquad \lambda_0 + \lambda_1 = 1. \tag{4.9}$$

The ground state P_0 and the excited state P_1 of the system are

$$P_0 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.10}$$

We shall also use the operators

$$A_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{4.11}$$

As a candidate for an evolution equation on V we consider

$$\frac{d\rho}{dt} = Z\rho \tag{4.12}$$

where $Z: V \to V$ is defined by

$$\begin{aligned} Z(\rho) = & -i[\omega_0 P_0 + \omega_1 P_1, \rho] \\ & + c_0\{2A_- \rho A_+ - A_+ A_- \rho - \rho A_+ A_-\} \\ & + c_1\{2A_+ \rho A_- - A_- A_+ \rho - \rho A_- A_+\} \end{aligned} \tag{4.13}$$

where $\omega_0, \omega_1 \in \mathbb{R}$ and $c_0, c_1 \geqslant 0$. In this equation the first term on the right-hand side is a Hamiltonian term, the second term tends to force ρ into the ground state and the third term tends to force ρ into the excited state.

Theorem 4.1 *The operator Z is the generator of a dynamical semigroup on V.*

Proof Since V is finite-dimensional Z is bounded and the formula

$$T_t(\rho) = e^{Zt}\rho = \sum_{n=0}^{\infty} \frac{t^n Z^n \rho}{n!} \tag{4.14}$$

defines a continuous one-parameter semigroup on V. The condition of Eq. (4.3) is equivalent to the infinitesimal condition

$$\frac{d}{dt}\operatorname{tr}[T_t \rho] = 0 \tag{4.15}$$

or

$$\operatorname{tr}[Z(\rho)] = 0 \tag{4.16}$$

for all $\rho \in V$, which may be verified by a trivial computation.

The positivity condition of Eq. (4.2) is more difficult. We put $Z = Z_1 + Z_2$ where

$$Z_1(\rho) = X\rho + \rho X^* \tag{4.17}$$

and

$$X = -i\omega_0 P_0 - i\omega_1 P_1 - c_0 A_+ A_- - c_1 A_- A_+ \tag{4.18}$$

while

$$Z_2(\rho) = 2c_0 A_- \rho A_+ + 2c_1 A_- \rho A_- . \tag{4.19}$$

Now

$$e^{Z_1 t}(\rho) = e^{Xt}\rho\, e^{X^* t} \tag{4.20}$$

is clearly a positive linear map on V. Since Z_2 is positive and

$$e^{Z_2 t} = \sum_{n=0}^{\infty} \frac{Z_2^n t^n}{n!} \tag{4.21}$$

it follows that $e^{Z_2 t}$ is positive if $t \geqslant 0$. Finally by Lie's formula or (A8)

$$e^{Zt} = \lim_{n\to\infty} (e^{Z_1 t/n}\, e^{Z_2 t/n})^n \tag{4.22}$$

is, for $t \geqslant 0$, a limit of positive linear maps and therefore positive itself.

QED

If ρ is diagonal, that is

$$\rho = \lambda_0 P_0 + \lambda_1 P_1 \tag{4.23}$$

then

$$\frac{d\rho}{dt} = \lambda_1 c_0(2P_1 - 2P_0) + \lambda_0 c_1(2P_0 - 2P_1) \tag{4.24}$$

so $\rho(t)$ is diagonal at all subsequent times.

If we write

$$\rho(t) = \lambda_0(t)P_0 + \lambda_1(t)P_1 \tag{4.25}$$

then

$$\begin{aligned} \lambda_0'(t) &= -2c_1\lambda_0(t) + 2c_0\lambda_1(t), \\ \lambda_1'(t) &= 2c_1\lambda_0(t) - 2c_0\lambda_1(t). \end{aligned} \tag{4.26}$$

Solution of these equations yields

$$\lim_{t\to\infty} \lambda_0(t) = \frac{c_0}{c_0 + c_1}, \qquad \lim_{t\to\infty} \lambda_1(t) = \frac{c_1}{c_0 + c_1} \tag{4.27}$$

as the steady state occupation probability of the unexcited and excited states respectively.

If the initial state ρ is not diagonal we can proceed as follows. V has real dimension 4 so $Z: V \to V$ can be written as a real 4×4 matrix. Further information about the time evolution of a non-diagonal initial state ρ may be obtained by finding the eigenvalues and eigenvalues of this matrix. We shall give a much more thorough analysis of this type of problem in chapter 6.

2.5 A more general formulation*

Instead of defining a state space to be the space of all self-adjoint trace class operators on a Hilbert space, one may take a more axiomatic approach. Advantages of this are that it allows one to see more clearly the relationship between quantum theory and probability theory, and secondly that the axiomatic approach fits in naturally with the use of C^*-algebras, which have proved so helpful in the development of quantum statistical mechanics and quantum field theory.

We start by defining an (abstract) state space as a real Banach space V partially ordered by a cone V^+, and satisfying:

(*i*) V^+ is a closed set in V; (5.1)

(*ii*) if $x, y \in V^+$ then $\|x\| + \|y\| = \|x + y\|$; (5.2)

(*iii*) if $x \in V$ and $\varepsilon > 0$ then there exist $x_1, x_2 \in V^+$ such that $x = x_1 - x_2$ and

$$\|x_1\| + \|x_2\| < \|x\| + \varepsilon. \tag{5.3}$$

In any state space the norm is linear on the positive cone and can be uniquely extended to a positive linear functional $\tau: V \to \mathbb{R}$ which satisfies

$$|\tau(x)| \leqslant \|x\| \tag{5.4}$$

for all $x \in V$ and

$$\tau(x) = \|x\| \tag{5.5}$$

for all $x \in V^+$. The states are defined as the elements of

$$\{x \in V^+ : \tau(x) = 1\} \tag{5.6}$$

and form a convex set whose extreme points, if any, are called pure states.

We list a number of important examples of abstract state spaces.

(1) We let V be the space of self-adjoint trace class operators on a Hilbert space $\mathscr{H}$, with the trace norm.

(2) If X is a compact Hausdorff space let V be the space of all real finite signed measures μ on the Baire σ-field $\mathscr{F}$ of X. Let V^- be the cone of non-negative measures and let

$$\|\mu\| = \sup\{|\mu(E) - \mu(X \setminus E)| : E \in \mathscr{F}\}. \tag{5.7}$$

(3) Let $\mathscr{A}$ be a C^*-algebra and let V be the space of all bounded linear functionals ϕ on $\mathscr{A}$ which are self-adjoint in the sense that $\phi(x^*) = \overline{\phi(x)}$ for all $x \in \mathscr{A}$. Define

$$V^+ = \{\phi \in V : \phi(x^*x) \geqslant 0 \quad \text{all } x \in \mathscr{A}\} \tag{5.8}$$

and let the norm be the dual space norm.

(4) Let $\mathscr{A}$ be a von Neumann algebra and let V be the space of all ultra-weakly continuous linear functionals on $\mathscr{A}$. This class of states is smaller than that of (3) but still forms a state space.

Example (1) gives one the usual states of quantum theory and example (2) those of probability theory. Both are contained within example (3), the first by taking $\mathscr{A}$ as the algebra of all compact operators on $\mathscr{H}$, and the second by putting $\mathscr{A} = C(X)$. One may regard (3) as a special case of (4) if one takes the von Neumann algebra of (4) as the universal enveloping algebra (second dual space) of the C^*-algebra of (3).

We note that many of the definitions of this book may be made for an abstract state space V. For example the Banach dual space V^* may be partially ordered by defining $\phi \geqslant \psi$ to mean that $\phi(x) \geqslant \psi(x)$ for all $x \in V^-$. The set $\mathscr{E}$ of effects can then be defined as

$$\mathscr{E} = \{\phi \in V^* : 0 \leqslant \phi \leqslant \tau\} \tag{5.9}$$

which is a convex set and compact for the weak* topology of V^*. $\mathscr{E}$ is also a partially ordered set with the orthocomplementation

$$\phi^\perp \equiv \tau - \phi. \tag{5.10}$$

We define the questions to be the extreme effects. The set $\mathscr{Q}$ of all questions inherits the partial ordering of $\mathscr{E}$ and is closed under the orthocomplementation. The following theorem forges a link between the axiomatic operational approach to quantum theory and the axiomatic lattice-theoretic approach.

Theorem 5.1 *Suppose that V is a finite-dimensional state space and that for any $\phi_0, \phi_1 \in \mathscr{E}$ there exists $\phi_2 \in \mathscr{E}$ with $\phi_2 \geqslant \phi_0, \phi_2 \geqslant \phi_1$ and*

$$K_0(\phi_2) \supseteq K_0(\phi_0) \cap K_0(\phi_1) \tag{5.11}$$

where

$$K_0(\phi_i) = \{x \in V^+ : \phi_i(x) = 0\}. \tag{5.12}$$

Then the set of questions is a complete orthocomplemented lattice.

We precede the proof by a lemma.

Lemma 5.2 *Under the above conditions an element ϕ of $\mathscr{E}$ is an extreme point if and only if*

$$\phi = \max\{\psi \in \mathscr{E} : K_0(\psi) \supseteq K_0(\phi)\}. \tag{5.13}$$

Proof Suppose ϕ satisfies this last equation and $\phi = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2$ where $\phi_1, \phi_2 \in \mathscr{E}$. If $x \in V^+$ then $\phi(x) = 0$ if and only if $\phi_1(x) = 0$ and $\phi_2(x) = 0$, so

$$K_0(\phi) = K_0(\phi_1) \cap K_0(\phi_2). \tag{5.14}$$

By Eq. (5.11) there exists $\phi_0 \geqslant \phi_1, \phi_2$ such that $K_0(\phi_0) \supseteq K_0(\phi)$ and by Eq. (5.13) it follows that $\phi \geqslant \phi_0$. Therefore

$$\phi_0 \leqslant \phi = \tfrac{1}{2}\phi_1 + \tfrac{1}{2}\phi_2 \leqslant \tfrac{1}{2}\phi_0 + \tfrac{1}{2}\phi_0 = \phi_0 \tag{5.15}$$

and the equality of the two ends of this equation implies that $\phi = \phi_1 = \phi_2$, so ϕ is extreme.

Conversely suppose that ϕ is an extreme point of $\mathscr{E}$. If $\{\psi_n\}_{n=1}^{\infty}$ is a dense subset of

$$\{\psi \in \mathscr{E} : K_0(\psi) \supseteq K_0(\phi)\} \tag{5.16}$$

then by Eq. (5.11) there is sequence ϕ_n in the same set with $\phi_n \geqslant \phi_{n-1}, \psi_n$. This sequence is monotonically increasing and converges to an element ϕ_0 of the set which may be seen to be the maximum element of the set. If $x \in K$, where K is the compact convex set

$$\{x \in V^+ : \tau(x) = 1\} \tag{5.17}$$

then

$$(\phi_0 - \phi)(x) < 1 \tag{5.18}$$

if $\phi(x) > 0$. If $\phi(x) = 1$ then $\phi_0(x) = 0$ and

$$(\phi_0 - \phi)(x) = 0. \tag{5.19}$$

Therefore the continuous function $(\phi_0 - \phi)$ on K never attains the value 1 and there exists $\varepsilon > 0$ such that

$$\begin{aligned} \phi_0 - \phi &= \tfrac{1}{2}(1 + \varepsilon)(\phi_0 - \phi) + \tfrac{1}{2}(1 - \varepsilon)(\phi_0 - \phi) \\ &= \tfrac{1}{2}u_1 + \tfrac{1}{2}u_2 \end{aligned} \tag{5.20}$$

where u_1, u_2 lie in $\mathscr{E}$. Now

$$K_0(u_i) = K_0(\phi_0 - \phi) \supseteq K_0(\phi_0) \supseteq K_0(\phi) \tag{5.21}$$

so by the definition of ϕ_0 it follows that $u_1, u_2 \leqslant \phi_0$. Therefore

$$\phi = \tfrac{1}{2}(\phi_0 - u_1) + \tfrac{1}{2}(\phi_0 - u_2) \tag{5.22}$$

where $(\phi_0 - u_i) \in \mathscr{E}$. Since ϕ is extreme $\phi = \phi_0 - u_i$ so $u_1 = u_2$ and $\phi = \phi_0$. Therefore ϕ satisfies Eq. (5.13). QED

Proof of theorem Since $\mathscr{Q}$ is closed under the orthocomplementation it is only necessary to prove that any set $S \subseteq \mathscr{Q}$ has a least upper bound in $\mathscr{Q}$. By selecting a monotonic sequence as before we may show that the set

$$T = \left\{\psi \in \mathscr{E} : K_0(\psi) \supseteq \bigcap_{\phi \in S} K_0(\phi)\right\} \tag{5.23}$$

has a maximum element ϕ_0. If $K_0(\psi) \supseteq K_0(\phi_0)$ then ψ lies in T so $\psi \leqslant \phi_0$. Therefore ϕ_0 satisfies Eq. (5.13) and so lies in $\mathscr{Q}$. If $\phi \in S$ then also $\phi \in T$ so $\phi_0 \geqslant \phi$. Conversely if $\phi \leqslant \phi_1 \in \mathscr{Q}$ for all $\phi \in S$ then $K_0(\phi_1) \subseteq K_0(\phi)$ for all $\phi \in S$, so

$$K_0(\phi_1) \subseteq \bigcap_{\phi \in S} K_0(\phi) \subseteq K_0(\phi_0). \tag{5.24}$$

Since $\phi_1 \in \mathscr{Q}$, it is a solution of Eq. (5.13) and so $\phi_0 \leqslant \phi_1$. Therefore ϕ_0 is the least upper bound of S in $\mathscr{Q}$. QED

One of the real advantages of the abstract formulation is that it ties together quantum theory and probability theory. On many occasions subsequently we shall write down definitions which have been reached by generalising a concept from probability theory to an abstract state space and then re-interpreting the definition in quantum theory. In fact when we say that some idea is introduced by analogy with probability theory, we shall usually mean that the quantum theoretical and probabilistic ideas are both particular cases of one which can be formulated on an abstract state space.

A symmetry T of an abstract state space V is defined as a positive linear map $T: V \to V$ with positive inverse and such that

$$\tau(Tx) = \tau(x) \tag{5.25}$$

for all $x \in V$. The set of all symmetries clearly forms a group. One of the most interesting ways in which quantum theory appears to be a richer theory than probability theory is the size of its group of symmetries. In example (2) above if V is finite-dimensional every symmetry is induced by one of the finite number of permutations of the underlying set X. In quantum theory, however, even if the Hilbert space is finite-dimensional, the group of symmetries given by Corollary 3.2 is still a Lie group. This fact has had a very profound effect on the development of quantum theory.

Notes

1 Ideas on repeated measurement may be found in Goldberger and Watson (1964), Furry (1966) and Davies and Lewis (1970). Lemma 1.1 is due to Furry while the case where the projections do not commute was developed by Davies and Lewis.

2 The term "operation" was introduced by Haag and Kastler (1964), and "effect" by Ludwig (1967, 1968). Lemmas 2.3 and 2.4 are well-known results from operator theory.

3 Theorem 3.1 was proved by Davies (1969) using a result of Stormer (1963), but related results may be found in Haag and Kastler (1964) and Hellwig and Kraus (1969). Corollaries 3.2 and 3.3 are proved in Hunziker (1972) and Varadarajan (1968) and were greatly generalised by Kadison (1951), but in a slightly different form are due to Wigner (1959). For Theorem 3.4 see Kadison (1965) and Varadarajan (1970).

4 The first attempt at a justification of irreversible evolution laws was made by Pauli. More recently dynamical semigroups have been proposed as the basis for a fundamentally new approach to the dynamics of irreversible systems by Ingarden (1973), Sudarshan *et al.* (1961) and Mehra and Sudarshan (1972). The generators of such semigroups were characterised in abstract terms by Kossakowski (1972b). The example given appears in several of these papers and is also treated by Haake (1973).

5 This general approach to quantum theory was developed in an axiomatic form by Ludwig, and we refer to Hartkamper and Neumann (1974) and Ludwig (1970) for further references. Theorem 5.1 was proved by Ludwig (1968) and this was followed up by Dahn (1968, 1973). Operations were introduced into the axiomatic theory by Dahn (1968), Davies and Lewis (1970), and Mielnik (1969). Srinivas (1973) compared the operational approach with the older lattice-theoretic approach (Birkhoff and von Neumann, 1936; Mackey, 1963; Varadarajan, 1968). The theory of operations on C^*-algebras has been developed by (Edwards, 1970, 1971; Haag and Kastler, 1964; Hellwig and Kraus, 1970; Steinmann, 1968). The enormous importance of symmetry arguments in quantum theory is clear from Wigner (1959), for example. In the axiomatic approach the group of symmetries is also very important (Varadarajan, 1968, 1970). Assumptions about its size have been used to restrict the variety of possible axiomatic structures by Davies (1974b) and Jones (1976).

3

Continuous Measurements

3.1 Observables

In the conventional accounts of quantum theory an observable is determined by a self-adjoint operator A on a Hilbert space $\mathscr{H}$. The expectation (or mean) of A in a state ρ is then defined by

$$E\{\rho, A\} = \mathrm{tr}[\rho A] \tag{1.1}$$

and the variance is taken to be

$$\mathrm{var}\{\rho, A\} = \mathrm{tr}[\rho A^2] - \{\mathrm{tr}[\rho A]\}^2. \tag{1.2}$$

The spectral theorem gives a representation

$$A = \int_{-\infty}^{\infty} \lambda P(d\lambda) \tag{1.3}$$

where $P(\cdot)$ is a projection-valued measure on $\mathbb{R}$. It is generally supposed that the probability that a measurement of A in the state ρ gives a result in the Borél set E is

$$\mathrm{prob}\{\rho, E\} = \mathrm{tr}[\rho P(E)]. \tag{1.4}$$

This suggests redefining an observable to be a projection-valued measure, but our discussion in the last chapter motivates an even more general definition, given below. We hope that the reader will agree that the generality of our definition of an observable is justified by its applications throughout the book.

Definition 1.1 *Let Ω be a set with a σ-field $\mathscr{F}$ and let $\mathscr{H}$ be a Hilbert space. A positive-operator-valued (POV) measure on Ω is defined to be a map $A: \mathscr{F} \to \mathscr{L}_s(\mathscr{H})$ such that*

(*i*) $A(E) \geqslant A(\varnothing) = 0$ *for all* $E \in \mathscr{F}$; (1.5)

(*ii*) *if* $\{E_n\}$ *is a countable collection of disjoint sets in* $\mathscr{F}$ *then*

$$A\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} A(E_n) \tag{1.6}$$

where the series converges in the weak operator topology. The POV measure is called an observable if also

$$(iii) \qquad A(\Omega) = 1. \tag{1.7}$$

We note that the series in Eq. (1.6) also converges in the ultraweak and strong topologies. Given an observable the probability that the measurement of a state ρ yields a result in the set E is taken to be

$$\text{prob}\{\rho, E\} = \text{tr}[\rho A(E)]. \tag{1.8}$$

For each state ρ this is a probability measure on Ω. If $\Omega = \mathbb{R}$ then the expected value of a state ρ is

$$\begin{aligned} E\{\rho, A(\cdot)\} &= \int_{-\infty}^{\infty} \lambda \, \text{tr}[\rho A(d\lambda)] \\ &= \text{tr}[\rho B] \end{aligned} \tag{1.9}$$

where

$$B = \int_{-\infty}^{\infty} \lambda A(d\lambda) \tag{1.10}$$

if the integral converges. The observable $A(\cdot)$ therefore determines a self-adjoint operator B, but is not determined by it.

There are many technical variations on the definition of an observable, one being given in the following lemma, and another in its proof. The assumption that Ω is compact metrisable was discussed in chapter 1 section 11.

Lemma 1.2 *The weakly convergent integral formula*

$$\bar{A}(f) = \int_{\Omega} f(\omega) A(d\omega) \tag{1.11}$$

determines a one-one correspondence between the observables $A(\cdot)$ on the compact metrisable space Ω and the positive linear maps $\bar{A}: C_{\mathbb{R}}(\Omega) \to \mathscr{L}_s(\mathscr{H})$ such that $\bar{A}(1) = 1$. The observable $A(\cdot)$ is a projection-valued measure if and only if the linear map $\bar{A}$ is an algebra homomorphism.

Proof Given $\phi, \psi \in \mathscr{H}$ the map $E \to \langle A(E)\phi, \psi\rangle$ is a complex-valued measure on Ω so the integral

$$\int_{\Omega} f(\omega)\langle A(d\omega)\phi, \psi\rangle \tag{1.12}$$

converges if $f \in C_{\mathbb{R}}(\Omega)$. It defines a bounded bilinear form in ϕ and ψ, so there exists a bounded linear operator $\bar{A}(f)$ such that

$$\langle \bar{A}(f)\phi, \psi \rangle = \int_{\Omega} f(\omega)\langle A(d\omega)\phi, \psi \rangle \tag{1.13}$$

for all ϕ, $\psi \in \mathscr{H}$. The properties of $\bar{A}$ may be read off this formula. For example if $f \geqslant 0$ and $\phi = \psi$ then the integrand is non-negative so

$$\langle \bar{A}(f)\phi, \phi \rangle \geqslant 0 \tag{1.14}$$

for all $\phi \in \mathscr{H}$, which implies that $\bar{A}(f) \geqslant 0$.

Conversely suppose that $\bar{A}$ is given. Then for all ϕ, $\psi \in \mathscr{H}$ the map

$$f \to \langle \bar{A}(f)\phi, \psi \rangle \tag{1.15}$$

is a bounded linear functional on $C_{\mathbb{R}}(\Omega)$, so there is a complex measure $\mu_{\phi,\psi}$ on Ω such that

$$\langle \bar{A}(f)\phi, \psi \rangle = \int_{\Omega} f(\omega)\mu_{\phi,\psi}(d\omega). \tag{1.16}$$

The map ϕ, $\psi \to \mu_{\phi,\psi}$ is bounded and bilinear so for each Borel set E there is a bounded operator $A(E)$ on $\mathscr{H}$ such that

$$\mu_{\phi,\psi}(E) = \langle A(E)\phi, \psi \rangle. \tag{1.17}$$

The proof that $A(\cdot)$ is an observable is now a matter of straightforward verification.

Now suppose that $A(\cdot)$ is an observable and that $B(\Omega)$ is the space of all bounded Borel measurable functions on Ω. Then the formula

$$A^{\sim}(f) = \int_{\Omega} f(\omega)A(d\omega) \tag{1.18}$$

defines a positive linear map $A^{\sim} : B(\Omega) \to \mathscr{L}_s(\mathscr{H})$ which extends $\bar{A}$. Moreover by the Lebesgue dominated convergence theorem $A^{\sim}$ is σ-continuous in the sense that if $f_n \in B(\Omega)$ is a uniformly bounded sequence which converges pointwise to f then $A^{\sim}(f_n)$ converges in the weak operator topology to $A^{\sim}(f)$.

If $A(\cdot)$ is a projection-valued measure then it is immediate by linearity that

$$A^{\sim}(fg) = A^{\sim}(f)A^{\sim}(g) \tag{1.19}$$

whenever f, $g \in B(\Omega)$ are simple functions, that is take only a finite number of different values on Ω. Since simple functions are uniformly dense in $B(\Omega)$ it follows that Eq. (1.19) also holds for all $f, g \in B(\Omega)$ so that $A^{\sim}$ and *a fortiori* $\bar{A}$ are algebra homomorphisms.

Conversely if $\bar{A}$ is an algebra homomorphism and $f \in C_{\mathbb{R}}(\Omega)$ let

$$S = \{g \in B(\Omega) : A^{\sim}(fg) = A^{\sim}(f)A^{\sim}(g)\}. \tag{1.20}$$

Then S contains $C_{\mathbb{R}}(\Omega)$ and by the σ-continuity of $A^{\sim}$ is closed under the bounded pointwise convergence of sequences, and so equals $B(\Omega)$. Now if $g \in B(\Omega)$ let

$$T = \{f \in B(\Omega) : A^{\sim}(fg) = A^{\sim}(f)A^{\sim}(g)\}. \tag{1.21}$$

Then T contains $C_{\mathbb{R}}(\Omega)$ and must similarly equal $B(\Omega)$. We have now proved that $A^{\sim}$ is an algebra homomorphism. The observation that

$$A(E) = A^{\sim}(\chi_E) \tag{1.22}$$

where the characteristic function χ_E of E satisfies $\chi_E^2 = \chi_E$ completes the proof.
QED

We finally note that an observable could also be defined as a complex linear map $A' : C(\Omega) \to \mathscr{L}(\mathscr{H})$ such that $A'(f) = A'(\bar{f})^*$ for all $f \in C(\Omega)$, $A'(f) \geqslant 0$ if $f \geqslant 0$ and $A'(1) = 1$. In view of all the above equivalences we shall henceforth use the same symbol for A, $\bar{A}$, $A^{\sim}$ and A', hoping that this will not cause any confusion.

3.2 Simultaneous measurement

Let us suppose that Ω_1 and Ω_2 are compact metrisable spaces with Borel σ-fields $\mathscr{F}_1$ and $\mathscr{F}_2$ and that $\Omega_1 \times \Omega_2$ is their topological product with Borel σ-field $\mathscr{F}$. If one can measure together two quantities which take their values in Ω_1 and Ω_2 then the composite measurement should be describable by an observable $A(\cdot)$ on $\Omega_1 \times \Omega_2$. For every such observable one can define a marginal observable $A_1(\cdot)$ on Ω_1 by

$$A_1(E) = A(E \times \Omega_2) \tag{2.1}$$

with a similar definition for $A_2(\cdot)$ on Ω_2.

In the example in chapter 2 section 1 we defined the observable A on $\mathbb{Z} \times \mathbb{Z}$ by

$$A(E) = \sum \{P_i Q_j P_i : (i, j) \in E\}. \tag{2.2}$$

By a straight computation one finds that

$$A_1(E) = \sum_{i \in E} P_i = P(E) \tag{2.3}$$

for all $E \subseteq \mathbb{Z}$, while $A_2(E)$ cannot generally be given by such a simple expression.

The following theorem is only one of a large number of technically different proofs of the impossibility of simultaneous measurement in the conventional sense of non-commuting observables.

Theorem 2.1 *Let A be an observable on $\Omega_1 \times \Omega_2$ whose marginal observables are both projection-valued measures. Then A is a projection-valued measure and the marginal observables commute.*

Proof If $E_1 \subseteq \Omega_1$ and $E_2 = \Omega_1 \backslash E_1$ and $F_1 \subseteq \Omega_2$ and $F_2 = \Omega_1 \backslash F_1$ we define

$$A_{ij} = A(E_i \times F_j). \tag{2.4}$$

We also put

$$P_i = A(E_i \times \Omega_2), \qquad Q_j = A(\Omega_1 \times F_j) \tag{2.5}$$

so that P_i, Q_j are projections. Then $0 \leqslant A_{1j} \leqslant P_1$ so A_{1j} commutes with P_1 and $0 \leqslant A_{2j} \leqslant P_2$ so A_{2j} commutes with $P_2 = 1 - P_1$ and hence also with P_1. Therefore $Q_j = A_{1j} + A_{2j}$ commutes with P_1. Since $0 \leqslant A_{ij} \leqslant P_i$ and $0 \leqslant A_{ij} \leqslant Q_j$ and P_i commutes with Q_j it follows that $0 \leqslant A_{ij} \leqslant P_i Q_j$.

Hence

$$1 = A_{11} + A_{12} + A_{21} + A_{22} \leqslant P_1 Q_1 + P_1 Q_2 + P_2 Q_1 + P_2 Q_2 = 1 \tag{2.6}$$

which implies that

$$A_{ij} = P_i Q_j = Q_j P_i. \tag{2.7}$$

Since the class of all sets D in $\Omega_1 \times \Omega_2$ for which $A(D)$ is a projection contains all sets of the form $D = \bigcup_{r=1}^{n} E_r \times F_r$, it follows that it coincides with $\mathscr{F}$. QED

3.3 Approximate position measurements

In preparation for defining a joint observable for position and momentum on phase space, we introduce the idea of an approximate position observable. For simplicity we take $\mathscr{H} = L^2(\mathbb{R})$, although the ideas can easily be extended to include spin and to any dimension. The usual position observable is the projection-valued measure $Q(\cdot)$ on $\mathbb{R}$ defined by

$$\{Q(E)\psi\}(x) = \chi_E(x)\psi(x) \tag{3.1}$$

where χ_E is the characteristic function of E.

Now let f be a probability density on $\mathbb{R}$, that is a non-negative measurable function on $\mathbb{R}$ such that

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{3.2}$$

Then for any bounded measurable function g on $\mathbb{R}$ the convolution

$$(f \circ g)(x) = \int_{-\infty}^{\infty} f(y)g(x - y)dy \tag{3.3}$$

is a bounded continuous function on $\mathbb{R}$.

Theorem 3.1 *The weakly convergent integral*

$$Q_f(E) = \int_{-\infty}^{\infty} (f \circ \chi_E)(x)Q(dx) \tag{3.4}$$

defines an observable on $\mathbb{R}$. For any sets E and F in $\mathbb{R}$ the operators $Q_f(E)$ and $Q_f(F)$ commute.

Proof This is a trivial corollary of Lemma 1.2 and its proof.

The observable Q_f will be called an approximate position observable, the approximation depending on the function f. The usual position observable is the limiting case where f is the Dirac delta function. If f has support in $\{x \in \mathbb{R}: |x| \leqslant \delta\}$ then every point of the support of $f \circ \chi_E$ lies within a distance δ of E. The continuous function $f \circ \chi_E$ is therefore in some sense close to χ_E so $Q_f(E)$ is close to $Q(E)$.

Recall that the mean and variance of the state ρ are defined by

$$E\{\rho, Q_f\} = \int_{-\infty}^{\infty} x \operatorname{tr}[\rho Q_f(dx)], \tag{3.5}$$

and

$$\operatorname{var}\{\rho, Q_f\} = \int_{-\infty}^{\infty} x^2 \operatorname{tr}[\rho Q_f(dx)] - \left\{\int_{-\infty}^{\infty} x \operatorname{tr}[\rho Q_f(dx)]\right\}^2 \tag{3.6}$$

provided these integrals converge.

Theorem 3.2 *Suppose that the probability density f on $\mathbb{R}$ has finite mean and variance, and let $\rho = |\psi\rangle\langle\psi|$ where the unit vector ψ lies in Schwartz space. Then ρ has finite mean and variance for Q and Q_f and*

$$E\{\rho, Q_f\} = E\{\rho, Q\} + E(f), \tag{3.7}$$

$$\operatorname{var}\{\rho, Q_f\} = \operatorname{var}\{\rho, Q\} + \operatorname{var}(f). \tag{3.8}$$

Proof We see by direct computation that

$$
\begin{aligned}
E\{\rho, Q_f\} &= \int_{-\infty}^{\infty} x\{f \circ (|\psi|^2)\}(x)dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y)f(y)|\psi(x)|^2\, dx\, dy \\
&= E\{\rho, Q\} + E\{f\}
\end{aligned}
\tag{3.9}
$$

while

$$
\begin{aligned}
\int_{-\infty}^{\infty} x^2 \operatorname{tr}[\rho Q_f(dx)] &= \int_{-\infty}^{\infty} x^2\{f \circ (|\psi|^2)\}(x)dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y)^2 f(y)|\psi(x)|^2\, dx\, dy \\
&= \int_{-\infty}^{\infty} x^2 |\psi(x)|^2\, dx + \int_{-\infty}^{\infty} y^2 f(y)dy \\
&\quad + 2\int_{-\infty}^{\infty} x|\psi(x)|^2\, dx \int_{-\infty}^{\infty} yf(y)dy
\end{aligned}
\tag{3.10}
$$

which yields Eq. (3.8) after a little rearrangement. QED

The above theorem shows that if f has zero expectation, then every state has the same expectation for Q and Q_f, but that the variance of ρ with respect to Q_f is always greater than its "true" variance with respect to Q.

We finally comment that the idea of an approximate momentum observable is defined in an exactly analogous manner.

3.4 Observables on phase space

The following theorem justifies our introduction above of approximate position and momentum observables.

Theorem 4.1 *If $\mathscr{H} = L^2(\mathbb{R})$ there exists an observable A on $\mathbb{R} \times \mathbb{R}$ whose marginal observables A_1 and A_2 are respectively approximate position and momentum observables.*

Proof Let $\alpha \in L^2(\mathbb{R})$ be a vector of norm one whose expectation is zero with respect to P and Q where

$$(P\psi)(q) = -i\psi'(q), \qquad (Q\psi)(q) = q\psi(q). \tag{4.1}$$

If we define

$$\alpha_{xy}(q) = e^{iyq}\alpha(q - x) \tag{4.2}$$

then an elementary computation shows that

$$\langle P\alpha_{xy}, \alpha_{xy}\rangle = y, \qquad \langle Q\alpha_{xy}, \alpha_{xy}\rangle = x \tag{4.3}$$

so that α_{xy} can be considered as "localised" at the point (x, y) in phase space.

If $\rho \in V$ we define the continuous function $\rho(x, y)$ on phase space by

$$\rho(x, y) = \frac{1}{2\pi}\langle \rho\alpha_{xy}, \alpha_{xy}\rangle \tag{4.4}$$

noting that if $\rho \geqslant 0$ then $\rho(x, y) \geqslant 0$ for all $x, y \in \mathbb{R}$. We claim that if ρ is a normalised state then $\rho(x, y)$ is a probability density on $\mathbb{R}^2$.

By Eq. (1.4.2) it is sufficient to prove this when $\rho = |\psi\rangle\langle\psi|$ and $\|\psi\| = 1$. In this case

$$\int_{\mathbb{R}^2} \rho(x, y) dx\, dy = \int_{\mathbb{R}^2} |(2\pi)^{-1/2}\langle\psi, \alpha_{xy}\rangle|^2\, dy\, dx. \tag{4.5}$$

Now

$$(2\pi)^{-1/2}\langle\psi, \alpha_{xy}\rangle = (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-iyq}\{\psi(q)\overline{\alpha(q - x)}\}dq. \tag{4.6}$$

Therefore by the Plancherel theorem

$$\begin{aligned}\int_{\mathbb{R}^2} |(2\pi)^{-1/2}\langle\psi, \alpha_{xy}\rangle|^2\, dy\, dx &= \int_{\mathbb{R}^2} |\psi(q)\overline{\alpha(q - x)}|^2\, dq\, dx \\ &= \int_{\mathbb{R}^2} |\psi(q)|^2|\alpha(x)|^2\, dq\, dx = 1. \end{aligned}\tag{4.7}$$

We come next to the definition of the POV measure $A(\cdot)$. For every Borel set $E \subseteq \mathbb{R}^2$ the map

$$\rho \to \int_E \rho(x, y) dx\, dy \tag{4.8}$$

is a positive bounded linear functional on V. Therefore there exists a positive linear operator $A(E)$ on $\mathscr{H}$ such that

$$\mathrm{tr}[\rho A(E)] = \int_E \rho(x, y) dx\, dy \tag{4.9}$$

for all $\rho \in V$. It is straightforward to check that $A(\cdot)$ is an observable. Note that one may also define

$$A(E) = (2\pi)^{-1}\int_E |\alpha_{xy}\rangle\langle\alpha_{xy}| dx\, dy \tag{4.10}$$

where this integral converges in the weak operator topology. Moreover Eq. (4.7) is equivalent to

$$1 = (2\pi)^{-1} \int_{\mathbb{R}^2} |\alpha_{xy}\rangle\langle\alpha_{xy}| dx\, dy. \tag{4.11}$$

This leaves only the calculation of the marginal observables. If $\psi \in L^2(\mathbb{R})$ and E is a Borel set in $\mathbb{R}$ then

$$\begin{aligned}
\langle A_1(E)\psi, \psi\rangle &= \mathrm{tr}[|\psi\rangle\langle\psi| A(E \times \mathbb{R})] \\
&= (2\pi)^{-1} \int_{x\in E} \int_{y\in\mathbb{R}} |\langle\psi, \alpha_{xy}\rangle|^2\, dy\, dx \\
&= \int_{x\in E} \int_{q\in\mathbb{R}} |\psi(q)\overline{\alpha(q-x)}|^2\, dq\, dx \\
&= \int_{\mathbb{R}^2} |\psi(q)|^2 \chi_E(x) |\alpha(q-x)|^2\, dq\, dx \\
&= \int_{\mathbb{R}} |\psi(q)|^2 (\chi_E \circ |\alpha|^2)(q) dq
\end{aligned} \tag{4.12}$$

so

$$A_1(E) = \int_{\mathbb{R}} (\chi_E \circ |\alpha|^2)(q) Q(dq). \tag{4.13}$$

The proof that

$$A_2(F) = \int_{\mathbb{R}} (\chi_F \circ |\hat{\alpha}|^2)(k) P(dk) \tag{4.14}$$

is similar. QED

The observable A constructed above depends on the unit vector α, which can be regarded as specifying certain properties and limitations of the apparatus constructed to make the joint measurement of position and momentum. The most important of the limitations is the Heisenberg uncertainty principle, which takes the following form (with the normalisation $\hbar = 1$).

Theorem 4.2 *If A is the joint observable for position and momentum constructed above with marginal observables A_1 and A_2 then for every state ρ on $\mathscr{H}$*

$$\mathrm{var}\{\rho, A_1\} \cdot \mathrm{var}\{\rho, A_2\} \geqslant \tfrac{1}{4}. \tag{4.15}$$

Proof By Theorem 3.2

$$\operatorname{var}\{\rho, A_1\} \geqslant \int_{\mathbb{R}} q^2 |\alpha(q)|^2 \, dq - \left\{ \int_{\mathbb{R}} q |\alpha(q)|^2 \, dq \right\}^2 \tag{4.16}$$

and

$$\operatorname{var}\{\rho, A_2\} \geqslant \int_{\mathbb{R}} k^2 |\hat{\alpha}(k)|^2 \, dk - \left\{ \int_{\mathbb{R}} k |\hat{\alpha}(k)|^2 \, dk \right\}^2 \tag{4.17}$$

from which the result follows from the more usual version of the uncertainty principle. QED

We finally remark that Theorem 4.1 may be used to define generalised sine and cosine operators

$$C = \int x(x^2 + y^2)^{-1/2} A(dx\, dy), \tag{4.18}$$

$$S = \int y(x^2 + y^2)^{-1/2} A(dx\, dy). \tag{4.19}$$

These do not commute, although they satisfy

$$-1 \leqslant C \leqslant 1, \qquad -1 \leqslant S \leqslant 1 \tag{4.20}$$

and the operator $(C + iS)$ is not unitary. In spite of this these operators have aroused a lot of interest, in the case where α is chosen to be the ground state of the harmonic oscillator Hamiltonian, because of their simple expressions in terms of the eigenstates of the harmonic oscillator Hamiltonian. The necessary computation is most easily done by using the ideas of chapter 8.

3.5 The Wigner distribution

The above ideas are closely related to the Wigner distribution, which however contravenes the fundamental requirement that all probability distributions should be non-negative. It seems that all proposals for simultaneous measurement of position and momentum either introduce "negative probabilities", or have some degree of arbitrariness—in our case the function α, which is characteristic of the apparatus.

We continue to consider the probability distributions on phase space constructed in Theorem 4.1. We start by defining the Weyl operator $W(u, v)$ on $L^2(\mathbb{R})$ by

$$\begin{aligned} \{W(u, v)\psi\}(x) &= (e^{iuQ + ivP}\psi)(x) \\ &= e^{iuv/2}\, e^{ixu}\psi(x + v). \end{aligned} \tag{5.1}$$

Theorem 5.1 *If $\rho(x, y)$ is the probability density on phase space associated with the state ρ by means of the unit vector α then*

$$\int_{\mathbb{R}^2} \rho(x, y)\, e^{ixu+iyv}\, dx\, dy = \operatorname{tr}[\rho W(u, v)] \cdot \langle \alpha, W(u, v)\alpha \rangle. \tag{5.2}$$

Proof If $\rho = |\psi\rangle\langle\psi|$ and $\hat{\psi}$ is the Fourier transform of ψ then by Eq. (4.6)

$$(2\pi)^{-1/2}\, e^{ixy}\langle \psi, \alpha_{xy} \rangle = (2\pi)^{-1} \int_{\mathbb{R}^2} e^{-ixu-iyv} h(u, v)du\, dv \tag{5.3}$$

where

$$h(u, v) = \hat{\psi}(-u)\overline{\alpha(v)}\, e^{-iuv}. \tag{5.4}$$

This and all subsequent formulae have to be interpreted in the appropriate sense determined by Fourier analysis; in this case $h \in L^2(\mathbb{R}^2)$ so the integral converges in the mean square sense.

Because the Fourier transform of a convolution is a product

$$\rho(x, y) = (2\pi)^{-2} \int_{\mathbb{R}^2} e^{-ixu-iyv} k(u, v)du\, dv \tag{5.5}$$

where

$$k(u, v) = \int_{\mathbb{R}^2} h(-u_0, -v_0)h(u - u_0, v - v_0)du_0\, dv_0 \tag{5.6}$$

and so

$$k(u, v) = \int_{\mathbb{R}^2} \rho(x, y)\, e^{ixu+iyv}\, du\, dv \tag{5.7}$$

this integral converging absolutely because $\rho \in L^1(\mathbb{R}^2)$.

Now

$$\begin{aligned} k(u, v) &= \int_{\mathbb{R}^2} \overline{\hat{\psi}(u_0)}\alpha(-v_0)e^{iu_0v_0}\hat{\psi}(u_0 - u) \\ &\qquad \times \overline{\alpha(v - v_0)}e^{-i(u-u_0)(v-v_0)}\, du_0\, dv_0 \\ &= E(u, v)F(u, v) \end{aligned} \tag{5.8}$$

where

$$\begin{aligned} E(u, v) &= \int_{\mathbb{R}} e^{-iuv/2}\overline{\hat{\psi}(u_0)}\hat{\psi}(u_0 - u)\, e^{iu_0v}\, du_0 \\ &= \int_{\mathbb{R}} e^{iuv/2}\, e^{ixu}\psi(x + v)\overline{\psi(x)}\, dx \\ &= \langle W(u, v)\psi, \psi \rangle \end{aligned} \tag{5.9}$$

and

$$
\begin{aligned}
F(u, v) &= \int_{\mathbb{R}} e^{-iuv/2} \alpha(-v_0) \overline{\alpha(v - v_0)}\, e^{iuv_0}\, dv_0 \\
&= \int_{\mathbb{R}} \alpha(v_0) \{ e^{iuv/2} \alpha(v + v_0) e^{iuv_0} \}^{-}\, dv_0 \\
&= \langle \alpha, W(u, v) \alpha \rangle. \qquad (5.10)
\end{aligned}
$$

This concludes the proof when ρ is a pure state. The general result follows by use of Eq. (1.4.2). QED

The Wigner distribution $\rho^{\sim}$ of a state ρ is defined formally by

$$\int_{\mathbb{R}^2} \rho^{\sim}(x, y) e^{ixu + iyv}\, dx\, dy = \mathrm{tr}[\rho\, e^{iuQ + ivP}]. \qquad (5.11)$$

$\rho^{\sim}$ is, however, not generally a probability density.

If σ is defined by

$$\int_{\mathbb{R}^2} \sigma(x, y) e^{ixu - iyv}\, dx\, dy = \langle \alpha, W(u, v) \alpha \rangle \qquad (5.12)$$

then upon taking Fourier transforms of Eq. (5.2) one obtains, subject to convergence of the integrals,

$$\rho(x, y) = \int_{\mathbb{R}^2} \rho^{\sim}(x - x_0, y - y_0) \sigma(x_0, y_0) dx_0\, dy_0 \qquad (5.13)$$

so that the probability density ρ is the result of averaging the improper but canonical distribution $\rho^{\sim}$ by the function σ, which depends on the arbitrarily chosen unit vector α.

Notes

1 This definition of an observable was introduced by Davies and Lewis (1970) and Holevo (1972, 1973).

2 For other proofs of the impossibility of simultaneous measurement of non-commuting observables see von Neumann (1955), Varadarajan (1962, 1968), Urbanik (1961), Gudder (1968) and Cohen (1966). The relations between several of these are discussed by Park and Margenau (1968).

3 This definition of an approximate position observable was given in Davies (1970a).

4 The idea of Theorem 4.1 is due to Husimi (1940) but it was related to the theory of positive-operator-valued measures by Davies (1970a) and Holevo (1973). Davies (1970a) and Ford and Lewis (1974) also established its connection with the question of repeated measurement. Husimi's ideas were variously rediscovered and developed in (Gordon and Louisell, 1966; Kano, 1964; Kuryshkin, 1972, 1973; McKenna and Frisch, 1966; Moyal, 1949; She and Heffner, 1966; Urbanik, 1961). The uncertainty

principle in this form was given by Kuryshkin (1972, 1973) and Davies (1970a). For the generalised sine and cosine operators see (Carruthers and Nieto, 1968; Lerner *et al.*, 1970; Mathews and Eswaran, 1974) and further references there.

5 The Wigner distribution appeared first in Wigner (1932). The references of the last section are only a small selection from the enormous literature on the subject.

4

Operation-Valued Measures

4.1 Operations and measures

If V is the state space of a Hilbert space $\mathscr{H}$ we shall use two topologies on the space $\mathscr{L}(V)$ of bounded linear maps from V to itself. The first is the norm topology and the second is the strong topology, for which we write $T_n \xrightarrow{s} T$ if

$$\lim_n \|T_n\rho - T\rho\|_{\mathrm{tr}} = 0 \tag{1.1}$$

for all $\rho \in V$.

We have already defined an operation on V as an element T of the cone $\mathscr{L}^+(V)$ of all positive linear maps on V which satisfies the further condition

$$0 \leqslant \mathrm{tr}[T\rho] \leqslant \mathrm{tr}[\rho] \tag{1.2}$$

for all $\rho \in V^+$. If T_1 and T_2 are operations then so is $T_1 T_2$ where

$$(T_1 T_2)(\rho) = T_1\{T_2(\rho)\} \tag{1.3}$$

so the set $\mathscr{O}$ of operations is a semigroup. It is also a convex set in $\mathscr{L}(V)$.

Definition 1.1 *If Ω is a set with a σ-field $\mathscr{F}$ we define a positive-map-valued (PMV) measure on Ω to be a map $\mathscr{E}: \mathscr{F} \to \mathscr{L}^+(V)$ such that*

(i) $$\mathscr{E}(E) \geqslant \mathscr{E}(\varnothing) = 0 \text{ for all } E \in \mathscr{F}; \tag{1.4}$$

(ii) *if $\{E_n\}$ is a countable collection of disjoint sets in $\mathscr{F}$ then*

$$\mathscr{E}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathscr{E}(E_n) \tag{1.5}$$

where the sum is convergent in the strong topology. We say that the PMV measure $\mathscr{E}$ is an instrument if it satisfies the further condition

(iii) $$\mathrm{tr}[\mathscr{E}(\Omega)\rho] = \mathrm{tr}[\rho] \tag{1.6}$$

for all $\rho \in V$.

The definition of an instrument corresponds to the idea of a measurement which accepts a state, measures some property, and emits an output state conditional on the value observed. It is a stronger notion than that of an observable. As with observables there are many technical variations of the definition. By Lemma 2.2.2 we might, for example, suppose that instruments act on $\mathscr{L}_s(\mathscr{H})$ instead of V. In the following theorem we suppose that Ω is a compact metrisable space and $\mathscr{F}$ is the σ-field of Borel sets.

Theorem 1.2 *The formula*

$$\bar{\mathscr{E}}(f, \rho) = \int_{\Omega} f(\omega)\mathscr{E}(d\omega)\rho \tag{1.7}$$

sets up a one-one correspondence between the PMV measures on Ω and the positive bilinear maps

$$\bar{\mathscr{E}}: C_{\mathbb{R}}(\Omega) \times V \to V. \tag{1.8}$$

Moreover $\mathscr{E}$ is an instrument if and only if

$$\mathrm{tr}[\bar{\mathscr{E}}(1, \rho)] = \mathrm{tr}[\rho] \tag{1.9}$$

for all $\rho \in V$.

Proof It is a matter of straightforward computation to check that if $\mathscr{E}$ is a PMV measure then $\bar{\mathscr{E}}$ is a bilinear map which is positive in the sense that if $f \geqslant 0$ and $\rho \geqslant 0$ then $\bar{\mathscr{E}}(f, \rho) \geqslant 0$.

Conversely suppose that we are given $\bar{\mathscr{E}}$. Then for all $\rho \in V^+$ and $A \in \mathscr{C}_s(\mathscr{H})^+$ the map

$$f \to \mathrm{tr}[\bar{\mathscr{E}}(f, \rho)A] \tag{1.10}$$

is a positive linear map so there exists a measure $\mu_{\rho, A}$ on Ω such that

$$\mathrm{tr}[\bar{\mathscr{E}}(f, \rho)A] = \int_{\Omega} f(\omega)\mu_{\rho, A}(d\omega). \tag{1.11}$$

For each $\rho \in V^+$ and $E \in \mathscr{F}$ the map

$$A \to \mu_{\rho, A}(E) \tag{1.12}$$

is positive and linear on $\mathscr{C}_s(\mathscr{H})$ so there exists an element of V^+ which we call $\mathscr{E}_E(\rho)$ such that

$$\mu_{\rho, A}(E) = \mathrm{tr}[A\mathscr{E}_E(\rho)]. \tag{1.13}$$

For each $E \in \mathscr{F}$ and $A \in \mathscr{C}_s(\mathscr{H})^+$ the map

$$\rho \to \mu_{\rho, A}(E) \tag{1.14}$$

is linear so the map

$$\rho \to \mathscr{E}_E(\rho) \tag{1.15}$$

is positive and linear on V.

We have now shown that for all $f \in C_{\mathbb{R}}(\Omega)^+$, $\rho \in V^+$ and $A \in \mathscr{C}_s(\mathscr{H})^+$

$$\operatorname{tr}[\bar{\mathscr{E}}(f, \rho)A] = \int_\Omega f(\omega)\operatorname{tr}[A\mathscr{E}(d\omega)\rho] \tag{1.16}$$

so that

$$\bar{\mathscr{E}}(f, \rho) = \int_\Omega f(\omega)\mathscr{E}(d\omega)\rho \tag{1.17}$$

where the measure $E \to \mathscr{E}(E)\rho$ is countably additive in the weak* topology of V for each $\rho \in V^+$. By Eq. (1.3.9) $\mathscr{E}(\cdot)\rho$ is actually norm countably additive for all $\rho \in V^+$. The verification of the last statement of the theorem is trivial.

QED

Theorem 1.3 *If $\mathscr{E}$ is an instrument on Ω with values in $V = \mathscr{T}_s(\mathscr{H})$ then there is a unique observable $A(\cdot)$ on Ω such that*

$$\operatorname{tr}[\rho A(E)] = \operatorname{tr}[\mathscr{E}(E)\rho] \tag{1.18}$$

for all $E \in \mathscr{F}$ and $\rho \in V$.

Proof For all $E \in \mathscr{F}$ the right-hand side determines a positive linear functional on V and hence a positive operator $A(E)$ such that Eq. (1.18) holds. The verification that $A(\cdot)$ is an observable is straightforward. QED

4.2 Composition of instruments

We wish to show that if $\mathscr{E}^1$ and $\mathscr{E}^2$ are instruments on Ω_1 and Ω_2 respectively with values in a state space V, then it is possible to define their composition, an instrument $\mathscr{E}$ on $\Omega_1 \times \Omega_2$ which represents the measurement of first $\mathscr{E}^2$ and then $\mathscr{E}^1$.

If $\Omega_1 = \Omega_2 = \mathbb{Z}$ this is easy. For each $(m, n) \in \mathbb{Z}^2$ we define the operation $\mathscr{E}_{mn}$ by

$$\mathscr{E}_{mn}(\rho) = \mathscr{E}^1_m\{\mathscr{E}^2_n(\rho)\} \tag{2.1}$$

and for a general set $E \subseteq \mathbb{Z}^2$ we define

$$\mathscr{E}(E)(\rho) = \sum \{\mathscr{E}_{mn}(\rho) : (m, n) \in E\}. \tag{2.2}$$

The verification that $\mathscr{E}$ is an instrument is straightforward.

The general result is not nearly so easy and we confine ourselves to the case where Ω_1 and Ω_2 are compact metrisable spaces. As we commented in chapter 1 section 11 this includes all physically interesting cases.

Lemma 2.1 *Let Ω_1 and Ω_2 be compact metrisable spaces and $B: C_{\mathbb{R}}(\Omega_1) \times C_{\mathbb{R}}(\Omega_2) \to \mathbb{R}$ a bilinear function which is positive in the sense that if $f_1 \geqslant 0$*

and $f_2 \geqslant 0$ then $B(f_1, f_2) \geqslant 0$. There exists a measure μ on $\Omega_1 \times \Omega_2$ such that

$$B(f_1, f_2) = \int_{\Omega_1 \times \Omega_2} f_1(\omega_1) f_2(\omega_2) \mu(d\omega_1\, d\omega_2) \tag{2.3}$$

for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$ and $f_2 \in C_{\mathbb{R}}(\Omega_2)$.

Proof If the norm of a bilinear function is defined by

$$\|B\| = \sup\{|B(f_1, f_2)|: \|f_1\| \leqslant 1 \quad \text{and} \quad \|f_2\| \leqslant 1\} \tag{2.4}$$

then by an argument similar to that of Lemma 2.2.1 every positive bilinear function is norm bounded. It is sufficient to prove the lemma in the case where $B(1, 1) = 1$, in which case μ will be a probability measure.

The set K of all positive bilinear functions such that $B(1, 1) = 1$ is a compact convex set in the weak topology, which is defined as the smallest topology of K for which all the (linear) maps on K

$$B \to B(f_1, f_2) \tag{2.5}$$

are continuous. We claim that every extreme point of K is of the form

$$B_0(f_1, f_2) = f_1(\omega_1) f_2(\omega_2) \tag{2.6}$$

for some $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$.

To prove this fix $f_2 \in C_{\mathbb{R}}(\Omega_2)^+$ and consider the positive linear functional

$$\phi: f_1 \to B_0(f_1, f_2) \tag{2.7}$$

on $C_{\mathbb{R}}(\Omega_1)$, which is given by a measure μ on Ω_1. Either μ is a constant multiple of the evaluation at some point, or there exist $h_1, h_2 \in C_{\mathbb{R}}(\Omega_1)^+$ such that $h_1 + h_2 = 1$ and the positive linear functionals ϕ_1, ϕ_2 on $C_{\mathbb{R}}(\Omega_1)$ defined by

$$\phi_i(f) = \phi(h_i f) \tag{2.8}$$

are not scalar multiples of ϕ. If we define the positive bilinear forms B_1, B_2 by

$$B_i(f_1, f_3) = B_0(h_i f_1, f_3) \tag{2.9}$$

for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$ and $f_3 \in C_{\mathbb{R}}(\Omega_2)$, then B_1, B_2 are not scalar multiples of B_0 but $B_0 = B_1 + B_2$. We have now shown that if B_0 is an extreme point of K then for every $f_2 \in C_{\mathbb{R}}(\Omega_2)^+$ there exists a point $\omega_{f_2} \in \Omega_1$ such that

$$B_0(f_1, f_2) = c(f_2) f_1(\omega_{f_2}) \tag{2.10}$$

for some constant $c(f_2)$ and all $f_1 \in C_{\mathbb{R}}(\Omega_1)$. The point ω_{f_2} is actually independent of f_2 as we see by examining the formula

$$c(f_2) f_1(\omega_{f_2}) + c(f_3) f_1(\omega_{f_3}) = c(f_2 + f_3) f_1(\omega_{f_2 + f_3}) \tag{2.11}$$

valid for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$. We have now shown that

$$B_0(f_1, f_2) = c(f_2)f_1(\omega_1) \tag{2.12}$$

for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$ and $f_2 \in C_{\mathbb{R}}(\Omega_2)$. Interchanging the roles of f_1 and f_2 we obtain

$$c(f_2) = f_2(\omega_2) \tag{2.13}$$

for some $\omega_2 \in \Omega_2$ and all $f_2 \in C_{\mathbb{R}}(\Omega_2)$.

Finally since $B \in K$ and the extreme points are contained within the compact set S of bilinear functions of the form

$$B_0(f_1, f_2) = f_1(\omega_1)f_2(\omega_2) \tag{2.14}$$

there exists by (A6) a probability measure μ on S such that B is the barycentre of μ. But this means precisely that Eq. (2.3) is satisfied. QED

Theorem 2.2 *Let $\mathscr{E}^i$ be an instrument on the compact metrisable space Ω_i for $i = 1, 2$. Then there exists a unique instrument $\mathscr{E}$ on $\Omega_1 \times \Omega_2$ such that*

$$\mathscr{E}(E_1 \times E_2)\rho = \mathscr{E}^1(E_1)\mathscr{E}^2(E_2)\rho \tag{2.15}$$

for all $\rho \in V$ and all Borel sets E_1 and E_2.

Proof Using the bilinear form definition of an instrument, Eq. (2.15) is equivalent to

$$\mathscr{E}(f_1 f_2, \rho) = \mathscr{E}^1(f_1, \mathscr{E}^2(f_2, \rho)) \tag{2.16}$$

for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$ and $f_2 \in C_{\mathbb{R}}(\Omega_2)$.

If $\rho \in V^+$ and $A \in \mathscr{C}_s(\mathscr{H})^+$ then

$$B(f_1, f_2) = \operatorname{tr}[A\mathscr{E}^1(f_1, \mathscr{E}^2(f_2, \rho))] \tag{2.17}$$

is a positive bilinear form so by Lemma 2.1 there exists a measure $\mu_{A,\rho}$ on $\Omega_1 \times \Omega_2$ such that

$$\operatorname{tr}[A\mathscr{E}^1(f_1, \mathscr{E}^2(f_2, \rho))] = \int_{\Omega_1 \times \Omega_2} f_1(\omega_1)f_2(\omega_2)\mu_{A,\rho}(d\omega_1\, d\omega_2) \tag{2.18}$$

for all $f_1 \in C_{\mathbb{R}}(\Omega_1)$ and $f_2 \in C_{\mathbb{R}}(\Omega_2)$. If $\rho \in V^+$ and E is any Borel set in $\Omega_1 \times \Omega_2$ then

$$A \to \mu_{A,\rho}(E) \tag{2.19}$$

is a positive linear functional on $\mathscr{C}_s(\mathscr{H})$ so there exists an element of V^+ which we call $\mathscr{E}(E)\rho$ such that

$$\mu_{A,\rho}(E) = \operatorname{tr}[A\mathscr{E}(E)\rho] \tag{2.20}$$

for all $A \in \mathscr{C}_s(\mathscr{H})$. Since $\rho \to \mu_{A,\rho}(E)$ is a positive linear functional for all $A \in \mathscr{C}_s(\mathscr{H})^+$, $\rho \to \mathscr{E}(E)\rho$ is a positive linear map which we call $\mathscr{E}(E)$. The verification that $\mathscr{E}$ is an instrument satisfying Eq. (2.15) is now straightforward. QED

4.3 Strongly repeatable instruments

In most of the discussions associating measurements with projection-valued measures, two fundamental conditions are imposed. The first is the "simplification" to measurements which only take a discrete set of values, the consideration of which we defer. The second is the assumption that the measurement is repeatable, in the sense that if the measurement is repeated immediately after being made, then the same value is observed both times with probability one. We are now in a position to give precise mathematical expression to these conditions.

Let $\mathscr{E}$ be an instrument on a discrete space, say the set $\mathbb{Z}$ of integers. We say $\mathscr{E}$ is strongly repeatable if it satisfies the following conditions:

(*i*) $\mathscr{E}_m\mathscr{E}_n\rho = \delta_{mn}\mathscr{E}_n\rho$ for all $\rho \in V$; (3.1)

(*ii*) if $\mathrm{tr}[\mathscr{E}_n\rho] = \mathrm{tr}[\rho]$ and $\rho \in V^+$ then $\mathscr{E}_n\rho = \rho$; (3.2)

(*iii*) If $B \in \mathscr{L}_s(\mathscr{H})^+$ and $\mathrm{tr}[B\mathscr{E}_n(\rho)] = 0$ for all $\rho \in V^+$ and all $n \in \mathbb{Z}$ then $B = 0$. (3.3)

Condition (*i*) is the repeatability hypothesis, condition (*ii*) is a hypothesis of minimum disturbance, and condition (*iii*) would be violated if for example the range of the instrument lay in some super-selection sector.

Theorem 3.1 *The formula*

$$\mathscr{E}(E)\rho = \sum_{n\in E} P_n\rho P_n \tag{3.4}$$

sets up a one-one correspondence between the strongly repeatable instruments $\mathscr{E}$ on $\mathbb{Z}$ and the projection-valued measures $P(\cdot)$ on $\mathbb{Z}$.

We first prove a technical lemma.

Lemma 3.2 *Define an order ideal F in V to be a subset of V^+ which is a cone such that if $0 \leqslant \phi \leqslant \psi \in F$ then $\phi \in F$. The formula*

$$F = \{\rho \in V^+ : \rho = P\rho = \rho P\} \tag{3.5}$$

sets up a one-one correspondence between the trace-norm closed order ideals F in V and the projections P on $\mathscr{H}$.

Proof By Eq. (1.4.2) every $\rho \in F$ has a decomposition

$$\rho = \sum_{n=1}^{\infty} \lambda_n |\psi_n\rangle\langle\psi_n|. \tag{3.6}$$

Since

$$0 \leqslant |\psi_n\rangle\langle\psi_n| \leqslant \lambda_n^{-1}\rho \tag{3.7}$$

it follows that $|\psi_n\rangle\langle\psi_n| \in F$. Since F is closed and the sum in Eq. (3.6) is convergent in the trace norm, it is clear that F is determined by its pure states. We have now only to show that

$$K = \{\psi \in \mathscr{H} : |\psi\rangle\langle\psi| \in F\} \tag{3.8}$$

is a closed linear subspace of $\mathscr{H}$ and then let P be the projection onto K.

The fact that K is closed under scalar multiplication is obvious, while the formula

$$0 \leqslant |\psi_1 + \psi_2\rangle\langle\psi_1 + \psi_2| \leqslant 2|\psi_1\rangle\langle\psi_1| + 2|\psi_2\rangle\langle\psi_2| \tag{3.9}$$

shows that it is closed under addition. If $\psi_n \to \psi$ in norm then $|\psi_n\rangle\langle\psi_n| \to |\psi\rangle\langle\psi|$ in trace norm, so K is a closed set.

The converse, that every set F of the form of Eq. (3.5) is a norm closed order ideal, is obvious. QED

Proof of theorem If $n \in \mathbb{Z}$ then the set

$$F_n = \{\rho \in V^+ : \mathscr{E}_n\rho = \rho\} \tag{3.10}$$

is a norm closed cone in V^+, which we claim is an order ideal.

If $0 \leqslant \rho_1 \leqslant \rho$ and $\rho_2 = \rho - \rho_1$ then

$$0 \leqslant \mathrm{tr}[\mathscr{E}_n(\rho_i)] \leqslant \mathrm{tr}[\rho_i] \tag{3.11}$$

but

$$\begin{aligned} \mathrm{tr}[\mathscr{E}_n(\rho_1)] + \mathrm{tr}[\mathscr{E}_n(\rho_2)] &= \mathrm{tr}[\mathscr{E}_n(\rho)] = \mathrm{tr}[\rho] \\ &= \mathrm{tr}[\rho_1] + \mathrm{tr}[\rho_2] \end{aligned} \tag{3.12}$$

Therefore $\mathrm{tr}[\mathscr{E}_n(\rho_1)] = \mathrm{tr}[\rho_1]$, and $\rho_1 \in F_n$ by Eq. (3.2).

By Lemma 3.2 there exists a closed subspace L_n of $\mathscr{H}$ such that

$$F_n = \{\rho \in V^+ : \rho = P_n\rho = \rho P_n\} \tag{3.13}$$

where P_n is the projection onto L_n.

We show that if $m \neq n$ then L_m and L_n are orthogonal. If $\eta_m \in L_m$ and $\eta_n \in L_n$ are unit vectors then

$$\eta \to \langle \mathscr{E}_m(|\eta\rangle\langle\eta|)\eta_m, \eta_m\rangle \tag{3.14}$$

is a bounded quadratic form on $\mathscr{H}$, so there exists a self-adjoint operator S on $\mathscr{H}$ such that

$$\langle \mathscr{E}_m(|\eta\rangle\langle\eta|)\eta_m, \eta_m\rangle = \langle S\eta, \eta\rangle \tag{3.15}$$

for all $\eta \in \mathscr{H}$. Now

$$\langle S\eta_m, \eta_m\rangle = 1 \tag{3.16}$$

and $0 \leqslant S \leqslant 1$, so $S\eta_m = \eta_m$. On the other hand

$$\begin{aligned} 1 &= \mathrm{tr}[|\eta_n\rangle\langle\eta_n|] \\ &= \mathrm{tr}[\mathscr{E}_n(|\eta_n\rangle\langle\eta_n|)] \\ &= \mathrm{tr}\left[\sum_{r=-\infty}^{\infty} \mathscr{E}_r(|\eta_n\rangle\langle\eta_n|)\right] \\ &= \mathrm{tr}[\mathscr{E}_{\mathbb{Z}}(|\eta_n\rangle\langle\eta_n|)] \\ &= 1 \end{aligned} \tag{3.17}$$

so $\mathscr{E}_m(|\eta_n\rangle\langle\eta_n|) = 0$ and $\langle S\eta_n, \eta_n\rangle = 0$, which implies that $S\eta_n = 0$. But eigenvectors belonging to distinct eigenvalues of S are orthogonal so $\langle \eta_m, \eta_n\rangle = 0$.

Let Q be the projection onto the orthogonal complement of $\sum_{n\in\mathbb{Z}} \oplus L_n$. If $v \in V^+$ then $\mathscr{E}_n(\mathscr{E}_n v) = \mathscr{E}_n v$ so $\mathscr{E}_n v \in F_n$ and

$$\mathrm{tr}[Q\mathscr{E}_n(v)] = 0 \tag{3.18}$$

by Eq. (3.3) it follows that $Q = 0$ so

$$\mathscr{H} = \sum_{n\in\mathbb{Z}} \oplus L_n \tag{3.19}$$

and the projections $\{P_n\}$ determine a projection-valued measure on $\mathbb{Z}$.

It remains to determine the form of the map $\rho \to \mathscr{E}_n(\rho)$. If

$$G_n = \{\rho \in V^+ : \mathscr{E}_n(\rho) = 0\} \tag{3.20}$$

then G_n is a norm closed order ideal in V which contains F_m for all $m \neq n$. Therefore there exists a closed subspace M_n in $\mathscr{H}$ such that $M_n \supseteq L_m$ for all $m \neq n$ and $\psi \in M_n$ if and only if $\mathscr{E}_n(|\psi\rangle\langle\psi|) = 0$. Since $M_n \supseteq \sum_{m\neq n} \oplus L_m = L_n^\perp$ and $M_n \cap L_n = 0$ it follows that $M_n = L_n^\perp$. Now let e_1 and e_2 be unit vectors in L_m and L_n respectively. Then if $\alpha_1, \alpha_2 \in \mathbb{C}$

$$\begin{aligned} 0 &\leqslant |\alpha_1 e_1 + \alpha_2 e_2\rangle\langle\alpha_1 e_1 + \alpha_2 e_2| \\ &\leqslant 2\alpha_1\bar{\alpha}_1|e_1\rangle\langle e_1| + 2\alpha_2\bar{\alpha}_2|e_2\rangle\langle e_2| \end{aligned} \tag{3.21}$$

so

$$0 \leqslant \mathscr{E}_n(|\alpha_1 e_1 + \alpha_2 e_2\rangle\langle\alpha_1 e_1 + \alpha_2 e_2|)$$
$$\leqslant 2\alpha_1\bar{\alpha}_1 |e_1\rangle\langle e_1| \tag{3.22}$$

and

$$\mathscr{E}_n(|\alpha_1 e_1 + \alpha_2 e_2\rangle\langle\alpha_1 e_1 + \alpha_2 e_2|) = \gamma |e_1\rangle\langle e_1| \tag{3.23}$$

for some real number γ. The map

$$\alpha \equiv (\alpha_1, \alpha_2) \to \gamma \tag{3.24}$$

is a quadratic form on $\mathbb{C}^2$ so there exists a self-adjoint 2×2 matrix T such that

$$\gamma = \langle T\alpha, \alpha\rangle \tag{3.25}$$

for all $\alpha \in \mathbb{C}^2$. Now if $\alpha = (1, 0)$ then $\langle T\alpha, \alpha\rangle = 1$ and if $\alpha = (0, 1)$, $\langle T\alpha, \alpha\rangle = 0$. Since also $0 \leqslant T \leqslant 1$ it follows that $T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and hence that $\gamma = \alpha_1\bar{\alpha}_1$ for all $\alpha \in \mathbb{C}^2$. Therefore

$$\mathscr{E}_n(|\alpha_1 e_1 + \alpha_2 e_2\rangle\langle\alpha_1 e_1 + \alpha_2 e_2|)$$
$$= \alpha_1\bar{\alpha}_1 |e_1\rangle\langle e_1|$$
$$= P_n|\alpha_1 e_1 + \alpha_2 e_2\rangle\langle\alpha_1 e_1 + \alpha_2 e_2| P_n \tag{3.26}$$

from which it follows by Eq. (3.6) that

$$\mathscr{E}_n(\rho) = P_n \rho P_n \tag{3.27}$$

for all $\rho \in V$. QED

It might appear that a formula similar to Eq. (3.4) could also be used to obtain a repeatable instrument starting with a more general projection-valued measure. It turns out, however, that there is no way of defining the integral so that it converges. Justifying such a negative statement is rather problematical, and we have to be content with a rather indirect argument. In the discrete case if $\rho = \mathscr{E}(\mathbb{Z})\rho'$ then

$$P(E)\rho = \rho P(E) \tag{3.28}$$

for all sets $E \subseteq \mathbb{Z}$. One would expect that the same property would hold for any $\mathscr{E}$ defined by an integral version of Eq. (3.4). This is, however, not universally possible as we now show.

Theorem 3.3 *Let $d\omega$ be a continuous measure on the compact metrisable space Ω and define the projection-valued measure $P(\cdot)$ on $\mathscr{H} = L^2(\Omega)$ by*

$$\{P(E)\psi\}(x) = \chi_E(x)\psi(x) \tag{3.29}$$

for all $\psi \in \mathscr{H}$ *and all Borel sets* $E \subseteq \Omega$. *Then the only* $\rho \in V$ *such that*

$$\rho P(E) = P(E)\rho \tag{3.30}$$

for all $E \subseteq \Omega$ *is* $\rho = 0$.

Proof Let ψ be a non-zero element of $L^2(\Omega)$ such that $\rho\psi = \lambda\psi$ for some eigenvalue λ of ρ and define

$$L = \text{lin}\{P(E)\psi : E \subseteq \Omega\} \tag{3.31}$$

Since $\psi \neq 0$ there exist two disjoint Borel sets E_1 and F_1 of non-zero measure both contained in $\{x : \psi(x) \neq 0\}$. Given the Borel set F_n of non-zero measure there exist two disjoint Borel sets E_{n+1} and F_{n+1} of non-zero measure both contained in F_n. By induction we obtain an infinite sequence $\{E_n\}$ of disjoint Borel sets of non-zero measure all contained in $\{x : \psi(x) \neq 0\}$. The vectors $P(E_n)\psi$ are then all non-zero and, having disjoint supports, are linearly independent. Therefore L is infinite-dimensional.

An easy calculation based on Eq. (3.30) shows that L is contained in the eigenspace of ρ associated with λ. On the other hand since ρ is trace class every non-zero eigenvalue has finite multiplicity. This shows that zero is the only eigenvalue of ρ, so $\rho = 0$. QED

The above results show that in measurement theory discrete and continuous projection-valued measures have very different properties. This runs completely counter to the frequent suggestion that for foundational purposes one need only consider discrete observables, the continuous ones being approximated by discrete ones in some manner.

We finally comment that although Theorem 3.1 gives precise operational conditions under which the famous formula

$$\rho \to \rho' \equiv \sum_{n=1}^{\infty} P_n \rho P_n \tag{3.32}$$

for "reduction of the wave-packet" arises, it does not pretend to solve the severe philosophical problems associated with this formula. Considerations relevant to the latter problem are presented in chapters 9 and 10.

4.4 Conditional expectations*

The ideas of the last section may be put into the framework of conditional expectations on von Neumann algebras. Let $\mathscr{V}$ be a von Neumann algebra in $\mathscr{L}(\mathscr{H})$. A normal conditional expectation of $\mathscr{L}(\mathscr{H})$ onto $\mathscr{V}$ is a linear map $\mathscr{E}$ of $\mathscr{L}(\mathscr{H})$ onto $\mathscr{V}$ such that

(*i*) $\mathscr{E}(X^*) = \mathscr{E}(X)^*$ for all $X \in \mathscr{L}(\mathscr{H})$; (4.1) (*ii*) $\mathscr{E}(X) = X$ if and only if $X \in \mathscr{V}$; (4.2)

(*iii*) if $X \geqslant 0$ then $\mathscr{E}(X) \geqslant 0$; (4.3)

(*iv*) if $X_1, X_2 \in \mathscr{V}$ and $Y \in \mathscr{L}(\mathscr{H})$ then

$$\mathscr{E}(X_1 Y X_2) = X_1 \mathscr{E}(Y) X_2; \tag{4.4}$$

(*v*) if $X_n \uparrow X$ then $\mathscr{E}(X_n) \uparrow \mathscr{E}(X)$. (4.5)

As an example let $P(\cdot)$ be a projection-valued measure on $\mathbb{Z}$ and define

$$\mathscr{V} = \{A \in \mathscr{L}(\mathscr{H}) \colon AP_n = P_n A \text{ for all } n \in \mathbb{Z}\} \tag{4.6}$$

Then $\mathscr{V}$ is a von Neumann algebra and the map $\mathscr{E}$ on $\mathscr{L}(\mathscr{H})$ defined by

$$\mathscr{E}(A) = \sum_{n \in \mathbb{Z}} P_n A P_n \tag{4.7}$$

is a normal conditional expectation of $\mathscr{L}(\mathscr{H})$ onto $\mathscr{V}$.

We suppose for the rest of this section that Ω is a compact metrisable space and that $\mathscr{H}$ is a separable Hilbert space. If $P(\cdot)$ is a projection-valued measure on Ω we say $P(\cdot)$ is discrete if there is a countable set $E \subseteq \Omega$ such that $P(\Omega \setminus E) = 0$. We say that $P(\cdot)$ is continuous if for every Borel set E in Ω such that $P(E) \neq 0$ there exists a Borel set $F \subseteq E$ such that $P(F) \neq 0$ and $P(E \setminus F) \neq 0$.

Lemma 4.1 *If $P(\cdot)$ is a projection-valued measure on Ω there is a countable set $E_0 \subseteq \Omega$ such that $E \to P(E \cap E_0)$ is a discrete projection-valued measure with values in $\mathscr{L}(P(E_0)\mathscr{H})$ and $E \to P(E \setminus E_0)$ is a continuous projection-valued measure with values in $\mathscr{L}(P(\Omega \setminus E_0)\mathscr{H})$.*

Proof Since $\mathscr{H}$ is separable there exists a countable set of unit vectors $\{\psi_n\}$ which is dense in the unit sphere of $\mathscr{H}$. If we define

$$\mu(E) = \sum_{n=1}^{\infty} 2^{-n} \langle P(E)\psi_n, \psi_n \rangle \tag{4.8}$$

then μ is a probability measure on Ω. By Eq. (1.11.10) there exists a countable set $E_0 \subseteq \Omega$ such that $E \to \mu(E \cap E_0)$ is discrete and $E \to \mu(E \setminus E_0)$ is continuous. Since E_0 is countable $E \to P(E \cap E_0)$ is a discrete projection-valued measure. We show that $E \to P(E \setminus E_0)$ is continuous.

If $P(E \setminus E_0) \neq 0$ then there exists n such that $\langle P(E \setminus E_0)\psi_n, \psi_n \rangle \neq 0$, so $\mu(E \setminus E_0) \neq 0$. Therefore there exists $F \subseteq E$ such that $\mu(F \setminus E_0) \neq 0$ and $\mu((E \setminus F) \setminus E_0) \neq 0$. But the first inequality implies that $P(F \setminus E_0) \neq 0$ while the second implies that $P((E \setminus F) \setminus E_0) \neq 0$. Therefore $E \to P(E \setminus E_0)$ is continuous. QED

Theorem 4.2 *Let $P(\cdot)$ be a projection-valued measure on Ω and let $\mathscr{V}$ be the von Neumann algebra*

$$\mathscr{V} = \{A \in \mathscr{L}(\mathscr{H}): AP(E) = P(E)A \text{ for all } E\}. \tag{4.9}$$

Then there exists a normal conditional expectation $\mathscr{E}$ of $\mathscr{L}(\mathscr{H})$ onto $\mathscr{V}$ if and only if $P(\cdot)$ is discrete.

Proof If $P(\cdot)$ is discrete then a suitable expectation $\mathscr{E}$ is given by Eq. (4.7).

If $\mathscr{E}$ exists and P is the projection associated with the set $\Omega \setminus E_0$ of Lemma 4.1 then P lies in the centre of $\mathscr{V}$. Therefore if $A \in \mathscr{L}(\mathscr{H})$ and $A = PA = AP$ it follows that $\mathscr{E}(A) = P\mathscr{E}(A) = \mathscr{E}(A)P$ and hence that $\mathscr{E}$ determines a normal conditional expectation of $\mathscr{L}(P\mathscr{H})$ onto the von Neumann algebra

$$\mathscr{W} = \{A \in \mathscr{L}(P\mathscr{H}): A = P(E)A = AP(E) \text{ for all } E\} \tag{4.10}$$

These considerations show that it is sufficient to prove that $\mathscr{E}$ cannot exist if $P(\cdot)$ is a continuous projection-valued measure.

Since $\mathscr{E}$ is positive and normal $\mathscr{E} = \mathscr{F}^*$ for some positive linear map $\mathscr{F}: V \to V$. If $\rho_0 \in V$, $\rho = \mathscr{F}\rho_0$, $A \in \mathscr{L}(\mathscr{H})$ and $P = P(E)$ for some Borel set E in Ω, then since P lies in the centre of $\mathscr{V}$

$$\begin{aligned}
&\operatorname{tr}[A\{P\rho P + (1 - P)\rho(1 - P)\}] \\
&\quad = \operatorname{tr}[\{PAP + (1 - P)A(1 - P)\}\mathscr{F}\rho_0] \\
&\quad = \operatorname{tr}[\mathscr{E}\{PAP + (1 - P)A(1 - P)\}\rho_0] \\
&\quad = \operatorname{tr}[\{P\mathscr{E}(A)P + (1 - P)\mathscr{E}(A)(1 - P)\}\rho_0] \\
&\quad = \operatorname{tr}[\mathscr{E}(A)\rho_0] \\
&\quad = \operatorname{tr}[A\rho].
\end{aligned} \tag{4.11}$$

Since A is arbitrary

$$\rho = P\rho P + (1 - P)\rho(1 - P) \tag{4.12}$$

which implies that

$$P(E)\rho = \rho P(E) \tag{4.13}$$

for all Borel sets E in Ω. By a slight variation of Theorem 3.3, $\rho = 0$. Therefore $\mathscr{F} = 0$ and so $\mathscr{E} = 0$, which contradicts Eq. (4.2). QED

4.5 Covariant instruments

One is often interested in describing measurements which are symmetric under the action of a group, such as the Euclidean or orthogonal groups. Before doing this we review some relevant ideas from group theory.

If G is a locally compact group and X a locally compact space, we call X a G-space if there is a jointly continuous map $X \times G \to X$ such that

$$(xg_1)g_2 = x(g_1g_2) \tag{5.1}$$

for all $x \in X$ and $g_1, g_2 \in G$. The map is called an action of G on X, which is said to be transitive if for all $x_1, x_2 \in X$ there exists $g \in G$ such that $x_1g = x_2$. If X is a G-space the stability subgroup H_x of a point $x \in X$ is

$$H_x = \{g \in G : xg = x\} \tag{5.2}$$

which is always closed in G. If X is a transitive G-space and H is the stability subgroup of some point x of X there is a one-one correspondence between the right cosets Hg of H and the points xg of X. If X and G are also second countable this is a homeomorphism and a G-space isomorphism of X with the set G/H of right cosets of H given its quotient topology and the induced action of G. By the definition of the quotient topology the formula

$$f^{\sim}(g) = f(Hg) \tag{5.3}$$

sets up a one-one correspondence between the continuous functions f on $X = G/H$ and the continuous functions $f^{\sim}$ on G which are constant on right cosets.

If G is a compact metrisable group and X is a transitive G-space then X is compact and there are G-invariant Borel measures dg and dx on G and X. These are unique if we suppose (as we always shall) that they have total mass one.

If G is a locally compact group *a* (strongly continuous unitary) representation of G on a Hilbert space $\mathscr{H}$ is defined as a homomorphism U from G to the group of unitary operators on $\mathscr{H}$ which is continuous for the strong operator topology of $\mathscr{L}(\mathscr{H})$. The representation is called irreducible if the only operators A such that

$$AU_g = U_gA \tag{5.4}$$

for all $g \in G$ are the multiples of the identity operator.

If X is a G-space and U a representation of G on a Hilbert space $\mathscr{H}$ we say that a POV measure $A(\cdot)$ on X with values in $\mathscr{L}_s(\mathscr{H})$ is covariant if

$$A(E_g) = U_g^*A(E)U_g \tag{5.5}$$

for all $g \in G$ and all Borel sets $E \subseteq X$. Similarly we call a PMV measure $\mathscr{E}$ on X covariant if

$$\mathscr{E}(E_g, \rho) = U_g^*\mathscr{E}(E, U_g\rho U_g^*)U_g \tag{5.6}$$

for all $\rho \in V$, $g \in G$ and $E \subseteq X$.

The special case of covariant projection-valued measures has been exhaustively investigated by group theorists under the name "system of imprimitivity", a principal result being a complete classification of all systems of imprimitivity on any transitive G-space. We shall not use these results but quote one easy consequence of the imprimitivity theorem.

Lemma 5.1 *If $P(\cdot)$, U is a transitive system of imprimitivity on the G-space X with values in $\mathscr{L}(\mathscr{H})$ and X is not finite, then $\mathscr{H}$ is infinite-dimensional.*

As an example if X is the unit sphere in $\mathbb{R}^3$ and G is the orthogonal group $SO(3)$ acting on X by rotation, there is no finite-dimensional system of imprimitivity on X.

Theorem 5.2 *If U is a representation of the compact metrisable group G on the n-dimensional Hilbert space $\mathscr{H}$ and $X = G/H$ is a transitive G-space then the formula*

$$\mathscr{E}(f, \rho) = \int_G f^{\sim}(g) U_g^* T(U_g \rho U_g^*) U_g dg \tag{5.7}$$

sets up a one-one correspondence between the covariant PMV measures $\mathscr{E}$ on X and the positive linear maps $T: V \to V$ such that

$$T(U_h \rho U_h^*) = U_h T(\rho) U_h^* \tag{5.8}$$

for all $h \in H$. Moreover $\mathscr{E}$ is an instrument if and only if

$$\mathrm{tr}[\rho] = \mathrm{tr}[T\rho] \tag{5.9}$$

for all $\rho \in V$ such that $U_g \rho = \rho U_g$ for all $g \in G$.

Proof We normalise the invariant measures on G, H, X so that they all have mass one, and let x_0 be the point of X corresponding to the coset He.

It is straightforward to show that Eq. (5.7) determines a covariant PMV measure on X for all positive linear maps $T: V \to V$. If f_n is a sequence of non-negative functions on X with $\int_X f_n(x)dx = 1$ and such that the supports of f_n decrease to the point x_0, then the non-negative functions $f_n^{\sim}$ on G determine probability measures on G which converge weakly to the Haar measure of H. Therefore

$$\lim_{n \to \infty} \mathscr{E}(f_n, \rho) = \int_H U_h^* T(U_h \rho U_h^*) U_h \, dh. \tag{5.10}$$

If T satisfies Eq. (5.8) then

$$\lim_{n \to \infty} \mathscr{E}(f_n, \rho) = T(\rho) \tag{5.11}$$

which shows that T is determined by $\mathscr{E}$.

We now have to show the existence of a representation of the said form for any given covariant PMV measure $\mathscr{E}$ on X. We start by defining a positive linear map $\theta: C_{\mathbb{R}}(G) \to C_{\mathbb{R}}(X)$ by

$$(\theta f)(Hg) = \int_H f(hg)dh \tag{5.12}$$

and then consider $\mathscr{F}: C_{\mathbb{R}}(G) \times V \to V$ defined by

$$\mathscr{F}(f, \rho) = \int_X (\theta f)(x)\mathscr{E}(dx)\rho. \tag{5.13}$$

One sees by direct verification that $\mathscr{F}$ is a PMV measure on G which is covariant in the sense that

$$\mathscr{F}(sf, \rho) = U_s\mathscr{F}(f, U_s^*\rho U_s)U_s^* \tag{5.14}$$

for all $s \in G$, where

$$(sf)(g) = f(gs) \tag{5.15}$$

for all $g, s \in G$ and all $f \in C_{\mathbb{R}}(G)$.

Since the positive linear functional

$$f \to \mathrm{tr}[\mathscr{F}(f, 1)] \tag{5.16}$$

on $C_{\mathbb{R}}(G)$ is invariant, there exists a constant α such that

$$\mathrm{tr}[\mathscr{F}(f, 1)] = \alpha \int_G f(g)dg \tag{5.17}$$

for all $f \in C_{\mathbb{R}}(G)$. If $A \in V^+$ and $f \in C_{\mathbb{R}}(G)^+$ then

$$\begin{aligned} 0 &\leqslant \mathrm{tr}[\mathscr{F}(f, A)] \\ &\leqslant \mathrm{tr}[\mathscr{F}(f, \|A\|1)] \\ &= \alpha\|A\| \int_G f(g)dg. \end{aligned} \tag{5.18}$$

Let $\{f_n\}$ be an approximate identity on G, that is a sequence of non-negative continuous functions on G with integral 1, whose supports decrease to the identity e. Let S be a countable dense set in V which is closed under sums, rational scalar multiples and the taking of the positive and negative part $\rho^{\pm}$ of any $\rho \in S$. If $\rho \in S^+ \equiv S \cap V^+$ then the sequence $\mathscr{F}(f_n, \rho)$ lies in the compact set

$$\{A \in V^+ : \mathrm{tr}[A] \leqslant \alpha\|\rho\|\} \tag{5.19}$$

and so has a convergent subsequence. Since S^+ is countable one may even find a single subsequence, which by an abuse of notation we shall again call $\{f_n\}$, such that $\mathscr{F}(f_n, \rho)$ converges as $n \to \infty$ for all $\rho \in S^+$. Since S is positively generated the limit exists for all $\rho \in S$. If we put

$$T_0(\rho) = \lim_{n\to\infty} \mathscr{F}(f_n, \rho) \tag{5.20}$$

then T_0 is a rationally-linear map from S to V which is positive and satisfies

$$\|T_0(\rho)\|_{tr} \leqslant \alpha\|\rho\| \leqslant \alpha\|\rho\|_{tr} \tag{5.21}$$

for all $\rho \in S^+$ and hence all $\rho \in S$. Since T_0 is bounded it has a unique extension to a positive linear map $T_1 : V \to V$, which has the property that

$$T_1(\rho) = \lim_{n\to\infty} \mathscr{F}(f_n, \rho) \tag{5.22}$$

for all $\rho \in V$.

If $f \in C_{\mathbb{R}}(G)$ and we define the convolution

$$(f_n \circ f)(g) = \int_G f_n(gs^{-1})f(s)ds \tag{5.23}$$

then we may also write

$$f_n \circ f = \int_G f(s)(s^{-1}f_n)ds \tag{5.24}$$

where the integral is a norm-convergent integral in $C_{\mathbb{R}}(G)$. Since $f_n \circ f$ converges uniformly to f as $n \to \infty$

$$\begin{aligned}
\mathscr{F}(f, \rho) &= \lim_{n\to\infty} \mathscr{F}(f_n \circ f, \rho) \\
&= \lim_{n\to\infty} \int_G f(s)\mathscr{F}(s^{-1}f_n, \rho)ds \\
&= \lim_{n\to\infty} \int_G f(s)U_s^*\mathscr{F}(f_n, U_s\rho U_s^*)U_s\, ds \\
&= \int_G f(s)U_s^*T_1(U_s\rho U_s^*)U_s\, ds
\end{aligned} \tag{5.25}$$

by the Lebesgue dominated convergence theorem.

Now if $f \in C_{\mathbb{R}}(X)$ then $f^{\sim} \in C_{\mathbb{R}}(G)$ is invariant under left translations by elements of H. Therefore

$$\begin{aligned}
\mathscr{E}(f, \rho) &= \int_H \int_G f^{\sim}(hs) U_s^* T_1(U_s \rho U_s^*) U_s \, ds \, dh \\
&= \int_H \int_G f^{\sim}(s) U_s^* U_h T_1(U_h^* U_s \rho U_s^* U_h) U_h^* U_s \, ds \, dh \\
&= \int_G f^{\sim}(s) U_s^* T(U_s \rho U_s^*) U_s \, ds \qquad (5.26)
\end{aligned}$$

where T, defined by

$$T(A) = \int_H U_h T_1(U_h^* A U_h) U_h^* \, dh \qquad (5.27)$$

clearly satisfies Eq. (5.8).

Finally

$$\begin{aligned}
\operatorname{tr}[\mathscr{E}(1, \rho)] &= \int_G \operatorname{tr}[U_g^* T(U_g \rho U_g^*) U_g] dg \\
&= \int_G \operatorname{tr}[T(U_g \rho U_g^*)] dg \\
&= \operatorname{tr}[T(\sigma)] \qquad (5.28)
\end{aligned}$$

where $\sigma = \int_G U_g \rho U_g^* \, dg$ is an element of V which commutes with U_g for all $g \in G$ and satisfies $\operatorname{tr}[\sigma] = \operatorname{tr}[\rho]$. If T satisfies Eq. (5.9) then

$$\operatorname{tr}[\mathscr{E}(1, \rho)] = \operatorname{tr}[\sigma] = \operatorname{tr}[\rho]. \qquad (5.29)$$

On the other hand if $\mathscr{E}$ is an instrument and $\rho \in V$ commutes with U_g for all $g \in G$ then $\rho = \sigma$, so Eq. (5.9) is satisfied. QED

Theorem 5.3 *If U is a representation of a compact metrisable group G on the finite-dimensional Hilbert space $\mathscr{H}$ and $X = G/H$ is a transitive G-space then the formula*

$$A(f) = \int_G f^{\sim}(g) U_g^* B U_g \, dg \qquad (5.30)$$

where $f \in C_{\mathbb{R}}(X)$, sets up a one-one correspondence between the covariant POV measures on X and the positive operators B on $\mathscr{H}$ such that $BU_h = U_h B$ for all $h \in H$.

Proof The quickest method is to deduce this from the previous theorem. Given the covariant POV measure A the formula

$$\mathscr{E}(f, \rho) = 1 \cdot \mathrm{tr}[A(f)\rho] \tag{5.31}$$

defines a covariant PMV measure on X. By Eq. (5.11) the map T of Theorem 5.2 satisfies

$$T(\rho) = c(\rho) \cdot 1 \tag{5.32}$$

for all $\rho \in V$, where $c(\rho)$ is a constant depending on ρ, and is actually a positive linear functional of ρ. Therefore there exists a positive operator B on $\mathscr{H}$ such that

$$T(\rho) = \mathrm{tr}[B\rho] \cdot 1 \qquad \text{QED} \tag{5.33}$$

We finally comment that one can often use Eq. (5.7) and Eq. (5.30) to construct covariant PMV and POV measures even when $\mathscr{H}$ is infinite-dimensional and G is not compact. What is, however, no longer true is that every such measure is of one of these forms.

4.6 Position measurements

In the measurement of a continuous quantity, such as position, one may suppose either that the particle is absorbed during the measurement, or that it emerges in a state perturbed by the measurement. In the second case the uncertainty principle suggests that the more accurately the position is measured the greater is the perturbation of the momentum of the outgoing state, and that there is not any canonical instrument appropriate to this situation.

A conventional way of treating this problem is to partition the position space $\mathbb{R}^3$ into a countable number of disjoint sets $\{E_n\}$ of small diameter and to take the outgoing state to be

$$\rho' = \sum_{n=1}^{\infty} P(E_n)\rho P(E_n) \tag{6.1}$$

where $P(E_n)$ is the projection associated with the set E_n. As well as depending on a rather arbitrarily chosen partition of E_n this method has the disadvantage that it destroys the translational and rotational symmetry of the original problem. We show below that it is possible to construct a mathematical model of the above situation which retains its group symmetry properties.

We take $\mathscr{H} = L^2(\mathbb{R}^3)$ to represent a single spinless particle and construct an instrument on $\mathbb{R}^3$ starting with a function α which represents the characteristics of the measuring apparatus. We define a representation U of $\mathbb{R}^3$ on $\mathscr{H}$ by translation

$$(U_y\psi)(x) = \psi(x + y) \tag{6.2}$$

where $\psi \in \mathscr{H}$. Supposing α is bounded we define the operator A_y on $\mathscr{H}$ by

$$(A_y\psi)(x) = \alpha(x - y)\psi(x) \tag{6.3}$$

where $\psi \in \mathscr{H}$.

Theorem 6.1 *If α is bounded and $\int_{\mathbb{R}^3} |\alpha(x)|^2\, d^3x = 1$ then the formula*

$$\mathscr{E}_\alpha(E, \rho) = \int_E A_x\rho A_x^*\, d^3x \tag{6.4}$$

defines an instrument on $\mathbb{R}^3$ which is covariant with respect to the translation group. The observable $A_\alpha(\cdot)$ associated with the instrument is the approximate position observable given by

$$A_\alpha(E)\psi(x) = \{\chi_E \circ (|\alpha|^2)\}(x)\psi(x) \tag{6.5}$$

for all $\psi \in \mathscr{H}$, and is also covariant.

Proof We first note that

$$A_y = U_{-y}A_0 U_y \tag{6.6}$$

for all $y \in \mathbb{R}^3$. Therefore if $a \in \mathbb{R}^3$

$$\begin{aligned}\mathscr{E}_\alpha(E + a, \rho) &= \int_{E-a} U_{-x}A_0 U_x\rho U_{-x}A_0^*U_x\, d^3x \\ &= \int_E U_{-y-a}\, A_0 U_{y+a}\rho U_{-y-a}A_0^*U_{y+a}\, d^3y \\ &= U_a^*\mathscr{E}_\alpha(E, U_a\rho U_a^*)U_a \end{aligned}\tag{6.7}$$

and $\mathscr{E}_\alpha$ is covariant. To establish that the integral is always finite, and to find the observable, we suppose $\rho = |\psi\rangle\langle\psi|$ and calculate

$$\begin{aligned}\mathrm{tr}[\mathscr{E}_\alpha(E, \rho)] &= \int_E \|A_y\psi\|^2\, d^3y \\ &= \int_{y\in E}\int_{x\in\mathbb{R}^3} |\alpha(x - y)\psi(x)|^2\, dy\, dx \\ &= \int_{\mathbb{R}^3\times\mathbb{R}^3} \chi_E(y)|\alpha(x - y)|^2|\psi(x)|^2\, dy\, dx \\ &= \int_{\mathbb{R}^3} (\chi_E \circ |\alpha|^2)(x)|\psi(x)|^2\, dx. \end{aligned}\tag{6.8}$$

If $E = \mathbb{R}^3$ then $\chi_E \circ |\alpha|^2 = 1$ so

$$\mathrm{tr}[\mathscr{E}_\alpha(\mathbb{R}^3, \rho)] = \mathrm{tr}[\rho] \tag{6.9}$$

for the pure state $\rho = |\psi\rangle\langle\psi|$ and hence for all states by Eq. (3.6). Therefore $\mathscr{E}_\alpha$ is an instrument. By Eq. (6.8) its observable is given by Eq. (6.5), and is clearly covariant. QED

We finally comment that if α is rotation invariant, then the instrument $\mathscr{E}_\alpha$ and associated observable A_α are actually covariant with respect to the full Euclidean group.

Notes

1 The convex and semigroup structure of the set of operations was studied by Kraus (1971). The definition of an instrument was given by Davies and Lewis (1970).

2 While Lemma 2.1 must surely be well known, we have not been able to find a precise reference. A version of Theorem 2.2 was first proved by Davies and Lewis (1970). W. Wils in a private communication and Wright commented that it was a consequence of general theorems on products of vector-valued measures (Thomas, 1970; Wright, 1973).

3 The ideas involved in Theorem 3.1 have been known for a long time but in this precise form it is due to Davies and Lewis (1970). Von Neumann (1955) gave a lengthy discussion of the formula for reduction of the wave-packet, which has always been a source of great interest. More recently Hepp (1972) and Frigerio (1974) justified its use for certain simple models.

4 Conditional expectations were introduced by Umegaki (1954) and Nakamura and Turumaru (1954). Condition *(iv)* in our definition was shown to be a consequence of the other conditions by Tomiyama (1957). Theorem 4.2 is a special case of results of de Korvin (1971) and Størmer (1972).

5 The classification of systems of imprimitivity was carried out in a series of papers by Mackey. For this and the general background on G-spaces we refer to Varadarajan (1970). Wightman (1962) investigated the relationship of these ideas to the problem of the non-localisability of the photon (Newton and Wigner, 1949). Jauch and Piron (1967) suggested two ways of localising the photon, firstly by using a covariant POV measure and secondly by using a projection-valued set function which was not additive. The second was followed up by Amrein (1969) while the physical acceptability of the first was stressed by Kraus (1971). Theorems 5.2 and 5.3 are special cases of theorems of Davies (1970a). A complete classification of all covariant POV measures on the transitive G-space of a Lie group G was given by Poulson (1970).

6 This material is due to Davies (1970a). It is related to the simultaneous observables on phase space in chapter 3 section 4. The ideas were further developed by Ford and Lewis (1974).

5

Statistical Properties of Quantum Fields

5.1 Measurements in quantum optics

When the foundations of quantum theory were first laid, it was usually implicitly assumed that measurements were made on single particles over such short time intervals that the wave functions did not change appreciably during the measurement. Since 1956, however, there has been a great development of interest in the statistics of photon fields, which has grown with the development of the coherent photon sources known as lasers. In this field one commonly measures quantities which by their very nature involve the development in time of a system of many particles. These features have stimulated fresh thought about the foundations of measurement theory.

Let us consider a typical experiment. If a weak beam of light is directed into a photon counter, one may record the instants at which photons arrive at the counter. The number of photons which arrive during a given period of time is an integer-valued random variable. One may also find the probability distribution of the time taken for the first photon to arrive at the counter from a given starting time. It is possible, both in principle and experimentally, that this distribution will depend on whether the starting time is specified by a clock, or chosen to be the time of arrival of some photon in the beam.

If the beam is split into two parts which are directed into different counters, there may be a statistical correlation between the photon arrivals recorded by the two counters. More complicated experiments with several counters registering in delayed coincidence give rise to a hierarchy of space-time correlation functions for the photon beam.

Experimentally the interest of these correlation functions lies in the fact that they are very different for incoherent (e.g. incandescent) light sources and for lasers. Theoretically they have been successfully explained by the use of a class of states called quasiclassical or coherent states, together with a method for calculating the relevant correlation functions.

The above developments still present a challenge from the foundational point of view because the probability distributions encountered are not related to any projection-valued measures on the relevant Fock space. In this chapter we shall describe a mathematical model for the above situations using the operational ideas already developed. Quite apart from the conceptual problems there are technical difficulties raised by the unbounded nature of boson field operators. We shall in fact assume that all relevant operators are bounded for most of the chapter, so that much of the theory will be adequate only for fermion fields. The extension to boson fields will be outlined at the end.

5.2 Definition of quantum stochastic processes

In this section we specify the requirements on any mathematical structure sufficiently complicated to accommodate the type of measurements described above. These structures, which we call quantum stochastic processes (QSPs) are then classified in section 3. Only in section 4 do we return to the particular QSPs relevant to quantum optics.

Throughout this chapter we suppose that V is the state space of a separable Hilbert space $\mathscr{H}$ and that the apparatus records events represented by points $(x, t) \in X \times (0, \infty)$, where $t > 0$ is the time at which the event occurs and the value space X provides some further information about the event, such as its location in space. We suppose that X is a compact metrisable space, which includes all physically relevant cases by chapter 1 section 11.

If $t \geqslant 0$ the sample space X_t is defined to be the set of all sample points, in turn defined as sequences $\{(x_i, t_i): 1 \leqslant i \leqslant n\}$ of events of arbitrary finite length n such that $0 < t_1 < \cdots < t_n \leqslant t$. The case $n = 0$ gives rise to the sample point z with no events. The space X_t is a locally compact second countable topological space. Given any $s, t \geqslant 0$ there is a continuous map λ of $X_s \times X_t$ onto X_{s+t} defined by

$$\lambda\{(x_i, s_i)_{i=1}^{m}, (y_j, t_j)_{j=1}^{n}\} = (z_k, u_k)_{k=1}^{m+n} \tag{2.1}$$

where $(z_k, u_k) = (y_k, t_k)$ if $1 \leqslant k \leqslant n$ and $(z_k, u_k) = (x_{k-n}, s_{k-n} + t)$ if $n < k \leqslant m + n$.

We now define a quantum stochastic process (QSP) to be a family of PMV measures $\mathscr{E}^t$ defined on X_t for $t \geqslant 0$ and satisfying

(*i*) $\operatorname{tr}[\mathscr{E}^t(X_t, \rho)] = \operatorname{tr}[\rho]$ for all $\rho \in V$ and $t \geqslant 0$; (2.2)

(*ii*) $\lim_{t \to 0} \mathscr{E}^t(X_t, \rho) = \rho$ for all $\rho \in V$; (2.3)

(*iii*) for all $\rho \in V$ and $s, t \geqslant 0$

$$\mathscr{E}^t(F, \mathscr{E}^s(E, \rho)) = \mathscr{E}^{s+t}(\lambda(F \times E), \rho). \tag{2.4}$$

The last equation states both that the evolution after time t depends only on the state at time t and that the evolution is homogeneous in time. It is analogous to the Chapman–Kolmogorov equation in probability theory. The state $\mathscr{E}^t(E, \rho)$ is the state at time t conditional upon the events having occurred up to time t determining a sample point in E.

If $\mathscr{E}$ is a QSP the formula

$$T_t(\rho) = \mathscr{E}^t(X_t, \rho) \tag{2.5}$$

defines a dynamical semigroup T_t on V. There is yet another semigroup of interest. If

$$S_t(\rho) = \mathscr{E}^t(z, \rho) \tag{2.6}$$

then S_t is a positive linear map on V for all $t \geqslant 0$ and

(i)
$$0 \leqslant \mathrm{tr}[S_t(\rho)] \leqslant \mathrm{tr}[\rho] \tag{2.7}$$

for all $\rho \in V^+$ and $t \geqslant 0$;

(ii)
$$S_t S_s(\rho) = S_{t+s}(\rho) \tag{2.8}$$

for all $\rho \in V$ and $s, t \geqslant 0$. It is not obvious however, whether this semigroup is strongly continuous.

We say that the QSP has a bounded interaction rate if there is a constant K such that

$$\mathrm{tr}[\mathscr{E}^t(X_t \setminus \{z\}, \rho)] \leqslant Kt\, \mathrm{tr}[\rho] \tag{2.9}$$

for all $\rho \in V^+$ and $t > 0$. In order to investigate the implications of this we introduce the Borel sets

$$A_t^n = \{\text{all points in } X_t \text{ with exactly } n \text{ events}\} \tag{2.10}$$

and

$$B_t^n = \bigcup_{m=n}^{\infty} A_t^m = \{\text{all points in } X_t \text{ with at least } n \text{ events}\} \tag{2.11}$$

observing that

$$A_{s+t}^q = \bigcup_{m+n=q} \lambda(A_s^m \times A_t^n). \tag{2.12}$$

Let $m \geqslant n$ and let F_m^n be the family of all subsets a of $(1, \ldots, m)$ containing exactly n points. For $a \in F_m^n$ and $1 \leqslant r \leqslant m$ define $C_{a,r} \subseteq X_{m^{-1}t}$ by

$$C_{a,r} = \begin{cases} X_{m^{-1}t} & \text{if } r \notin a \\ B_{m^{-1}t}^1 & \text{if } r \in a \end{cases} \tag{2.13}$$

then define $D^n_{t,m} \subseteq X_t$ by

$$D^n_{t,m} = \bigcup_{a \in F^n_m} \lambda(C_{a,1} \times \cdots \times C_{a,m}). \tag{2.14}$$

Clearly $D^n_{t,m} \subseteq B^n_t$; conversely the characteristic functions of those sets satisfy

$$\chi(B^n_t)(x) = \lim_{m \to \infty} \chi(D^n_{t,m})(x) \tag{2.15}$$

for all $x \in X_t$. The hypothesis of a bounded interaction rate implies that

$$\mathrm{tr}[\mathscr{E}^t(D^n_{t,m}, \rho)] \leqslant \frac{m!K^n t^n \,\mathrm{tr}[\rho]}{(m-n)!n!m^n} \tag{2.16}$$

for all $\rho \in V^+$. Therefore by Fatou's lemma

$$\mathrm{tr}[\mathscr{E}^t(B^n_t, \rho)] \leqslant \frac{K^n t^n \,\mathrm{tr}[\rho]}{n!}. \tag{2.17}$$

Since $z = \lambda(z \times \cdots \times z)$ we also can deduce from Eq. (2.9) that if $\rho \in V^+$

$$\mathrm{tr}[\mathscr{E}^t(z, \rho)] \geqslant \left(1 - \frac{Kt}{m}\right)^m \mathrm{tr}[\rho] \tag{2.18}$$

and letting $m \to \infty$ we obtain

$$\mathrm{tr}[S_t(\rho)] \geqslant e^{-Kt}\,\mathrm{tr}[\rho]. \tag{2.19}$$

Since

$$T_t(\rho) = S_t(\rho) + \mathscr{E}^t(B^1_t, \rho) \tag{2.20}$$

it follows from Eq. (2.9) that if $\rho \in V^+$ then

$$\lim_{t \to 0} S_t(\rho) = \lim_{t \to 0} T_t(\rho) = \rho \tag{2.21}$$

so S_t is a strongly continuous one-parameter semigroup on V.

The dynamical semigroup T_t has the property that for any partition $\{E_n\}_{n=1}^{\infty}$ of X_t and any $\rho \in V^+$

$$T_t(\rho) = \sum_{n=1}^{\infty} \mathscr{E}^t(E_n, \rho). \tag{2.22}$$

This suggests except in totally degenerate cases that T_t transforms pure states into mixed states and that as t increases $T_t(\rho)$ becomes more and more mixed. Such an argument does not apply to the semigroup S_t. Certainly as t increases $\rho \to S_t(\rho)$ is a transformation giving us more and more information about ρ, but this information is of rather a minimal kind, that a certain type of interaction between the quantum system and the measuring apparatus has not occurred up to time t. It is therefore reasonable to suppose that the evolution S_t is of the "simplest kind". Motivated by Theorem 2.3.1 we take this to mean that there exists a strongly continuous one parameter contraction semigroup B_t on $\mathscr{H}$ such that

$$S_t(\rho) = B_t \rho B_t^* \tag{2.23}$$

for all $\rho \in V$ and all $t \geqslant 0$. By Eq. (2.19) it is apparent that

$$e^{-Kt/2}\|\psi\| \leqslant \|B_t \psi\| \leqslant \|\psi\| \tag{2.24}$$

for all $\psi \in \mathscr{H}$ and all $t \geqslant 0$. We shall say more about the infinitesimal generator Y of B_t below.

5.3 Classification of quantum stochastic processes

We assume that $\mathscr{E}$ is a QSP satisfying Eq. (2.9) and Eq. (2.23), and continue with our previous notation. If π is the map $(x, t) \to x$ from A_t^1 onto X then we define

$$\mathscr{J}^t(E, \rho) = t^{-1}\mathscr{E}^t(\pi^{-1}E, \rho) \tag{3.1}$$

so that $\mathscr{J}^t$ is a PMV measure on the value space X such that

$$0 \leqslant \mathrm{tr}[\mathscr{J}^t(X, \rho)] \leqslant K\, \mathrm{tr}[\rho] \tag{3.2}$$

for all $\rho \in V^+$ and all $t > 0$. Equation (2.17) implies that

$$T_t(\rho) = S_t(\rho) + t\mathscr{J}^t(X, \rho) + O(t^2) \tag{3.3}$$

as $t \to 0$. We wish to show that $\mathscr{J}^t$ converges as $t \to 0$, but this is technically difficult. We shall see that it is helpful to consider $\mathscr{J}^t$ as a positive bilinear map

$$\mathscr{J}^t : C_{\mathbb{R}}(X) \times V \to V \tag{3.4}$$

which is permissible by Theorem 4.1.2.

Lemma 3.1 *If $\rho_0 \in V$ is such that*

$$t_n^{-1}(B_{t_n}\rho_0 B_{t_n}^* - \rho_0) \equiv t_n^{-1}\{S_{t_n}(\rho_0) - \rho_0\} \tag{3.5}$$

converges in the weak operator topology to a limit in V for some sequence $t_n \to 0$, then ρ_0 lies in the domain dom(W) *of the infinitesimal generator W of S_t.*

Proof Given a sequence $t_n \to 0$ we define

$$\mathscr{D}' = \{\rho \in V : t_n^{-1}\{S_{t_n}(\rho) - \rho\} \text{ converges in the weak operator topology to a limit in } V\} \tag{3.6}$$

and define $W'(\rho)$ as the said limit for all $\rho \in \mathscr{D}'$. Then $\mathscr{D}' \supseteq \mathrm{dom}(W)$ and W' is an extension of W. If we can show that $(1 - W')$ is one–one then since $(1 - W)$ maps $\mathrm{dom}(W)$ one–one onto V by (A7), it follows that $\mathrm{dom}(W) = \mathscr{D}'$.

Suppose that there exists some non-zero $\rho \in \mathscr{D}'$ such that $(1 - W')\rho = 0$, and let $\psi \in \mathscr{H}$ be such that $\langle \rho\psi, \psi \rangle \neq 0$. If we define $f: \mathbb{R}^+ \to \mathbb{R}$ by

$$f(t) = \langle S_t(\rho)\psi, \psi \rangle / \langle \rho\psi, \psi \rangle \tag{3.7}$$

then f is continuous, $f(0) = 1$, and for all $t \geqslant 0$

$$\begin{aligned} \lim_{n\to\infty} t_n^{-1}\{f(t + t_n) - f(t)\} &= \lim_{n\to\infty} \frac{\langle t_n^{-1}(S_{t_n}\rho - \rho)B_t^*\psi, B_t^*\psi \rangle}{\langle \rho\psi, \psi \rangle} \\ &= \frac{\langle W'(\rho)B_t^*\psi, B_t^*\psi \rangle}{\langle \rho\psi, \psi \rangle} \\ &= \frac{\langle S_t(\rho)\psi, \psi \rangle}{\langle \rho\psi, \psi \rangle} \\ &= f(t). \end{aligned} \tag{3.8}$$

The set $\{t \geqslant 0: f(t) \geqslant e^{t/2}\}$ is therefore non-empty, closed and has no right-hand end point. This contradicts the fact that f is a bounded function, and leads to the conclusion that $(1 - W')$ is one–one. QED

Lemma 3.2 *The domains of the infinitesimal generators Z of T_t and W of S_t are equal. Moreover $\mathscr{J}^t(X, \rho)$ converges in trace norm as $t \to 0$ for all $\rho \in V$.*

Proof If $\rho \in V$ and $\rho^{\pm}$ are defined by Eq. (1.1.3) then by Eq. (3.2) the operators $\mathscr{J}^t(X, \rho^{\pm})$ are uniformly bounded, and there exists a sequence $t_n \to 0$ such that $\mathscr{J}^{t_n}(X, \rho^{\pm})$ converge in the weak operator topology to limits $\sigma_{\pm} \in \mathscr{L}_s(\mathscr{H})^+$. (A1) and Eq. (3.2) imply that $\sigma_{\pm} \in V^+$ and

$$\mathrm{tr}[\sigma_{\pm}] \leqslant K\, \mathrm{tr}[\rho^{\pm}] \tag{3.9}$$

If $\sigma = \sigma_+ - \sigma_-$ then $\sigma \in V$, $\|\sigma\|_{\mathrm{tr}} \leqslant K\|\rho\|_{\mathrm{tr}}$ and σ is the limit in the weak operator topology of $\mathscr{J}^{t_n}(X, \rho)$ as $t_n \to 0$.

Now suppose $\rho \in \mathrm{dom}(Z)$. By the above and the formula

$$t^{-1}(T_t\rho - \rho) = t^{-1}(S_t\rho - \rho) + \mathscr{J}^t(X, \rho) + O(t) \tag{3.10}$$

there exists a sequence $t_n \to 0$ such that $t_n^{-1}(S_{t_n}\rho - \rho)$ converges in the weak operator topology to a limit in V. By Lemma 3.1 it follows that $\rho \in \mathrm{dom}(W)$. By Eq. (3.10) it now follows that $\mathscr{J}^t(X, \rho)$ converges in trace norm as $t \to 0$ if $\rho \in \mathrm{dom}(Z)$. But $\mathrm{dom}(Z)$ is dense in V so $\mathscr{J}^t(X, \rho)$ converges in trace norm to a limit in V as $t \to 0$ for all $\rho \in V$. Again by Eq. (3.10) we conclude that the domains of Z and W are equal. QED

Lemma 3.3 *There exists a sequence $t_n \to 0$ of the form $t_n = 2^{-m_n}$ and a PMV measure $\mathscr{J}$ on X such that for all $f \in C_{\mathbb{R}}(X)$ and all $\rho \in V$, $\mathscr{J}^{t_n}(f, \rho)$ converges in trace norm to $\mathscr{J}(f, \rho)$.*

Proof By the separability of $C_{\mathbb{R}}(X)$ and V, the uniform boundedness of $\mathscr{J}^{2^{-n}}$, and the compactness of

$$\{A : \mathscr{L}_s(\mathscr{H}) : \|A\| \leqslant 1\} \tag{3.11}$$

in the weak operator topology, there exists a sequence $t_n = 2^{-m_n}$ and a bilinear map

$$\mathscr{J} : C_{\mathbb{R}}(X) \times V \to \mathscr{L}_s(\mathscr{H}) \tag{3.12}$$

such that for all $f \in C_{\mathbb{R}}(X)$ and $\rho \in V$, $\mathscr{J}^{t_n}(f, \rho)$ converges in the weak operator topology to $\mathscr{J}(f, \rho)$.

If $f \in C_{\mathbb{R}}(X)^+$ and $\rho \in V^+$ then $\mathscr{J}(f, \rho) \geqslant 0$ and

$$\begin{aligned} \operatorname{tr}[\mathscr{J}(f, \rho)] &\leqslant \liminf_{n\to\infty} \operatorname{tr}[\mathscr{J}^{t_n}(f, \rho)] \\ &\leqslant K \operatorname{tr}[\rho] < \infty \end{aligned} \tag{3.13}$$

by (A1), so $\mathscr{J}$ has its range in V and is a PMV measure. Moreover if $\rho \in V^+$ and $0 \leqslant f \leqslant c1$ then

$$\begin{aligned} \operatorname{tr}[\mathscr{J}(c1, \rho)] &= \lim_{n\to\infty} \operatorname{tr}[\mathscr{J}^{t_n}(c1, \rho)] \\ &\geqslant \liminf_{n\to\infty} \operatorname{tr}[\mathscr{J}^{t_n}(f, \rho)] + \liminf_{n\to\infty} \operatorname{tr}[\mathscr{J}^{t_n}(1 - f, \rho)] \\ &\geqslant \operatorname{tr}[\mathscr{J}(f, \rho)] + \operatorname{tr}[\mathscr{J}(1 - f, \rho)] \\ &= \operatorname{tr}[\mathscr{J}(1, \rho)] \end{aligned} \tag{3.14}$$

which implies that

$$\lim_{n\to\infty} \operatorname{tr}[\mathscr{J}^{t_n}(f, \rho)] = \operatorname{tr}[\mathscr{J}(f, \rho)]. \tag{3.15}$$

Therefore by (A1), $\mathscr{J}^{t_n}(f, \rho)$ converges in trace norm to $\mathscr{J}(f, \rho)$. QED

We now define the total interaction rate R of the QSP to be the operator $R \in \mathscr{L}_s(\mathscr{H})^+$ such that for all $\rho \in V$

$$\begin{aligned} \operatorname{tr}[\rho R] &= \operatorname{tr}[\mathscr{J}(X, \rho)] \\ &= \lim_{t\to 0} \operatorname{tr}[\mathscr{J}^t(X, \rho)]. \end{aligned} \tag{3.16}$$

This name is justified by Eq. (3.18) below.

Lemma 3.4 *If $\psi \in \mathrm{dom}(Y)$ where Y is the infinitesimal generator of B_t then*

$$\langle R\psi, \psi\rangle = -2\,\mathrm{Re}\langle Y\psi, \psi\rangle. \tag{3.17}$$

If $\rho \in V$ then $t \to \mathrm{tr}[S_t(\rho)]$ is differentiable and

$$\frac{d}{dt}\mathrm{tr}[S_t(\rho)] = -\mathrm{tr}[RS_t(\rho)]. \tag{3.18}$$

Proof If $\psi \in \mathrm{dom}(Y)$ then

$$\begin{aligned}\lim_{t\to 0} & t^{-1}\{S_t(|\psi\rangle\langle\psi|) - |\psi\rangle\langle\psi|\} \\ &= \lim_{t\to 0}\{t^{-1}|B_t\psi - \psi\rangle\langle\psi| + t^{-1}|B_t\psi\rangle\langle B_t\psi - \psi|\} \\ &= |Y\psi\rangle\langle\psi| + |\psi\rangle\langle Y\psi| \end{aligned} \tag{3.19}$$

so $|\psi\rangle\langle\psi| \in \mathrm{dom}(W)$. By Eq. (3.10) it follows that

$$\langle Y\psi, \psi\rangle + \langle\psi, Y\psi\rangle + \mathrm{tr}[\mathscr{J}(X, |\psi\rangle\langle\psi|)] = 0 \tag{3.20}$$

so Eqs. (3.17) and (3.18) are proved for all states $\rho = |\psi\rangle\langle\psi|$ with $\psi \in \mathrm{dom}(Y)$. Therefore if ρ lies in

$$L = \mathrm{lin}\{|\psi\rangle\langle\psi| : \psi \in \mathrm{dom}(Y)] \tag{3.21}$$

we see that

$$\mathrm{tr}[S_t(\rho)] = \mathrm{tr}[\rho] - \int_0^t \mathrm{tr}[S_s(\rho)R]ds \tag{3.22}$$

this being the integral version of Eq. (3.18). But L is dense in V so Eq. (3.22) holds for all $\rho \in V$. Differentiating this now yields Eq. (3.18). QED

Theorem 3.5 *Let $\mathscr{E}$ be a QSP on X, V where X is a compact metrisable space and V is the state space of a separable Hilbert space $\mathscr{H}$. Let $\mathscr{E}$ satisfy Eq. (2.9) and Eq. (2.23). Then $\mathscr{E}$ is uniquely determined by the PMV measure $\mathscr{J}$ and the infinitesimal generator Y on $\mathscr{H}$, which are related by the equation*

$$\mathrm{tr}[\mathscr{J}(X, |\psi\rangle\langle\psi|)] = -2\,\mathrm{Re}\langle Y\psi, \psi\rangle \tag{3.23}$$

for all $\psi \in \mathrm{dom}(Y)$. We call Y, $\mathscr{J}$ the infinitesimal generators of the QSP $\mathscr{E}$.

Proof If $t > 0$, $f \in C_{\mathbb{R}}(X)$ and $\rho \in V$ then by Eqs (2.4) and (3.1)

$$2^{-m}[2^m t]\mathscr{J}^{2^{-m}[2^m t]}(f, \rho) = \sum_{r=1}^{[2^m t]} 2^{-m} S_{2^{-m}([2^m t]-r)} \mathscr{J}^{2^{-m}}(f, S_{2^{-m}(r-1)}\rho) \tag{3.24}$$

where $[2^m t]$ denotes the largest integer less than or equal to $2^m t$. By Lemma 3.3 letting $m = m_n$ and $n \to \infty$

$$t\mathscr{J}^t(f, \rho) = \int_0^t S_{t-s}\mathscr{J}(f, S_s\rho)ds. \tag{3.25}$$

By Theorem 4.1.2 there exists a PMV measure $\mathscr{J}$ on X such that

$$t\mathscr{J}^t(E, \rho) = \int_0^t S_{t-s}\mathscr{J}(E, S_s\rho)ds \tag{3.26}$$

for all $t \geqslant 0$, $\rho \in V$ and all Borel sets $E \subseteq X$. It is a consequence of this formula that

$$\mathscr{J}(E, \rho) = \lim_{t \to 0} \mathscr{J}^t(E, \rho) \tag{3.27}$$

in the trace norm, for all $E \subseteq X$ and $\rho \in V$, which both greatly improves on Lemma 3.3 and shows that $\mathscr{J}$ is uniquely determined by $\mathscr{E}$.

Now if $0 < r_1 < t_1 < \cdots < r_m < t_m \leqslant t$ and $\{E_i\}_{i=1}^m$ are Borel sets in X, we let $E \subseteq X_t$ be the Borel set

$$E = \{(x_i, s_i)_{i=1}^m : r_i < s_i \leqslant t_i \quad \text{and} \quad x_i \in E_i\}. \tag{3.28}$$

Then by Eq. (2.4)

$$\begin{aligned}\mathscr{E}^t(E, \rho) &= \prod_{i=1}^m (t_i - r_i)\{S_{t-t_m}\mathscr{J}_{E_m}^{t_m - r_m}S_{r_m - t_{m-1}} \cdots S_{r_2 - t_1}\mathscr{J}_{E_1}^{t_1 - r_1}S_{r_1}\rho\}\\ &= \int_{r_i < s_i \leqslant t_i} S_{t-t_m}(S_{t_m - s_m}\mathscr{J}_{E_m}S_{s_m - r_m}) \cdots\\ &\qquad S_{r_2 - t_1}(S_{t_1 - s_1}\mathscr{J}_{E_1}S_{s_1 - r_1})S_{r_1}\rho\, ds_1 \cdots ds_m\end{aligned} \tag{3.29}$$

or

$$\mathscr{E}^t(E, \rho) = \int_{r_i < s_i \leqslant t_i} S_{s-s_m}\mathscr{J}_{E_m}S_{s_m - s_{m-1}}\mathscr{J}_{E_{m-1}} \cdots \mathscr{J}_{E_1}S_{s_1}\rho\, ds_1 \cdots ds_m \tag{3.30}$$

This shows that for all $\rho \in V$, S_t and $\mathscr{J}$ determine $\mathscr{E}^t(E, \rho)$ for any set E of the form of Eq. (3.28), and hence for any Borel set E. QED

The next theorem shows that the stated conditions on Y and $\mathscr{J}$ are the only ones necessary for them to be the generators of a QSP.

Theorem 3.6 *Let X be a compact metrisable space and V the state space of a separable Hilbert space $\mathscr{H}$. Let Y be the infinitesimal generator of a strongly continuous one-parameter contraction semigroup B_t on $\mathscr{H}$ and $\mathscr{J}$ a PMV measure on X such that*

$$\mathrm{tr}[\mathscr{J}(X, |\psi\rangle\langle\psi|)] = -2\,\mathrm{Re}\langle Y\psi, \psi\rangle \tag{3.31}$$

for all $\psi \in \mathrm{dom}(Y)$. Then there exists a unique QSP $\mathscr{E}$ such that Y, $\mathscr{J}$ are its infinitesimal generators.

Proof We define $\mathscr{E}^t$ independently on each of the sets $A_t^n \subseteq X_t$. If $n = 0$ we define $\mathscr{E}^t(z, \rho)$ and $S_t(\rho)$ by

$$\mathscr{E}^t(z, \rho) = S_t(\rho) = B_t \rho B_t^* \tag{3.32}$$

for all $t \geqslant 0$ and $\rho \in V$. If $n \geqslant 1$ and $0 < t_1 < \cdots < t_n \leqslant t$ there is a unique PMV measure $\mathscr{J}_{t_1,\ldots,t_n}$ on X^n such that if $E_1, \ldots, E_n \subseteq X$

$$\mathscr{J}_{t_1,\ldots,t_n}(E_1 \times \cdots \times E_n, \rho) = S_{t-t_n}\mathscr{J}_{E_n}S_{t_n-t_{n-1}} \cdots \mathscr{J}_{E_1}S_{t_1}\rho \tag{3.33}$$

where in order to apply Theorem 4.2.2 we regard S_s as a PMV measure on a set of one element. If E is a Borel set in A_t^n and $0 < t_1 < \cdots < t_n \leqslant t$ we define $E_{t_1,\ldots,t_n} \subseteq X^n$ by

$$E_{t_1,\ldots,t_n} = \{(x_1, \ldots, x_n) : (x_r, t_r)_{r=1}^n \in E\}. \tag{3.34}$$

We then define the PMV measure $\mathscr{E}_n^t$ on A_t^n by

$$\mathscr{E}_n^t(E, \rho) = \int_{0 < t_1 < \cdots < t_n \leqslant t} \mathscr{J}_{t_1,\ldots,t_n}(E_{t_1,\ldots,t_n}, \rho)dt_1 \cdots dt_n. \tag{3.35}$$

If

$$0 \leqslant \mathrm{tr}[\mathscr{J}(X, \rho)] \leqslant K\,\mathrm{tr}[\rho] \tag{3.36}$$

for all $\rho \in V^+$ then it is immediate that

$$0 \leqslant \mathrm{tr}[\mathscr{E}_n^t(A_t^n, \rho)] \leqslant \frac{K^n t^n\,\mathrm{tr}[\rho]}{n!} \tag{3.37}$$

for all $\rho \in V^-$. We now define the PMV measure $\mathscr{E}^t$ on X_t by

$$\mathscr{E}^t(E, \rho) = \sum_{n=0}^{\infty} \mathscr{E}_n^t(E \cap A_t^n, \rho) \tag{3.38}$$

Equation (3.37) implies that this series converges. Indeed

$$0 \leqslant \mathrm{tr}[\mathscr{E}^t(X_t, \rho)] \leqslant e^{Kt}\,\mathrm{tr}[\rho] \tag{3.39}$$

for all $\rho \in V^+$.

Equation (2.4) is now immediate from Eq. (3.35). If we define $T_t: V \to V$ by

$$T_t(\rho) = \mathscr{E}^t(X_t, \rho) \tag{3.40}$$

Then it follows from Eq. (3.37) that for all $\rho \in V$

$$\begin{aligned} T_t(\rho) &= S_t(\rho) + \int_0^t S_{t-s} \mathscr{J}_X S_s \rho \, ds + O(t^2) \\ &= S_t(\rho) + t\mathscr{J}_X(\rho) + o(t) \end{aligned} \tag{3.41}$$

as $t \to 0$, which proves Eq. (2.3). Finally by Eqs (3.18) and (3.31)

$$\begin{aligned} \frac{d}{dt} \operatorname{tr}[T_t \rho]|_{t=0} &= -\operatorname{tr}[R\rho] + \operatorname{tr}[\mathscr{J}_X(\rho)] \\ &= 0 \end{aligned} \tag{3.42}$$

for all $\rho \in V$ so

$$\operatorname{tr}[T_t(\rho)] = \operatorname{tr}[\rho] \tag{3.43}$$

for all $\rho \in V$ and $t \geqslant 0$, which is Eq. (2.2). QED

5.4 Measurements on quantum fields

Now that we have obtained a general method of constructing QSPs, we return to the problem of describing counting measurements on a quantum field, which for technical reasons we take to be a fermion field.

We suppose that an individual fermion is described by a wave function in the Hilbert space

$$\mathscr{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^{2n+1} \tag{4.1}$$

and that the quantum field is represented by states on the fermion Fock space

$$\mathscr{F} = \mathbb{C} \oplus \mathscr{H} \oplus (\mathscr{H} \otimes_{\text{anti}} \mathscr{H}) \oplus \cdots \tag{4.2}$$

Let $x_1, \ldots, x_n$ be points in $\mathbb{R}^3$ representing the positions of n particle counters and let $f_1, \ldots, f_n \in \mathscr{H}$ be test functions with supports concentrated in small neighbourhoods of $x_1, \ldots, x_n$ respectively. The smeared annihilation operators $a(f_1), \ldots, a(f_n)$ are bounded operators on $\mathscr{F}$ which represent the effectiveness of the counters in absorbing particles from the field. By suitable choice of the test functions one can accommodate the possibility that the sensitivities of the counters depend on the directions and energies of the incoming particles.

The space X is taken to be the finite set $\{1, \ldots, n\}$, so that the sample space records the times at which particles arrive at each of the counters. The PMV measure $\mathscr{J}$ is defined by

$$\mathscr{J}(E, \rho) = \sum_{r \in E} a(f_r)\rho a(f_r)^* \tag{4.3}$$

The rate R is therefore the bounded operator

$$R = \sum_{r=1}^{n} a^*(f_r)a(f_r). \tag{4.4}$$

If H is the Hamiltonian describing the evolution of the quantum field in the absence of the detectors then Y is taken to be

$$Y = -iH - \tfrac{1}{2}R \tag{4.5}$$

this being the generator of a one-parameter contraction semigroup on $\mathscr{F}$ by Theorem 1.9.1 and Lemma 3.4. We can now use the theory of the last section to construct the relevant QSP on $V = \mathscr{T}_s(\mathscr{F})$. We note in passing that if $\rho_t = T_t(\rho)$ then, neglecting domain problems,

$$\rho_t' = -i[H, \rho_t] + \sum_{r=1}^{n} \{a(f_r)\rho_t a^*(f_r) - \tfrac{1}{2}a^*(f_r)a(f_r)\rho_t - \tfrac{1}{2}\rho_t a^*(f_r)a(f_r)\} \tag{4.6}$$

which is similar to the differential equation we obtained for the Wigner–Weisskopf atom in chapter 2 section 4.

The above constructions all apply formally to boson quantum fields. However, the smeared field operators $a(f_r)$ are unbounded, so R is unbounded. This means that the QSP does not have a bounded interaction rate and does not lie within the class analysed in section 3. Although such QSPs can be constructed the unboundedness of the various operators involved makes it a matter of much greater technical difficulty.

Now let us look more closely at the case where we have two identical counters placed at points x_1 and x_2 in $\mathbb{R}^3$ and let us suppose that they are associated with the annihilation operators A_1 and A_2 respectively. For an initial state ρ the output state at time $2t$ conditional upon there being only two events in this time, one at x_1 in $(0, t]$ and one at x_2 in $(t, 2t]$, is

$$t^2 A_2 A_1 \rho A_1^* A_2^* + o(t^2) \tag{4.7}$$

Now for reasons of causality we should be very surprised if the conditional output state depends on which of the events occurs first for very small t. Therefore if the QSP is of a causally sensible type we should have

$$A_2 A_1 \rho A_1^* A_2^* = A_1 A_2 \rho A_2^* A_1^* \tag{4.8}$$

for all $\rho \in V$. If the counters are non-directional we can also suppose that the QSP is covariant under the Euclidean group, which has a unitary representation on $\mathscr{F}$ induced by its representation on the single-particle space.

The above considerations are of an operational nature and should therefore apply not only to a representation of the CCRs or CARs, but to all abstract field theories. In the following theorem we suppose we are given a field theory covariant under a Lie group G and that the field operators A_x have a common dense invariant domain.

Theorem 4.1 *Let G be a Lie group and X a connected transitive G-space such that for all $x, y \in X$ there exists a $g \in G$ satisfying $xg = y$ and $yg = x$ simultaneously. Let U be a representation of G on a Hilbert space $\mathscr{F}$ and let $\mathscr{D}$ be a dense G-invariant subspace of $\mathscr{F}$ provided with a topology stronger than the norm topology and such that $g, \psi \to U_g\psi$ is jointly continuous from $G \times \mathscr{D}$ to $\mathscr{D}$. Let A_x be a family of continuous linear operators from $\mathscr{D}$ to $\mathscr{D}$ parametrized by $x \in X$ and such that for all $x \in X$ and $g \in G$*

$$A_{xg} = U_g^* A_x U_g. \tag{4.9}$$

Suppose that for all $x \neq y \in X$ and $\psi \in \mathscr{D}$

$$A_x A_y |\psi\rangle\langle\psi| A_y^* A_x^* = A_y A_x |\psi\rangle\langle\psi| A_x^* A_y^* \tag{4.10}$$

and that if $x \neq y$ there exists $\psi \in \mathscr{D}$ with $A_x A_y \psi \neq 0$. Then either

$$A_x A_y - A_y A_x = 0 \tag{4.11}$$

for all $x, y \in X$ or

$$A_x A_y + A_y A_x = 0 \tag{4.12}$$

for all $x, y \in X$.

Proof For all $x \neq y \in X$ and $\psi \in \mathscr{D}$ there exists a constant λ of absolute value 1 such that

$$A_x A_y \psi = \lambda A_y A_x \psi. \tag{4.13}$$

Simple calculations show that this constant must be independent of ψ so

$$A_x A_y = \lambda(x, y) A_y A_x \tag{4.14}$$

where

$$\lambda : M \equiv \{(x, y) : X \times X : x \neq y\} \to \{z \in \mathbb{C} : |z| = 1\}. \tag{4.15}$$

For each $x_0 \neq y_0$ the value $\lambda(x_0, y_0)$ is unique by the non-degeneracy condition. If $\xi, \eta \in \mathscr{D}$ are chosen so that $\langle A_{x_0} A_{y_0} \xi, \eta\rangle \neq 0$ then

$$\langle U_{g_1}^* A_{x_0} U_{g_1} U_{g_2}^* A_{y_0} U_{g_2} \xi, \eta\rangle = \lambda(x_0 g_1, y_0 g_2)\langle U_{g_2}^* A_{y_0} U_{g_2} U_{g_1}^* A_{x_0} U_{g_1} \xi, \eta\rangle \tag{4.16}$$

so λ is continuous in a neighbourhood of (x_0, y_0), and hence on the whole of M. The transformation properties of A_x imply both that if $x \neq y$ then $\lambda(x, y) = \lambda(y, x)^{-1}$ and that $\lambda(xg, yg) = \lambda(x, y)$ for all $g \in G$. If g is chosen to interchange x and y we deduce that $\lambda(x, y) = \pm 1$.

If $\dim X \geqslant 2$ then M is obtained by removing a manifold of $\dim X$ from a manifold of dimensions $(\dim X)^2$ and so is connected. Since λ is continuous with values ± 1 everywhere, it is constant. If $\dim X = 1$ then since it is connected either $X \simeq \{z \in \mathbb{C} : |z| = 1\}$ in which case M is again connected, or $X \simeq \mathbb{R}$. In the second case M has two components but $(x, y) \to (y, x)$ interchanges them, so λ is again constant. QED

5.5 Generators of dynamical semigroups

Among the other results section 3 contained the construction of a class of dynamical semigroups, which seems worth isolating.

If H is an (unbounded) self-adjoint operator on $\mathscr{H}$, the formula

$$U_t(\rho) = e^{-itH} \rho\, e^{itH} \tag{5.1}$$

defines a strongly continuous one-parameter group of isometries on the state space V, whose infinitesimal generator we denote by $-iad_H$.

Lemma 5.1 *The domain of ad_H is the set of all self-adjoint trace class operators ρ on $\mathscr{H}$ such that*

$$\rho\{\mathrm{dom}(H)\} \subseteq \mathrm{dom}(H) \tag{5.2}$$

and such that the operator $H\rho - \rho H$ on dom(H) *is norm bounded with an extension to a trace class operator σ on $\mathscr{H}$.*

Proof We denote by $\mathscr{D}$ the class of operators satisfying the conditions of the lemma and by dom(ad_H) the domain of ad_H.

If $\rho \in \mathrm{dom}(ad_H)$ and $\psi \in \mathrm{dom}(H)$ then

$$\begin{aligned}\frac{d}{dt} e^{-itH}(\rho\psi) &= \frac{d}{dt}(U_t\rho)(e^{-itH}\psi)\\ &= \{-iad_H(U_t\rho)\}(e^{-itH}\psi) + (U_t\rho)\{-iH\, e^{-itH}\psi\}\end{aligned} \tag{5.3}$$

so putting $t = 0$

$$-iH(\rho\psi) = (-iad_H\rho)\psi - i\rho(H\psi) \tag{5.4}$$

Therefore $\rho\psi \in \mathrm{dom}(H)$ and

$$(ad_H\rho)\psi = (H\rho - \rho H)\psi \tag{5.5}$$

for all $\psi \in \mathrm{dom}(H)$. Therefore $\rho \in \mathscr{D}$.

Conversely let $\rho \in \mathscr{D}$ and let $\rho_n = P_n \rho P_n$, where P_n is the spectral projection of H associated with the interval $[-n, n]$. Since P_n converge strongly to 1 as $n \to \infty$ and ρ is trace class, ρ_n converge in trace norm to ρ. Moreover H is bounded on $P_n\mathscr{H}$, and by restricting to this subspace one can see that $\rho_n \in \mathrm{dom}(ad_H)$ with

$$\begin{aligned} ad_H(\rho_n) &= -i[H, \rho_n] \\ &= -i(HP_n\rho P_n - P_n\rho P_n H) \\ &= -iP_n(H\rho - \rho H)P_n \\ &= -iP_n\sigma P_n \end{aligned} \tag{5.6}$$

where σ is the trace class operator in the statement of the lemma. Since P_n converges strongly to 1, $P_n\sigma P_n$ converges in trace norm to σ. Because the generator ad_H is a closed operator $\rho \in \mathrm{dom}(ad_H)$ and $ad_H(\rho) = -i\sigma$. QED

Theorem 5.2 *Let H be an (unbounded) self-adjoint operator on $\mathscr{H}$ and $\mathscr{J}$ a (bounded) positive linear map on the state space V. Then the operator Z on V defined by*

$$Z(\rho) = -iad_H(\rho) + \mathscr{J}(\rho) - \tfrac{1}{2}\{\mathscr{J}^*(1)\rho + \rho\mathscr{J}^*(1)\} \tag{5.7}$$

with domain equal to that of ad_H, is the generator of a dynamical semigroup on V.

Proof Since Z is a bounded perturbation of ad_H it is the generator of a one-parameter semigroup T_t on V. If $\rho \in \mathrm{dom}(Z)$ then $T_t(\rho) \in \mathrm{dom}(Z)$ for all $t \geqslant 0$ and

$$\begin{aligned} \frac{d}{dt}\mathrm{tr}[T_t\rho] &= \mathrm{tr}[Z(T_t\rho)] \\ &= 0 \end{aligned} \tag{5.8}$$

so

$$\mathrm{tr}[T_t\rho] = \mathrm{tr}[\rho]. \tag{5.9}$$

Since dom Z is dense in V and T_t is a bounded operator on V, Eq. (5.9) holds for all $\rho \in V$.

If W is defined by

$$W(\rho) = -iad_H(\rho) - \tfrac{1}{2}\{\mathscr{J}^*(1)\rho + \rho\mathscr{J}^*(1)\} \tag{5.10}$$

then by a calculation similar to that of Lemma 5.1, W is the generator of the semigroup S_t on V defined by

$$S_t(\rho) = B_t\rho B_t^* \tag{5.11}$$

where B_t is the one parameter contraction semigroup on $\mathscr{H}$ with generator $\{-iH - \frac{1}{2}\mathscr{J}^*(1)\}$. Now $Z = W + \mathscr{J}$ where $\mathscr{J}$ is bounded, so by (A8)

$$T_t\rho = \lim_{n\to\infty}\{S_{t/n}\, e^{\mathscr{J}t/n}\}^n\rho \tag{5.12}$$

for all $\rho \in V$, the limit being taken in the trace norm. But if $t \geqslant 0$ both S_t and $e^{\mathscr{J}t}$ are positive linear maps on V, so by Eq. (5.12), T_t is a positive linear map. QED

Notes

1 It is not appropriate to give here a detailed account of the experimental and theoretical development of this field, so we content ourselves with referring the reader to Glauber (1963, 1968) and Klauder and Sudarshan (1968).

2–4 These are taken from Davies (1969) where it is also shown that the condition (2.23) is a consequence of the apparently weaker condition that S_t is a pure operation for all $t \geqslant 0$. Local commutativity is discussed in an operational setting by Hellwig and Kraus (1970) and Steinmann (1968) but Theorem 4.1 is due to Davies (1969). QSPs for boson fields are constructed in Davies (1969) and (1971).

5 Much more general results of the type of Lemma 5.1 may be found in Bratteli and Robinson (1975) and Evans (1975).

6

The Generalised Wigner–Weisskopf Atom

6.1 Beam-foil spectroscopy

It is possible to describe many of the experimental results of atomic spectroscopy by means of a very simple model. One supposes that an electron has energy equal to one of the discrete resonant energies of some atomic Hamiltonian, and to be capable of transitions between these energies at a rate determined by calculations from quantum electrodynamics. The appropriate mathematical model is a Markov process whose state space is the set of eigenvalues of the atomic Hamiltonian.

A more sophisticated model is necessary if one wants to accommodate the possibility that the initial state of the electron is a superposition of two (or several) eigenstates rather than a mixture. The need for this more refined model is particularly clear in the field of beam-foil spectroscopy, where it is possible to observe the decay of an excited state of an atom over very short times, comparable to its half-life.

In beam-foil spectroscopy a beam of ions is accelerated to an energy of order 1 MeV and then made to pass through a very thin foil. The ions emerge on the other side as neutral but excited atoms. A photon detector is placed to detect radiation from the atoms perpendicular to their direction of motion at a specified distance downstream from the foil. By moving the detector and using the known velocity of the atoms, one can plot the rate of decay as a function of the time since impact over times of the order of 10^{-9} s. Instead of a simple exponential decay curve, one obtains typically a decay curve heavily modulated by oscillations with frequencies equal to the fine structure splittings of the atomic Hamiltonian.

The explanation of this phenomenon is that a typical atom traverses the foil so rapidly (10^{-13} s) that it can emerge in a superposition of two fine structure levels. The decay rate of the atom then changes periodically as the phase difference between the two components oscillates. There exist successful, even if partly phenomenological, models of this behaviour which involve generalising the Wigner–Weisskopf atom to a system with several energy levels.

It turns out that these phenomenological models are closely related to QSPs, although this is not obvious from the physical literature because only quantities related to the first decay are usually considered, and for these the full strength of the theory is not needed.

We devote this chapter to the study of the relevant class of finite-dimensional QSPs. The theory is analogous in the sense of chapter 2 section 5 to the theory of continuous time Markov chains in probability theory. In particular our definitions of irreducibility, transience and recurrence correspond exactly to the concepts of the same names in probability theory.

6.2 Irreducible and simple processes

For the remainder of this chapter we suppose that V is the state space of a separable Hilbert space $\mathscr{H}$ and that the value space X is locally compact and second countable. We suppose that $A : X \to \mathscr{L}(\mathscr{H})$ is a strongly continuous operator-valued function such that for all $\psi \in \mathscr{H}$

$$\int_X \|A_x\psi\|^2 \, dx \leqslant K\|\psi\|^2 \tag{2.1}$$

for some constant K and a Borel measure dx on X such that the measure of every non-empty open set in X is non-zero. The formula

$$\mathscr{J}(E, \rho) = \int_E A_x \rho A_x^* \, dx \tag{2.2}$$

then defines a PMV measure on X and the rate $R \in \mathscr{L}_s(\mathscr{H})$ satisfies

$$0 \leqslant R \equiv \int_X A_x^* A_x \, dx \leqslant K1. \tag{2.3}$$

If H is an (unbounded) self-adjoint operator on $\mathscr{H}$ and $Y = -iH - \frac{1}{2}R$ then Y is the generator of a one-parameter contraction semigroup B_t on $\mathscr{H}$, and Y, $\mathscr{J}$ are the generators of a QSP $\mathscr{E}$ on V by Theorem 5.3.6. We define the semigroups T_t and S_t as in chapter 5 section 2. For the remainder of this chapter we shall call a QSP constructed in the above manner a process.

If $\mathscr{E}$ is a process, a norm closed order ideal $I \subseteq V^+$ is called invariant with respect to $\mathscr{E}$ if $\rho \in I$ implies $T_t(\rho) \in I$ for all $t \geqslant 0$. If E is any Borel set in the sample space X_t then

$$0 \leqslant \mathscr{E}^t(E, \rho) \leqslant T_t(\rho) \in I \tag{2.4}$$

for any $\rho \in I$ so $\mathscr{E}^t(E, \rho) \in I$. If K is the closed subspace of $\mathscr{H}$ associated with the ideal I by Lemma 4.3.2 then $\mathscr{E}$ can be restricted to a process on the state space of K. We call the process $\mathscr{E}$ irreducible if it has no proper such restriction.

Lemma 2.1 *The process $\mathscr{E}$ is irreducible if and only if there is no proper closed subspace K of $\mathscr{H}$ such that $A_x K \subseteq K$ for all $x \in X$ and*

$$(1 - Y)\{K \cap \mathrm{dom}(Y)\} = K. \tag{2.5}$$

Proof Suppose that the closed order ideal I corresponding to the subspace $K \subseteq \mathscr{H}$ is invariant. Suppose that $\psi \in K$ and let f_n be a sequence of non-negative continuous functions on X such that $\int_X f_n(x)dx = 1$ and such that the supports of f_n decrease to the point $x \in X$. If the PMV measure $\mathscr{J}^t$ is defined by Eq. (5.3.1) then

$$0 \leqslant \mathscr{J}^t(f_n, |\psi\rangle\langle\psi|) \leqslant t^{-1}T_t(|\psi\rangle\langle\psi|) \tag{2.6}$$

and

$$\mathscr{J}(f_n, |\psi\rangle\langle\psi|) = \lim_{t\to 0} \mathscr{J}^t(f_n, |\psi\rangle\langle\psi|) \tag{2.7}$$

so

$$\mathscr{J}(f_n, |\psi\rangle\langle\psi|) \in I. \tag{2.8}$$

Also

$$A_x|\psi\rangle\langle\psi|A_x^* = \lim_{n\to\infty} \mathscr{J}(f_n, |\psi\rangle\langle\psi|) \tag{2.9}$$

so $A_x\psi \in K$. Similarly if $\psi \in K$ then

$$0 \leqslant B_t|\psi\rangle\langle\psi|B_t^* \leqslant T_t(|\psi\rangle\langle\psi|) \tag{2.10}$$

so $B_t\psi \in K$. The infinitesimal generator of B_t restricted to K has domain $K \cap \mathrm{dom}(Y)$ and Eq. (2.5) is a consequence of (A7).

Conversely suppose the conditions of the lemma are satisfied. Eq. (2.5) implies that $(1 - Y)^{-1}$ maps K into K. By Eq. (1.8.1), $(z1 - Y)^{-1}$ maps K into K for all $\mathrm{Re}\, z > 0$, so B_t leaves K invariant by Eq. (1.9.4). The invariance of the order ideal I associated with K may now be seen by examining the proof of Theorem 5.3.6. QED

We have already defined the sample spaces X_t if $0 \leqslant t < \infty$. We now define X_∞ as the set of all sequences $\{(x_i, t_i)\}_{i=1}^n$ of events such that $0 < t_i < t_{i+1}$ for all i, and either n is finite or $n = \infty$ and $\lim_{i\to\infty} t_i = \infty$. There are natural mappings $\pi : X_t \to X_s$ for all $0 \leqslant s \leqslant t \leqslant \infty$ obtained by dropping all events which occur after the time s. If $\rho \in V^+$ the measures P_ρ^t on X_t defined for $0 \leqslant t < \infty$ by

$$P_\rho^t(E) = \mathrm{tr}[\mathscr{E}^t(E, \rho)] \tag{2.11}$$

are compatible under these mappings, so there is a measure P_ρ on X_∞ compatible with all of them by Theorem 1.11.1. Instead of regarding $\mathscr{E}^t$ as defined on X_t we can regard it as defined on the σ-field $\mathscr{F}_t$ of Borel sets in X_∞ which are inverse images of Borel sets in X_t. This is only a change of view-point but we shall find it notationally helpful. If $\omega \in X_\infty$ is a sample point we denote by $x_n(\omega)$ and $t_n(\omega)$ the place and time of the nth event. If a sample point has fewer than n events we put $t_n(\omega) = \infty$. If U is a Borel set in X the following Borel sets in X_∞ will be needed.

$$A_t^n = \{t_n(\omega) \leqslant t,\ t_{n+1}(\omega) > t\}, \tag{2.12}$$

$$B_t^{n,U} = \{x_i(\omega) \in U \text{ on at least } n \text{ occasions with } t_i(\omega) \leqslant t\}, \tag{2.13}$$

except that if $t = \infty$ we demand that all relevant $t_i(\omega)$ are finite.

$$L_t^U = \bigcup_{i=1}^{\infty} \{t_i(\omega) \leqslant t,\ x_i(\omega) \in U,\ t_{i+1}(\omega) > t\}. \tag{2.14}$$

If t is an integral multiple of n^{-1} we define

$$C_t^n = \left\{\text{if } \frac{r-1}{n} < t_i(\omega) \leqslant \frac{r}{n} \leqslant t \text{ for any integers } r \text{ and } i \text{ then } t_{i+1}(\omega) > \frac{r}{n}\right\}. \tag{2.15}$$

Equation (5.2.17) states that for some constant K

$$\operatorname{tr}[\mathscr{E}^t(B_t^{n,X}, \rho)] \leqslant \frac{K^n t^n \operatorname{tr}[\rho]}{n!} \tag{2.16}$$

for all $\rho \in V^+$. We now construct approximate discrete skeletons to the continuous time process $\mathscr{E}$.

Lemma 2.2 *If $t > 0$ is an integer, $\rho \in V$, $E \in \mathscr{F}_t$ and $n \geqslant K$ then*

$$\|\mathscr{E}^t(E \cap C_t^n, \rho) - \mathscr{E}^t(E, \rho)\|_{\mathrm{tr}} \leqslant \frac{K^2 t \|\rho\|_{\mathrm{tr}}}{2n}. \tag{2.17}$$

Proof If $\rho \in V^+$ and $n \geqslant K$ then

$$\begin{aligned} \operatorname{tr}[\mathscr{E}^t(C_t^n, \rho)] &\geqslant \left(1 - \frac{K^2}{2n^2}\right)^{nt} \operatorname{tr}[\rho] \\ &\geqslant \left(1 - \frac{K^2 t}{2n}\right) \operatorname{tr}[\rho] \end{aligned} \tag{2.18}$$

by Eq. (2.16). Therefore

$$\begin{aligned} \|\mathscr{E}^t(E \cap C_t^n, \rho) - \mathscr{E}^t(E, \rho)\|_{\mathrm{tr}} &= \operatorname{tr}[\mathscr{E}^t(E \setminus C_t^n, \rho)] \\ &\leqslant \operatorname{tr}[\mathscr{E}^t(X_\infty \setminus C_t^n, \rho)] \leqslant \frac{K^2 t \operatorname{tr}[\rho]}{2n}. \end{aligned} \tag{2.19}$$

If $\rho \in V$ then by Eq. (1.3.6) and (1.3.7)

$$\|\mathscr{E}^t(E \cap C_t^n, \rho) - \mathscr{E}^t(E, \rho)\|_{\mathrm{tr}} \leqslant \frac{K^2 t\{\mathrm{tr}[\rho^+] + \mathrm{tr}[\rho^-]\}}{2n} = \frac{K^2 t\|\rho\|_{\mathrm{tr}}}{2n}. \qquad \text{QED} \quad (2.20)$$

We call an open set $U \subseteq X$ null if $P_\rho(B_\infty^{1,U}) = 0$ for all $\rho \in V^+$.

Lemma 2.3 *An open set $U \subseteq X$ is null if and only if $A_x = 0$ for all $x \in U$. If $\mathscr{E}$ is irreducible then either U is null or $P_\rho(B_\infty^{1,U}) > 0$ for all non-zero $\rho \in V^+$; if X is null then $\mathscr{H}$ is one-dimensional.*

Proof If $A_x \neq 0$ for some $x \in U$ then there exists ψ such that $A_x\psi \neq 0$ for that x and hence for all x in a subset of positive measure in U. If $\rho = |\psi\rangle\langle\psi|$ then

$$\mathscr{J}(U, \rho) = \int_{x \in U} |A_x\psi\rangle\langle A_x\psi|\, dx \neq 0. \tag{2.21}$$

By Eqs. (5.3.1) and (5.3.27)

$$\begin{aligned}\mathscr{J}(U, \rho) &= \lim_{t\to 0} \mathscr{J}^t(U, \rho)\\ &= \lim_{t\to 0} t^{-1}\mathscr{E}^t(\{0 < t_1(\omega) \leqslant t, x_1(\omega) \in U, t_2(\omega) > t\}, \rho)\end{aligned} \tag{2.22}$$

so the right-hand side is non-zero for small enough t. Therefore

$$\begin{aligned}P_\rho(B_\infty^{1,U}) &\geqslant P_\rho\{0 < t_1(\omega) \leqslant t, x_1(\omega) \in U, t_2(\omega) > t\}\\ &= \mathrm{tr}[\mathscr{E}^t(\{0 < t_1(\omega) \leqslant t, x_1(\omega) \in U, t_2(\omega) > t\}, \rho)] > 0.\end{aligned} \tag{2.23}$$

Conversely suppose that $\rho \in V^+$ and that

$$\begin{aligned}0 < P_\rho(B_\infty^{1,U}) &= \sum_{n=1}^\infty \sum_{m=1}^\infty P_\rho\{x_i(\omega) \notin U \text{ for } i < n \text{ and } x_n(\omega) \in U\\ &\qquad\qquad \text{and } m-1 < t_n(\omega) \leqslant m\}\\ &= \sum_{n=1}^\infty \sum_{m=1}^\infty \lim_{r\to\infty} P_\rho\{\omega \in C_m^r, x_i(\omega) \notin U \text{ for } i < n,\\ &\qquad\qquad x_n(\omega) \in U, m-1 < t_n(\omega) \leqslant m\}\end{aligned} \tag{2.24}$$

by Lemma 2.2. Therefore for some m, n, r and $\rho' \in V^+$ we obtain

$$\begin{aligned}0 &< \mathrm{tr}[\mathscr{E}^{1/r}(\{x_1(\omega) \in U, t_1(\omega) \leqslant r^{-1}, t_2(\omega) > r^{-1}\}, \rho')]\\ &= \mathrm{tr}[\mathscr{J}^{1/r}(U, \rho')]\\ &= r\int_0^{1/r} \mathrm{tr}[S_{1/r-s}\mathscr{J}(U, S_s\rho')]ds\end{aligned} \tag{2.25}$$

by Eq. (5.3.26). Therefore for some $0 < s < r^{-1}$

$$0 < \mathscr{J}(U, \rho'') = \int_U A_x \rho'' A_x^* \, dx \tag{2.26}$$

where $\rho'' = S_s \rho'$. It follows that $A_x \neq 0$ for some $x \in U$.

Suppose that $\mathscr{E}$ is irreducible and the open set U is not null. The set J of $\rho \in V^+$ such that $P_\rho(B_\infty^{1, U}) = 0$ is a norm closed order ideal. It is invariant because if $\rho \in J$ then

$$P_{T_t \rho}(B_\infty^{1, U}) = P_\rho\{x_i(\omega) \in U \text{ for some } i \text{ such that } t < t_i(\omega)\} \leqslant P_\rho(B_\infty^{1, U}) = 0. \tag{2.27}$$

Now $J \neq V^+$ since U is not null, so $J = 0$. If X is a null set then $A_x = 0$ for all $x \in X$ so $\mathscr{J} = 0$ and $R = 0$. Therefore $Y = -iH$ and $T_t(\rho) = e^{-itH} \rho \, e^{itH}$. If $\dim \mathscr{H} > 1$ there exists a proper projection P commuting with H and then $J = \{\rho \in V^+ : P\rho = \rho P = \rho\}$ is a proper closed invariant order ideal, so $\mathscr{E}$ cannot be irreducible. QED

No longer assuming irreducibility we say that a process $\mathscr{E}$ on X, V is simple if for every open set $U \subseteq X$ with compact closure the order ideal V_U^+ generated by the set of $\mathscr{E}^t(L_t^U, \rho)$ where $\rho \in V^+$ and $t > 0$ is finite dimensional. A slight modification of the proofs of Lemmas 2.1 and 2.3 leads to

Theorem 2.4 *A process $\mathscr{E}$ is simple if and only if for every open set U in X with compact closure there exists a finite-dimensional subspace K of $\mathscr{H}$ such that $A_x K \subseteq K$ for all $x \in U$, $K \subseteq \mathrm{dom}(Y)$ and $YK \subseteq K$.*

6.3 Recurrence and transience

Throughout this section we suppose that $\mathscr{E}$ is a simple irreducible process on X, V and that U, W are two open subsets of X with compact closures. If $E \subseteq X$ is a Borel set we say the process is m times recurrent at E for the state $\rho \in V^+$ if

$$P_\rho\{B_\infty^{m, E}\} = \mathrm{tr}[\rho]. \tag{3.1}$$

Lemma 3.1 *If $\mathscr{E}$ is singly recurrent at U for all $\rho \in V_U^+$ then it is infinitely recurrent at U for all $\rho \in V^+$, unless U is null.*

Proof Suppose that for some integer m and all $\rho \in V_U^+$ we have established that $P_\rho\{B_\infty^{m, U}\} = \mathrm{tr}[\rho]$. If $\varepsilon > 0$ then by the compactness of $\{\rho \in V_U^+ : \mathrm{tr}[\rho] = 1\}$ and Lemma 2.2 there exist integers t and n such that for all $\rho \in V^+$

$$P_\rho\{B_1^{m, U} \cap C_{2t}^n\} \geqslant \left(1 - \frac{\varepsilon}{2}\right) \mathrm{tr}[\rho]. \tag{3.2}$$

Therefore

$$P_\rho\{B_\infty^{2m,U}\} \geqslant P_\rho\{B_{2t}^{2m,U} \cap C_{2t}^n\}$$

$$\geqslant \sum_{p,q} P_\rho\Big\{x_i(\omega) \in U \text{ for the } m\text{th time where } \frac{q-1}{n} < t_i(\omega) \leqslant \frac{q}{n} \leqslant t,$$

$$x_j(\omega) \in U \text{ for the } 2m\text{th time where } \frac{p+q+1}{n} < t_j(\omega)$$

$$\leqslant \frac{p+q}{n} \leqslant 2t,\ \omega \in C_{2t}^n\Big\}. \tag{3.3}$$

If we define

$$\rho_q = \mathscr{E}^{q/n}\Big(\Big\{\omega \in C_{q/n}^n,\ x_i(\omega) \in U \text{ for the } n\text{th time where } \frac{q-1}{n} < t_i(\omega) \leqslant \frac{q}{n} \leqslant t\Big\}, \rho\Big) \tag{3.4}$$

then $0 \leqslant \rho_q \leqslant \mathscr{E}^{q/n}(L_{q/n}^U, \rho) \in V_U^+$ so $\rho_q \in V_U^+$. The right-hand side of Eq. (3.3) now becomes

$$= \sum_{p,q} P_{\rho_q}\Big\{\omega \in C_{2t-q/n}^n,\ x_j(\omega) \in U \text{ for the } n\text{th time where } \frac{p-1}{n} < t_j(\omega) \leqslant \frac{p}{n} \leqslant 2t - \frac{q}{n}\Big\}$$

$$= \sum_q P_{\rho_q}\{C_{2t-q/n}^n \cap B_{2t-q/n}^{m,U}\}$$

$$\geqslant \sum_q P_{\rho_q}\{C_{2t}^n \cap B_t^{m,U}\}$$

$$\geqslant \sum_q \left(1 - \frac{\varepsilon}{2}\right)\mathrm{tr}[\rho_q]$$

$$= \left(1 - \frac{\varepsilon}{2}\right)\sum_q P_\rho\Big\{\omega \in C_{q/n}^n,\ x_i(\omega) \in U \text{ for the } n\text{th time where } \frac{q-1}{n} < t_i(\omega) \leqslant \frac{q}{n} \leqslant t\Big\}$$

$$\geqslant \left(1 - \frac{\varepsilon}{2}\right) P_\rho\{C^n_{2t} \cap B^{m,U}_t\}$$

$$\geqslant \left(1 - \frac{\varepsilon}{2}\right)^2 \operatorname{tr}[\rho] \geqslant (1 - \varepsilon)\operatorname{tr}[\rho]. \tag{3.5}$$

As $\varepsilon > 0$ is arbitrary we obtain $P_\rho\{B^{2m,U}_\infty\} = \operatorname{tr}[\rho]$. Now we know that $P_\rho\{B^{1,U}_\infty\} = \operatorname{tr}[\rho]$ for all $\rho \in V^+_U$ so inductively we obtain $P_\rho\{B^{\infty,U}_\infty\} = \operatorname{tr}[\rho]$ for all $\rho \in V^+_U$. The set

$$J = \{\rho \in V^+ : P_\rho\{B^{\infty,U}_\infty\} = \operatorname{tr}[\rho]\} \tag{3.6}$$

is a norm closed invariant order ideal containing V^+_U, which is non-zero if U is not null. By irreducibility $J = V^+$. QED

Lemma 3.2 *If the process $\mathscr{E}$ is singly recurrent at U for all $\rho \in V^+_U$ then it is singly recurrent at W for all $\rho \in V^+_U$, unless W is null.*

Proof By Lemma 2.3 and the compactness of $\{\rho \in V^+_U : \operatorname{tr}[\rho] = 1\}$ there is a constant $k > 0$ such that $P_\rho\{B^{1,W}_\infty\} \geqslant k \operatorname{tr}[\rho]$ for all $\rho \in V^+_U$. Let k be the largest such constant and let $\rho_0 \in V^+_U$ be a state such that $P_{\rho_0}\{B^{1,W}_\infty\} = k \operatorname{tr}[\rho_0] = k$. Let $\varepsilon > 0$ and let t be a sufficiently large integer so that $P_{\rho_0}\{B^{1,W}_t\} > k - \varepsilon$. Using Lemma 3.1 let s be a sufficiently large integer that

$$P_{\rho_0}\{\text{if } t_i(\omega) \leqslant t \text{ then } x_i(\omega) \notin W, \text{ but } x_j(\omega) \in U \text{ for some } t < t_j(\omega) \leqslant t + s\} > 1 - k - \varepsilon. \tag{3.7}$$

Using Lemma 2.2 let n be a sufficiently large integer that

$$P_{\rho_0}\{\omega \in C^n_{s+t};\ \text{if } t_i(\omega) \leqslant t \text{ then } x_i(\omega) \notin W, \text{ but } x_j(\omega) \in U \text{ for some } t < t_j(\omega) \leqslant t + s\} > 1 - k - \varepsilon. \tag{3.8}$$

Then

$$P_{\rho_0}\{B^{1,W}_\infty\} \geqslant P_{\rho_0}\{B^{1,W}_t\} + \sum_{m=1}^{ns} P_{\rho_m}\{B^{1,W}_\infty\} \tag{3.9}$$

where

$$\rho_m = \mathscr{E}^{t+m/n}\left(\left\{\omega \in C^n_{t+m/n}, x_i(\omega) \notin W \text{ for } t_i(\omega) \leqslant t, x_i(\omega) \notin U \right.\right.$$
$$\left.\left.\text{for } t < t_i(\omega) \leqslant \frac{m-1}{n}, x_i(\omega) \in U \text{ for some } \frac{m-1}{n} < t_i(\omega) \leqslant \frac{m}{n}\right\}, \rho_0\right)$$
$$\leqslant \mathscr{E}^{t+m/n}(L^U_{t+m/n}, \rho_0) \in V^+_U. \tag{3.10}$$

Since $\rho_m \in V_U^+$, $P_{\rho_m}\{B_\infty^{1,W}\} \geqslant k\,\mathrm{tr}[\rho_m]$ and Eq. (3.9) becomes

$$\begin{aligned}
k &\geqslant k - \varepsilon + k \sum_{m=1}^{ns} \mathrm{tr}[\rho_m] \\
&= k - \varepsilon + k \sum_{m=1}^{ns} P_{\rho_0}\Big\{\omega \in C^n_{t+m/n},\ x_i(\omega) \notin W \text{ for } t_i(\omega) \leqslant t, \\
&\qquad x_i(\omega) \notin U \text{ for } t < t_i(\omega) \leqslant \frac{m-1}{n},\ x_i(\omega) \in U \text{ for some } \frac{m-1}{n} < t_i(\omega) \leqslant \frac{m}{n}\Big\} \\
&\geqslant k - \varepsilon + kP_{\rho_0}\{\omega \in C^n_{t+s},\ x_i(\omega) \notin W \text{ for } t_i(\omega) \leqslant t \text{ but} \\
&\qquad\qquad x_i(\omega) \notin U \text{ for some } t < t_i(\omega) \leqslant t + s\} \\
&> k - \varepsilon + k(1 - k - \varepsilon).
\end{aligned} \tag{3.11}$$

Letting $\varepsilon \to 0$ gives $0 \geqslant k(1-k)$, and as $k > 0$ we have $k = 1$, which is the required result. QED

Lemma 3.3 *Suppose U and W are not null. If $\mathscr{E}$ is singly recurrent at U for all $\rho \in V_U^+$ then it is infinitely recurrent at W for all $\rho \in V^+$.*

Proof Let $\rho \in V^+$ and for $\varepsilon > 0$ let t, n be large enough integers so that $P_\rho\{C_t^n \cap B_t^{1,U}\} \geqslant (1-\varepsilon)\mathrm{tr}[\rho]$. Then defining

$$\rho_m = \mathscr{E}^{m/n}\left(\left\{\omega \in C^n_{m/n},\ x_i(\omega) \notin U \text{ for } t_i(\omega) \leqslant \frac{m-1}{n} \text{ but } x_i(\omega) \in U \text{ for some } \frac{m-1}{n} < t_i(\omega) \leqslant \frac{m}{n}\right\}, \rho\right) \tag{3.12}$$

so that

$$0 \leqslant \rho_m \leqslant \mathscr{E}^{m/n}(L^U_{m/n}, \rho) \in V_U^+ \tag{3.13}$$

we obtain

$$\begin{aligned}
P_\rho\{B_\infty^{1,W}\} &\geqslant \sum_{m=1}^{nt} P_{\rho_m}\{B_\infty^{1,W}\} \\
&\geqslant \sum_{m=1}^{nt} \mathrm{tr}[\rho_m] \\
&\geqslant P_\rho\{C_t^n \cap B_t^{1,U}\} \\
&\geqslant (1-\varepsilon)\mathrm{tr}[\rho].
\end{aligned} \tag{3.14}$$

Letting $\varepsilon \to 0$ gives the equation $P_\rho\{B_\infty^{1,W}\} = \mathrm{tr}[\rho]$ for all $\rho \in V^+$. By Lemma 3.1 it follows that $P_\rho\{B_\infty^{\infty,W}\} = \mathrm{tr}[\rho]$ for all $\rho \in V^+$. QED

We turn to the study of processes which do not satisfy the hypothesis of Lemma 3.1. We define an integer-valued function $N_t^E : X_\infty \to [0, \infty]$ for every Borel set $E \subseteq X$ and every $t < \infty$ by

$$N_t^E(\omega) = \{\text{number of } i \text{ for which } x_i(\omega) \in E \text{ and } t_i(\omega) \leqslant t\} \tag{3.15}$$

while for $t = \infty$ we demand $t_i(\omega) < \infty$. For each state $\rho \in V^+$ the expected number of occurrences $N_t^E(\rho)$ of events within E up to time t is given by

$$N_t^E(\rho) = \int_{X_\infty} N_t^E(\omega) P\rho(d\omega) \tag{3.16}$$

so that

$$N_\infty^E(\rho) = \sum_{n=1}^{\infty} P_\rho\{x_n(\omega) \in E\}. \tag{3.17}$$

Lemma 3.4 *Suppose $\mathscr{E}$ is not singly recurrent at U for some $\rho_0 \in V_U^+$. Then there exist $0 < k < 1$ and an integer m such that for all $\rho \in V_U^+$ and all integers n*

$$P_\rho\{B_\infty^{mn, U}\} \leqslant k^n \operatorname{tr}[\rho]. \tag{3.18}$$

Proof Let $S_m \subseteq V_U^+$ be the set of $\rho \in V_U^+$ such that $\operatorname{tr}[\rho] = 1$ and $P_\rho\{B_\infty^{m, U}\} = 1$. Then $S_m \subseteq S_{m-1}$ and if $\rho_1 \in \bigcap_{m=1}^\infty S_m$ we have $\operatorname{tr}[\rho_1] = 1$ and $\rho \in J$ where

$$J = \{\rho \in V^+ : P_\rho\{B_\infty^{\infty, U}\} = \operatorname{tr}[\rho]\}. \tag{3.19}$$

But J is a norm closed invariant order ideal and $\rho_0 \notin J$ so by irreducibility $J = 0$. Therefore $\bigcap_{m=1}^\infty S_m = \varnothing$ and by compactness $S_m = \varnothing$ for some m. For that m there exists a constant $0 < k < 1$ such that for all $\rho \in V_U^+$ Eq. (3.18) holds if $n = 1$.

Suppose that we have established that Eq. (3.18) holds for some n and all $\rho \in V_U^+$. Then for any $\varepsilon > 0$ there exist integers r and t such that

$$\begin{aligned} P_\rho\{B_\infty^{m(n+1), U}\} - \varepsilon &\leqslant P_\rho\{B_t^{m(n+1), U} \cap C_t^r\} \\ &= \sum_{s=1}^{rt} P_{\rho_s}\{C_{t-s/r} \cap B_{t-s/r}^{mn, U}\} \end{aligned} \tag{3.20}$$

where $\rho_s \in V_U^+$ is defined by

$$\rho_s = \mathscr{E}^{s/r}\left(\left\{\omega \in C_{s/r}^r, x_i(\omega) \in U \text{ for the } m\text{th time where } \frac{s-1}{r} < t_i(\omega) \leqslant \frac{s}{r}\right\}, \rho\right). \tag{3.21}$$

The right-hand side of Eq. (3.20) is less than or equal to

$$\sum_{s=1}^{rt} k^n \operatorname{tr}[\rho_s] \leqslant k^n P_\rho\{B_\infty^{m,U}\}$$
$$\leqslant k^{n+1} \operatorname{tr}[\rho]. \tag{3.22}$$

Going to the limit as $\varepsilon \to 0$ gives the required result by induction. QED

Lemma 3.5 *Suppose $\mathscr{E}$ is not singly recurrent at U for some $\rho_0 \in V_U^+$. Then there exists a constant $\alpha_U < \infty$ such that $N_\infty^U(\rho) \leqslant \alpha_U \operatorname{tr}[\rho]$ for all $\rho \in V^+$.*

Proof For all $\rho \in V^+$ we have

$$N_\infty^U(\rho) = \lim_{t\to\infty} \lim_{n\to\infty} \int_{C_t^n \cap B_t^{1,U}} N_\infty^U(\omega) P_\rho(d\omega). \tag{3.23}$$

Also if we define $\rho_s \in V_U^+$ by

$$\rho_s = \mathscr{E}^{s/n}\left(\left\{\omega \in C_{s/n}^n, x_i(\omega) \notin U \text{ for } t_i(\omega) \leqslant \frac{s-1}{n}, x_i(\omega) \in U \right.\right.$$
$$\left.\left.\text{for some } \frac{s-1}{n} < t_i(\omega) \leqslant \frac{s}{n}\right\}, \rho\right) \tag{3.24}$$

and $\sigma = \sum_{s=1}^{nt} \rho_s \in V_U^+$, then $\operatorname{tr}[\sigma] \leqslant \operatorname{tr}[\rho]$ and

$$\int_{C_t^n \cap B_t^{1,U}} N_\infty^U(\omega) P_\rho(d\omega) = \sum_{s=1}^{nt} \left\{\operatorname{tr}[\rho_s] + \int_{C_{t-s/n}^n} N_\infty^U(\omega) P_{\rho_s}(d\omega)\right\}$$
$$\leqslant \operatorname{tr}[\rho] + \int_{X_\infty} N_\infty^U(\omega) P_\sigma(d\omega). \tag{3.25}$$

By Lemma 3.4 there exist constants k and m such that

$$P_\sigma\{N_\infty^U(\omega) \geqslant rm\} \leqslant k^r \operatorname{tr}[\sigma] \tag{3.26}$$

for all integers r. Therefore

$$\int_{X_\infty} N_\infty^U(\omega) P_\sigma(d\omega) \leqslant \sum_{r=0}^{\infty} mk^r \operatorname{tr}[\sigma] = \frac{m}{1-k} \operatorname{tr}[\sigma] \tag{3.27}$$

and

$$\int_{X_\infty} N_\infty^U(\omega) P_\rho(d\omega) \leqslant \left(1 + \frac{m}{1-k}\right) \operatorname{tr}[\rho]. \qquad \text{QED} \tag{3.28}$$

We summarise all the results obtained so far.

Theorem 3.6 *Call a process $\mathscr{E}$ on X, V recurrent if every state $\rho \in V^+$ is infinitely recurrent at every open set $U \subseteq X$ which is not null. Call it transient if every open subset $U \subseteq X$ with compact closure has a finite expected number of occurrences for every state $\rho \in V^+$. Then every simple irreducible process is either recurrent or transient.*

The problem of determining whether a given process is recurrent or transient is in general a difficult one, depending on detailed properties of H and $\mathscr{J}$, but some general guide is contained in the following results.

We call a process on X, V substantially finite-dimensional if there exists a finite rank projection P on $\mathscr{H}$ and a constant $\alpha > 0$ such that

(*i*) $T_t(P)P = PT_t(P)$ for all $t \geqslant 0$; (3.29)

(*ii*) $\mathrm{tr}[PT_t(P)] \geqslant \alpha$ for all $t \geqslant 0$. (3.30)

By taking $P = 1$ it is immediate that every finite-dimensional process is also substantially finite-dimensional.

Theorem 3.7 *Every substantially finite-dimensional, simple, irreducible process is recurrent.*

Proof If U_n in an increasing sequence of open sets with compact closures whose union is X then $N_\infty^{U_n}(\omega)$ converges monotonically to $N_\infty^X(\omega)$ for all $\omega \in X_\infty$. Therefore by Lemma 2.3

$$N_\infty^{U_n}(\rho) \to N_\infty^X(\rho) > 0. \tag{3.31}$$

for all non-zero $\rho \in V^-$.

If P and α are as described above then by the compactness of the set of ρ with $\mathrm{tr}[\rho] = 1$ in

$$J = \{\rho \in V^+ : P\rho = \rho P = \rho\} \tag{3.32}$$

there exist n and $\beta > 0$ such that for all $\rho \in J$

$$N_\infty^{U_n}(\rho) \geqslant \beta\, \mathrm{tr}[\rho]. \tag{3.33}$$

If $\mathscr{E}$ is transient application of the dominated convergence theorem to

$$\int_{X^\infty} N_\infty^{U_n}(\omega)P_\rho(d\omega) = \int_{X^\infty} N_t^{U_n}(\omega)P_\rho(d\omega) + \int_{X^\infty} N_\infty^{U_n}(\omega)P_{T_t\rho}(d\omega) \tag{3.34}$$

yields

$$\lim_{t\to\infty} \int_{X^\infty} N_\infty^{U_n}(\omega)P_{T_t\rho}(d\omega) = 0 \tag{3.35}$$

for all $\rho \in V^+$. If $\rho = P$ and $\rho' = PT_t(P) \in J$ then $0 \leqslant \rho' \leqslant T_t\rho$ by Eq. (3.29). Therefore if t is large enough

$$\int_{X^\infty} N_\infty^{U_n}(\omega)P_{\rho'}(d\omega) \leqslant \int_{X^\infty} N_\infty^{U_n}(\omega)P_{T_t\rho}(d\omega) < \beta\, \mathrm{tr}[\rho']. \tag{3.36}$$

This contradicts Eq. (3.33) so $\mathscr{E}$ cannot be transient. QED

An equilibrium state of a process $\mathscr{E}$ on X, V is by definition a state ρ such that $T_t\rho = \rho$ for all $t \geqslant 0$.

Theorem 3.8 *An equilibrium state of an irreducible process $\mathscr{E}$ has support equal to $\mathscr{H}$. A simple irreducible process with an equilibrium state must be recurrent. A finite-dimensional irreducible process possesses an equilibrium state, unique up to a constant multiple.*

Proof The first statement follows from the fact that if ρ_0 is an equilibrium state

$$J = \{\rho \in V^+ : \rho \leqslant \alpha\rho_0 \text{ some } \alpha\} \tag{3.37}$$

is an invariant order ideal, and so must be dense in V^-. For the second statement we observe as in Theorem 3.7 that if $\mathscr{E}$ is transient and U is an open set in X with compact closure then $N_\infty^U(\rho_0) = N_\infty^U(T_t\rho_0) \to 0$ as $t \to \infty$. Therefore $N_\infty^U(\rho_0) = 0$ for all such U and so $N_\infty^X(\rho_0) = 0$. Therefore $N_\infty^X(\rho) = 0$ for all $\rho \in J$ and hence all $\rho \in V^+$. By Lemma 2.3 this contradicts irreducibility.

If $\mathscr{H}$ is finite-dimensional then T_t is a one-parameter semigroup of affine endomorphisms of the compact convex set

$$X = \{\rho \in V^+ : \mathrm{tr}[\rho] = 1\}. \tag{3.38}$$

By (A4) there is at least one T-fixed point in X. If there is more than one fixed point then by (A5) there are at least two distinct extreme fixed points, which we call ρ_1 and ρ_2. Both ρ_1 and ρ_2 have support equal to $\mathscr{H}$ so there exists a constant α such that $\alpha\rho_1 \leqslant \rho_2 \leqslant \alpha^{-1}\rho_1$. But by the definition of extreme points this implies $\rho_1 = \rho_2$. QED

We finally collect several earlier results into one theorem.

Theorem 3.9 *Let $\mathscr{E}$ be a finite-dimensional process such that there is no proper subspace of $\mathscr{H}$ invariant under all A_x, $x \in X$. Then $\mathscr{E}$ is irreducible and there is a unique equilibrium state ρ_0 of $\mathscr{E}$. Moreover*

$$\lim_{t\to\infty} T_t(\rho) = \rho_0 \tag{3.39}$$

for all $\rho \in V^+$ with $\mathrm{tr}[\rho] = 1$.

Proof The irreducibility of $\mathscr{E}$ is a consequence of Lemma 2.1. If $\mathscr{S}$ is the multiplicative semigroup in $\mathscr{L}(\mathscr{H})$ generated by 1 and $\{A_x : x \in X\}$ then for all non-zero $\psi \in \mathscr{H}$ there exist $B_1, \ldots, B_n \in \mathscr{S}$ such that $B_1\psi, \ldots, B_n\psi$ span $\mathscr{H}$.

We next show that the mixed state $T_t(|\psi\rangle\langle\psi|)$ has support equal to $\mathscr{H}$ for all sufficiently small t. Suppose that none of the operators $B_1, \ldots, B_n$ is a product of more than m operators A_x. If $\rho = |\psi\rangle\langle\psi|$ then

$$T_{t_0}(\rho) \geqslant \sum_{r=0}^{n} \mathscr{E}^{t_0}(\{(i-1)r^{-1}t_0 < t_i(\omega) \leqslant ir^{-1}t_0 \text{ for } i = 1, \ldots, r\}, \rho)$$

$$= S_{t_0}(\rho) + \sum_{r=0}^{n} \int_{x_i \in X} \int_{(i-1)r^{-1}t_0 < t_i \leqslant ir^{-1}t_0}$$

$$\times |\psi_{t,x}\rangle\langle\psi_{t,x}| dt_1 \cdots dt_r \, dx_1 \cdots dx_r \tag{3.40}$$

where

$$\psi_{t,x} = B_{t_0 - t_r} A_{x_r} B_{t_r - t_{r-1}} A_{x_{r-1}} \cdots A_{x_1} B_{t_1} \psi \tag{3.41}$$

and $t = (t_1, \ldots, t_r)$, $x = (x_1, \ldots, x_r)$, by Eqs (5.3.35) and (5.3.38). The result now follows by taking t_0 small enough and using the continuity of $t, x \to \psi_{t,x}$.

Let X be the compact convex set $\{\rho \in V^+ : \mathrm{tr}[\rho] = 1\}$ and H the hyperplane $\{\rho \in V : \mathrm{tr}[\rho] = 1\}$. The interior of X in H consists of all ρ whose support is equal to $\mathscr{H}$, so ρ_0 is interior. Let m be the Minkowski functional of X in H with respect to ρ_0 as origin, that is

$$m(\rho) = \inf\{\lambda \in \mathbb{R}^+ : (\rho - \rho_0) \in \lambda(X - \rho_0)\}. \tag{3.42}$$

Then $m(\rho_0) = 0$, $m(\rho) < 1$ if and only if ρ is in the interior of X and $m(T_t\rho) \leqslant m(\rho)$ for all $\rho \in H$ and all $t \geqslant 0$. We have shown that if $\rho \in X$ is a pure state then $m(T_t\rho) < 1$ for sufficiently small $t > 0$, and hence for all $t > 0$. Therefore $m(T_1\rho) < 1$ for all $\rho \in X$ and by compactness there exists a constant $k < 1$ such that $m(T_1\rho) \leqslant k$ for all $\rho \in X$. Therefore $m(T_n\rho) \leqslant k^n$ for all $\rho \in X$ and all integers n, so $T_t\rho \to \rho_0$ as $t \to \infty$. QED

6.4 Transition processes

It is difficult to say very much more about QSPs without making further hypotheses about the infinitesimal generators. In this section we study a class of processes which we call transition processes because of their relevance to the evolution of simple quantum-mechanical systems emitting photons while undergoing transitions from one energy level to another. We define transition processes by certain global properties and then prove certain results about their infinitesimal generators, though it will become evident that we could just as easily have proceeded in the other direction.

A process $\mathscr{E}$ on X, V will be called a transition process if

(*i*) there is given a second countable locally compact Hausdorff space X_0 and a continuous map σ of X into $X_0 \times X_0$; (4.1)

(*ii*) if $\rho \in V^+$ then $P_\rho\{\text{for some } i, \sigma x_i(\omega) = (y_1, y_2), \sigma x_{i+1}(\omega) = (y_3, y_4) \text{ and } y_2 \neq y_3\} = 0.$ (4.2)

The first of these conditions was observed to hold for the hydrogen atom, and this indeed constituted the first major advance in atomic spectroscopy. The second cannot be observed directly since the emissions of an individual atom cannot normally be distinguished from those of neighbouring atoms. For the hydrogen atom we let S be the unit sphere in Euclidean space, let X_0 be the set of energy levels of the atom and take $X = X_0 \times X_0 \times S$. Then each event in X determines the energy levels before and after emission as well as the direction of emission of the photon. The Hamiltonian H is well known and the transition operators are calculated by perturbation methods using quantum electrodynamics.

Although the observed events are transitions, so that one does not have a repeatable observable on X in the sense of chapter 4 section 3, one may still obtain a projection-valued measure on X_0, corresponding to the energy observable. For any Borel set $E \subseteq X_0$ we define W_E^+ to be the norm closed order ideal generated by the set of all $\mathscr{E}^t(L_t^{\sigma^{-1}(X_0 \times E)}, \rho)$ where $\rho \in V^+$ and $t > 0$. Let P_E be the corresponding projection.

Theorem 4.1 *Let $\mathscr{E}$ be a finite-dimensional irreducible transition process on X. Then $P_{(\cdot)}$ is a projection-valued measure which commutes with the Hamiltonian H. If $x \in X$ and $\sigma x = (y, z)$ then A_x has support in $P_{\{y\}}\mathscr{H}$ and range in $P_{\{z\}}\mathscr{H}$.*

Proof If F is a Borel set in X_0 the formula

$$\mathrm{tr}[A\rho] = P_\rho\{x_1(\omega) \in \sigma^{-1}(F \times X_0)\} \tag{4.3}$$

valid for all $\rho \in V^+$, defines an operator A on $\mathscr{H}$ such that $0 \leqslant A \leqslant 1$. The set

$$J = \{\rho \in V^+ : \mathrm{tr}[A\rho] = \mathrm{tr}[\rho]\} \tag{4.4}$$

is a norm closed order ideal which contains all states of the form $\mathscr{E}^t(L_t^{\sigma^{-1}(X_0 \times F)}, \rho)$ because $\mathscr{E}$ is recurrent. Therefore $J \supseteq W_F^+$ and every $\rho \in W_F^+$ has support in the eigenspace of A corresponding to the eigenvalue one. Similarly if $E \cap F = \varnothing$ every $\rho \in W_E^+$ has support in the eigenspace of A corresponding to the eigenvalue zero. Therefore $P_E P_F = P_F P_E = 0$. It is immediate from their definitions that $P_\varnothing = 0$, that $P_E \leqslant P_F$ if $E \subseteq F$, and from recurrence that $P_{X_0} = 1$. Let $\{E_n\}$ be a countable disjoint family of Borel sets in X_0 with union E. We show that the projection $P_0 = P_E - \sum_{n=1}^{\infty} P_{E_n}$ equals zero. For all $\rho \in V^+$ and $t > 0$

$$\mathrm{tr}[P_0 \mathscr{E}^t(L_t^{\sigma^{-1}(X_0 \times E)}, \rho)] = \sum_{n=1}^{\infty} \mathrm{tr}[P_0 \mathscr{E}^t(L_t^{\sigma^{-1}(X_0 \times E_n)}, \rho)] = 0 \tag{4.5}$$

since $P_0 P_{E_n} = 0$ for all n. It follows that $\mathrm{tr}[P_0 \rho] = 0$ for all $\rho \in W_E^+$ so $P_0 P_E = 0$. But $P_0 \leqslant P_E$ so $P_0 = 0$. Therefore $P_{(\cdot)}$ is a projection-valued measure.

We next show that if $\sigma x = (y, z)$ and $A_x \neq 0$ then $P_{\{z\}} \neq 0$. Let U be any open neighbourhood of z and let $U' = \{y : \sigma y \in X_0 \times U\}$. If $A_x \psi \neq 0$ and U'' is any open neighbourhood of x contained in U' then W_U^+ contains

$$\rho = \int_{y \in U''} \int_{s=0}^{t} |\psi_{y,s}\rangle\langle\psi_{y,s}| \, dy \, ds \tag{4.6}$$

where $\psi_{y,s} = B_{t-s} A_y B_s \psi$. By the continuity of $\psi_{y,s}$ it follows that

$$|A_x \psi\rangle\langle A_x \psi| \in W_U^+ \tag{4.7}$$

so that $W_{\{z\}}^+$ and $P_{\{z\}}$ are non-zero.

Equation (4.7) implies that if $\sigma x = (y, z)$ the range of A_x is contained in $P_{\{z\}}\mathscr{H}$. By the transition hypothesis $A_x \rho A_x^* = 0$ for all $\rho \in W_{X_0 - \{y\}}^+$, so the support of A_x is contained in $P_{\{y\}}\mathscr{H}$. Therefore

$$A_x^* A_x = A_x^* A_x P_{\{y\}} = P_{\{y\}} A_x^* A_x \tag{4.8}$$

so the total interaction rate R commutes with the projection valued measure $P_{(\cdot)}$. Similar arguments show that $B_t P_{\{y\}}\mathscr{H} \subseteq P_{\{y\}}\mathscr{H}$ for all $y \in X_0$ and all $t \geqslant 0$. If U_t is the one-parameter unitary group generated by the Hamiltonian H, where $Y = -iH - \frac{1}{2}R$, then U_t may be constructed from B_t, R by Eq. (1.9.7). Therefore $U_t P_{\{y\}}\mathscr{H} \subseteq P_{\{y\}}\mathscr{H}$ for all $y \in X_0$ and $t \geqslant 0$. It follows by the finite-dimensionality of $P_{\{y\}}\mathscr{H}$ that $U_t P_{\{y\}} = P_{\{y\}} U_t$ for all $y \in X_0$ and all $t \in \mathbb{R}$. Therefore the projection-valued measure $P_{(\cdot)}$ commutes with the Hamiltonian H. QED

We finally point out that if X_0 is finite and $P_{\{y\}}$ is of rank one for all $y \in X_0$ then every probability measure μ on X_0 determines a state $\rho(\mu)$ on $\mathscr{H}$ by

$$\rho(\mu) = \sum_{y \in X_0} P_{\{y\}} \mu(y). \tag{4.9}$$

By Theorem 4.1 the set of such states is invariant under T_t so

$$T_t\{\rho(\mu)\} = \rho(\mu_t) \tag{4.10}$$

for some probability measure μ_t on X_0. The map $\mu \to \mu_t$ defines a Markov semigroup on X_0 in the sense of probability theory.

Notes

1 The subject of beam-foil spectroscopy is of recent origin, and we refer the reader to (Alguard and Drake, 1973; Andra, 1970, 1974; Bashkin, 1971; Macek, 1970) for further references. The importance of these experiments as a test of the superposition principle was stressed by Gerjuoy (1973).

2,3,4 This theory is taken from Davies (1970b), although many of the proofs are very similar to the proofs of analogous theorems in the theory of Markov chains, to be found in Chung (1967).

7

Contraction Semigroups on Hilbert Space

7.1 Introduction

A (strongly continuous) one parameter contraction semigroup B_t on a Hilbert space $\mathscr{H}$ is a family B_t of linear contractions defined for $t \geqslant 0$ and satisfying

$$(i) \quad B_0 = 1; \tag{1.1}$$

$$(ii) \quad B_s B_t = B_{s+t} \text{ for all } s, t \geqslant 0; \tag{1.2}$$

$$(iii) \quad \lim_{t \to 0} B_t \phi = \phi \text{ for all } \phi \in \mathscr{H}. \tag{1.3}$$

Lemma 1.1 *Every weakly continuous one parameter contraction semigroup B_t on a Hilbert space $\mathscr{H}$ is also strongly continuous. If B_t is strongly continuous then so is B_t^*.*

Proof Suppose B_t is a weakly continuous, and $\phi \in \mathscr{H}$. Then

$$\begin{aligned}\lim_{t \to 0} \|B_t \phi - \phi\|^2 &= \lim_{t \to 0} \{\|B_t \phi\|^2 + \|\phi\|^2 - 2\,\mathrm{Re}\langle B_t \phi \phi\rangle\} \\ &\leqslant \lim_{t \to 0} \{2\|\phi\|^2 - 2\,\mathrm{Re}\langle B_t \phi, \phi\rangle\} \\ &= 0 \end{aligned} \tag{1.4}$$

so B_t is also strongly continuous. If B_t is strongly continuous then B_t^* is weakly continuous, and hence strongly continuous. QED

We call an (unbounded) operator Y on $\mathscr{H}$ dissipative if

$$\langle \phi, Y\phi\rangle + \langle Y\phi, \phi\rangle \leqslant 0 \tag{1.5}$$

for all $\phi \in \mathrm{dom}(Y)$.

Lemma 1.2 *Let Y be the infinitesimal generator of a strongly continuous one-parameter semigroup B_t on $\mathscr{H}$. Then B_t is a contraction semigroup if and only if Y is dissipative.*

Proof If B_t is a contraction semigroup and $\phi \in \mathrm{dom}(Y)$ then

$$f(t) = \langle B_t\phi, B_t\phi\rangle \tag{1.6}$$

is monotonically decreasing and differentiable. Therefore

$$0 \geqslant f'(0) = \langle Y\phi, \phi\rangle + \langle \phi, Y\phi\rangle \tag{1.7}$$

Conversely if Y is dissipative and $\phi \in \mathrm{dom}(Y)$ then $B_t\phi \in \mathrm{dom}(Y)$ for all $t \geqslant 0$ and

$$\begin{aligned} f'(t) &= \langle YB_t\phi, B_t\phi\rangle + \langle B_t\phi, YB_t\phi\rangle \\ &\leqslant 0 \end{aligned} \tag{1.8}$$

so $f(t)$ is monotonically decreasing and

$$\|B_t\phi\| \leqslant \|\phi\| \tag{1.9}$$

for all $t \geqslant 0$. But dom (Y) is dense and B_t is bounded so Eq. (1.9) holds for all $\phi \in \mathscr{H}$. QED

In chapter 5 section 2 we defined a QSP on the state space V of $\mathscr{H}$. Given $\psi \in \mathscr{H}$ with $\|\psi\| = 1$ and $\rho = |\psi\rangle\langle\psi|$, the state at time $t \geqslant 0$ conditional upon no event having occurred up to time t was taken to be of the form

$$\mathscr{E}^t(z, \rho) = B_t\rho B_t^*. \tag{1.10}$$

The probability that no event has occurred by time t is

$$\mathrm{tr}[\mathscr{E}^t(z, \rho)] = \|B_t\psi\|^2 \tag{1.11}$$

and is a monotonically decreasing function of t. If $\psi \in \mathrm{dom}(Y)$ the probability density for the first event to occur at time t is

$$f(t) = -\frac{d}{dt}\|B_t\psi\|^2 = -2\,\mathrm{Re}\langle YB_t\psi, B_t\psi\rangle. \tag{1.12}$$

Moreover

$$1 = \int_0^\infty f(t)dt + \lim_{t\to\infty}\|B_t\psi\|^2 \tag{1.13}$$

so $\lim_{t\to\infty}\|B_t\psi\|^2$ is interpreted as the probability that no event ever occurs.

If $\mathscr{H}$ is finite-dimensional then $f(t)$ may be computed by first finding the eigenvalues and eigenvectors of the dissipative matrix Y. The eigenvalues have negative real parts, as is appropriate for a decaying system. The existence of non-zero imaginary parts to the eigenvalues allows the probability density $f(t)$ to have strongly oscillatory behaviour.

The theory of one-parameter contraction semigroups on a Hilbert space has also been used as a model for the decay of an unstable particle. While this can be justified in terms of the theory of QSPs by supposing that one is waiting for the emission of decay products, it is perhaps less dogmatic to investigate the semigroups without such theoretical encumbrances. For the remainder of this chapter we shall therefore work entirely at the Hilbert space level.

7.2 Dilations of contraction semigroups

Let G be a group, $\mathscr{H}$ a Hilbert space and $T: G \to \mathscr{L}(\mathscr{H})$ an operator-valued function on G. We say that T is positive definite if for all $g_1, \ldots, g_N \in G$ and $\psi_1, \ldots, \psi_N \in \mathscr{H}$

$$\sum_{m,n=1}^{N} \langle T(g_m^{-1} g_n)\psi_n, \psi_m \rangle \geqslant 0. \tag{2.1}$$

By taking $N = 1, 2$ it may be seen that this implies

$$T(e) \geqslant 0, \qquad T(g^{-1}) = T(g)^*. \tag{2.2}$$

Theorem 2.1 *If U_g is a representation of the group G on the Hilbert space $\mathscr{K}$ and P is the projection of $\mathscr{K}$ onto a subspace $\mathscr{H}$ then the restriction T_g of PU_g to $\mathscr{H}$ is a positive definite $\mathscr{L}(\mathscr{H})$-valued function on G with $T(e) = 1$. If U_g is strongly continuous then T_g is strongly continuous.*

Conversely if T_g is a positive definite $\mathscr{L}(\mathscr{H})$-valued function on G with $T_e = 1$ then there exists a representation U_g of G on a Hilbert space $\mathscr{K}$ containing $\mathscr{H}$ such that if P is the projection of $\mathscr{K}$ onto $\mathscr{H}$ then T_g is the restriction of PU_g to $\mathscr{H}$.

The representation is unique up to isomorphism if

$$\mathscr{K} = \overline{\operatorname{lin}}\{U_g\mathscr{H}: g \in G\} \tag{2.3}$$

in which case it is called the minimal unitary dilation of T_g. If T_g is a weakly continuous function of g then the representation U_g is strongly continuous.

Proof Given U_g and P let $g_1, \ldots, g_N \in G$ and $\psi_1, \ldots, \psi_N \in \mathscr{H}$. Then

$$\begin{aligned}\sum_{m,n=1}^{N} \langle T(g_m^{-1} g_n)\psi_n, \psi_m \rangle &= \sum_{m,n=1}^{N} \langle U(g_m^{-1} g_n)\psi_n, \psi_m \rangle \\ &= \sum_{m,n=1}^{N} \langle U(g_n)\psi_n, U(g_m)\psi_m \rangle \\ &= \left\| \sum_{n=1}^{N} U(g_n)\psi_n \right\|^2 \geqslant 0 \end{aligned} \tag{2.4}$$

so T_g is positive definite.

Given the positive definite operator-valued function T_g let $\mathscr{M}$ be the vector space of all functions $f: G \to \mathscr{H}$ of finite support, that is functions f which equal zero except on a finite set. Then the formula

$$\langle f_1, f_2 \rangle = \sum_{g,h \in G} \langle T(g^{-1}h) f_1(h), f_2(g) \rangle \tag{2.5}$$

defines an inner product on $\mathscr{M}$ except that $\langle f, f \rangle = 0$ does not imply $f = 0$. If $\mathscr{N} = \{f : \langle f, f \rangle = 0\}$ then there is induced a positive definite inner product on $\mathscr{M}/\mathscr{N}$. We define $\mathscr{K}$ to be the Hilbert space completion of $\mathscr{M}/\mathscr{N}$. If $\psi \in \mathscr{H}$ we define $J\psi \in \mathscr{K}$ to be the element corresponding to the function $f \in \mathscr{M}$ defined by $f(g) = 0$ if $g \neq e$ and $f(e) = \psi$. Since $T_e = 1$, J is an isometric embedding of $\mathscr{H}$ into $\mathscr{K}$. If $g \in G$ and $f \in \mathscr{M}$ we define

$$(V_g f)(x) = f(g^{-1}x). \tag{2.6}$$

Since $V_g : \mathscr{M} \to \mathscr{M}$ is isometric and linear it induces a unitary map U_g on $\mathscr{M}/\mathscr{N}$ and hence on $\mathscr{K}$. Moreover this is actually a representation of G on $\mathscr{K}$. If $\phi, \psi \in \mathscr{H}$ then

$$\begin{aligned} \langle PU_g(J\phi), (J\psi) \rangle &= \langle V_g \delta_e \phi, \delta_e \psi \rangle \\ &= \langle \delta_g \phi, \delta_e \psi \rangle \\ &= \langle T_g \phi, \psi \rangle. \end{aligned} \tag{2.7}$$

So T_g is indeed the restriction of PU_g to $\mathscr{H}$, if we delete the explicit reference to J. Now $\mathscr{M}$ is generated by $\{\delta_g \phi : g \in G, \phi \in \mathscr{H}\}$ so $\mathscr{K}$ is generated by $\{U_g(J\phi) : g \in G, \phi \in \mathscr{H}\}$ and Eq. (2.3) holds for the space $\mathscr{K}$ constructed as above.

Suppose $\mathscr{K}'$, U'_g is another unitary dilation satisfying Eq. (2.3). We define $S: \mathscr{M} \to \mathscr{K}'$ by

$$Sf = \sum_{g \in G} U'_g f(g) \tag{2.8}$$

Then

$$\begin{aligned} \|Sf\|^2 = \sum_{g,h \in G} \langle U'_g f(g), U'_h f(h) \rangle &= \sum_{g,h \in G} \langle U'_{h^{-1}g} f(g), f(h) \rangle \\ &= \sum_{g,h \in G} \langle T(h^{-1}g) f(g), f(h) \rangle \\ &= \|f\|^2 \end{aligned} \tag{2.9}$$

so S induces an isometric linear map $S' : \mathscr{M}/\mathscr{N} \to \mathscr{K}'$ and hence $S'' : \mathscr{K} \to \mathscr{K}'$. Moreover if $\phi \in \mathscr{H}$ then

$$S''(U_g J\phi) = S(\delta_g \phi) = U'_g \phi = U'_g S''(J\phi). \tag{2.10}$$

It follows by Eq. (2.3) applied to $\mathscr{K}'$ that S'' is a unitary map of $\mathscr{K}$ onto $\mathscr{K}'$ which implements the equivalence of the two unitary representations.

We finally prove that if T_g is a weakly continuous function of g then the representation U_g on $\mathscr{K}$ is weakly continuous, and hence strongly continuous by Lemma 1.1. If $f_1, f_2 \in \mathscr{M}$ then

$$\begin{aligned}\langle V_g f_1, f_2\rangle &= \sum_{g_1, g_2 \in G} \langle T(g_2^{-1}g_1)f_1(g^{-1}g_1), f_2(g_2)\rangle \\ &= \sum_{g_1, g_2 \in G} \langle T(g_2^{-1}gg_1)f_1(g_1), f_2(g_2)\rangle \qquad (2.11)\end{aligned}$$

and so is a continuous function of g. Therefore $g \to \langle U_g f_1, f_2\rangle$ is a continuous function of g for all f_1, f_2 in the dense subspace $\mathscr{M}/\mathscr{N}$ of $\mathscr{K}$ and hence for all $f_1, f_2 \in \mathscr{K}$. QED

In order to apply the above theorem to one-parameter contraction semigroups we have to verify the technical hypothesis.

Theorem 2.2 *Let C be a contraction on the Hilbert space $\mathscr{H}$ and define*

$$C(-n)^* = C(n) = C^n \qquad (2.12)$$

for all $n \in \mathbb{Z}^+$. Then $\{C(n)\}$ is a positive definite operator-valued function on $\mathbb{Z}$.

Proof Let $\mathscr{K} = l^2(\mathbb{Z}^+, \mathscr{H})$ be the Hilbert space of all square-summable $\mathscr{H}$-valued sequences on $\mathbb{Z}^+$ and let B be any operator on $\mathscr{H}$ such that

$$B^*B + C^*C = 1. \qquad (2.13)$$

Then let $D\colon \mathscr{K} \to \mathscr{K}$ be the isometric linear operator defined by

$$(D\psi)_n = \begin{cases} C\psi_0 & \text{if } n = 0, \\ B\psi_0 & \text{if } n = 1, \\ \psi_{n-1} & \text{if } n \geqslant 2. \end{cases} \qquad (2.14)$$

If $\psi\colon \mathbb{Z} \to \mathscr{H}$ is a function of finite support and $\psi'\colon \mathbb{Z} \to \mathscr{K}$ is defined by

$$(\psi'_m)_n = \begin{cases} \psi_m & \text{if } n = 0, \\ 0 & \text{if } n \geqslant 1, \end{cases} \qquad (2.15)$$

then

$$\sum_{m, r \in \mathbb{Z}} \langle C(m-r)\psi_m, \psi_r\rangle = \sum_{m, r \in \mathbb{Z}} \langle D(m-r)\psi'_m, \psi'_r\rangle \qquad (2.16)$$

where

$$D(-m)^* = D(m) = D^m \qquad (2.17)$$

for all $m \in \mathbb{Z}^+$. Therefore it is sufficient to prove that $D(m)$ is a positive definite operator-valued function on $\mathscr{K}$.

Suppose that $\psi'_m = 0$ for m less than some $a \in \mathbb{Z}$. Then

$$\begin{aligned}\sum_{m,r\in\mathbb{Z}} \langle D(m-r)\psi'_m, \psi'_r\rangle &= \sum_{m\geqslant r\geqslant 0} \langle D^{m-r}\psi'_{m-a}, \psi'_{r-a}\rangle \\ &\quad + \sum_{r>m\geqslant 0} \langle \psi'_{m-a}, D^{r-m}\psi'_{r-a}\rangle \\ &= \sum_{m\geqslant r\geqslant 0} \langle D^m\psi'_{m-a}, D^r\psi'_{r-a}\rangle \\ &\quad + \sum_{r>m\geqslant 0} \langle D^m\psi'_{m-a}, D^r\psi'_{r-a}\rangle \end{aligned} \tag{2.18}$$

since D is an isometry,

$$\begin{aligned} &= \sum_{m,r\geqslant 0} \langle D^m\psi'_{m-a}, D^r\psi'_{r-a}\rangle \\ &\geqslant 0. \end{aligned}$$

QED (2.19)

Theorem 2.3 *Let B_t be a strongly continuous one-parameter contraction semigroup on $\mathscr{H}$ and let $B_{-t} = B_t^*$ for all $t \geqslant 0$. Then $\{B_t\}_{t\in\mathbb{R}}$ is a positive definite operator-valued function on $\mathbb{R}$.*

Proof Suppose that $t_n \in \mathbb{R}$ and $\psi_n \in \mathscr{H}$ are given for $1 \leqslant n \leqslant N$. If $r > 0$ let $s_r(n)$ be the integer which minimises $|t_n - s_r(n)/r|$, and let $C_r(s) = B(s/r)$, so that

$$C_r(-s)^* = C_r(s) = C_r(1)^s \tag{2.20}$$

for all $s \in \mathbb{Z}^+$. Then by the strong continuity of the semigroup B_t

$$\begin{aligned}\sum_{m,n=1}^{N} \langle B(t_n - t_m)\psi_n, \psi_m\rangle &= \lim_{r\to+\infty} \sum_{m,n=1}^{N} \left\langle B\left\{\frac{s_r(n)}{r} - \frac{s_r(m)}{r}\right\}\psi_n, \psi_m\right\rangle \\ &= \lim_{r\to+\infty} \sum_{m,n=1}^{N} \langle C_r\{s_r(n) - s_r(m)\}\psi_n, \psi_m\rangle \\ &\geqslant 0 \end{aligned} \tag{2.21}$$

by Theorem 2.2. QED

7.3 Spectrum of the Hamiltonian

Lemma 3.1 *Let U_t be the minimal unitary dilation on $\mathscr{K}$ of a strongly continuous one-parameter contraction semigroup B_t on $\mathscr{H}$ which is strict in the sense that*

$$\lim_{t\to+\infty} \|B_t\phi\| = 0 \tag{3.1}$$

for all $\phi \in \mathscr{H}$. *Then the closed linear subspace* $\mathscr{L}$ *of* $\mathscr{K}$ *defined by*

$$\mathscr{L} = \overline{\text{lin}}\{U_t\mathscr{H} : t \leqslant 0\} \tag{3.2}$$

satisfies

(*i*) $U_t\mathscr{L} \subseteq \mathscr{L}$ *for all* $t \leqslant 0$; (3.3)

(*ii*) $\bigcup_{t\in\mathbb{R}} U_t\mathscr{L}$ *is dense in* $\mathscr{K}$; (3.4)

(*iii*) $\bigcap_{t\in\mathbb{R}} U_t\mathscr{L} = 0.$ (3.5)

Proof Condition (*i*) is immediate while condition (*ii*) is a consequence of Eq. (2.3).

If $s, t \geqslant 0$ and $\phi, \psi \in \mathscr{H}$ then by the semigroup property of B_t

$$PU_{s+t}\phi = PU_sPU_t\phi \tag{3.6}$$

so

$$\begin{aligned}\langle(1 - P)U_t\phi, U_{-s}\psi\rangle &= \langle PU_s(1 - P)U_t\phi, \psi\rangle \\ &= 0\end{aligned} \tag{3.7}$$

which implies by Eq. (3.2) that

$$(1 - P)U_t\phi \in \mathscr{L}^{\perp}. \tag{3.8}$$

If $\phi \in \mathscr{H}$ and $s \in \mathbb{R}$ then by Eq. (3.1)

$$\begin{aligned}U_s\phi &= \lim_{x\to+\infty} U_s\{\phi - U_{-x}B_x\phi\} \\ &= \lim_{x\to+\infty} \{U_{s-x}(1 - P)U_x\phi\} \in \overline{\text{lin}}\left\{\bigcup_{t\in\mathbb{R}} U_t\mathscr{L}^{\perp}\right\}.\end{aligned} \tag{3.9}$$

By Eq. (2.3) it follows that

$$\overline{\text{lin}}\left\{\bigcup_{t\in\mathbb{R}} U_t\mathscr{L}^{\perp}\right\} = \mathscr{K} \tag{3.10}$$

which is equivalent to Eq. (3.5). QED

Theorem 3.2 *Let* $U_t = e^{iHt}$ *be the minimal unitary dilation on* $\mathscr{K}$ *of some strongly continuous strict one-parameter contraction semigroup. Then there exists an auxiliary Hilbert space* $\mathscr{N}$ *and an isomorphism*

$$\mathscr{K} \simeq L^2(\mathbb{R}, \mathscr{N}) \tag{3.11}$$

onto the space of square integrable $\mathscr{N}$*-valued functions, under which* U_t *becomes the operator of translation by a distance* t,

$$(U_tf)(x) = f(x + t). \tag{3.12}$$

Consequently the spectrum of H is absolutely continuous and equals the whole real line.

Proof If P_t is the projection of $\mathscr{K}$ onto $U_t\mathscr{H}$ then Eqs. (3.3), (3.4) and (3.5) imply that

$$(i) \quad \text{If } s \leqslant t \text{ then } P_s \leqslant P_t; \tag{3.13}$$

$$(ii) \quad \lim_{t \to +\infty} P_t = 1; \tag{3.14}$$

$$(iii) \quad \lim_{t \to -\infty} P_t = 0; \tag{3.15}$$

where the limits are in the strong operator topology. Also we obtain from the same equations

$$(iv) \quad \lim_{t \to t_0} P_t = P_{t_0}; \tag{3.16}$$

$$(v) \quad U_a P_t U_a^* = P_{t+a}. \tag{3.17}$$

If we define the projection-valued measure $P(\cdot)$ on $\mathbb{R}$ by

$$P(E) = \int_{-\infty}^{\infty} \chi_E(t) P(dt) \tag{3.18}$$

then

$$U_a P(E) U_a^* = P(E + a) \tag{3.19}$$

so $P(\cdot)$ and U define a system of imprimitivity. By von Neumann's uniqueness theorem for representations of the CCRs (A3) a representation of the desired type exists.

If we take Fourier transforms with respect to $\mathbb{R}$ of Eqs. (3.11) and (3.12) we obtain

$$\mathscr{K} \simeq L^2(\mathbb{R}, \mathscr{N}) \tag{3.20}$$

and

$$(U_t f)(x) = e^{itx} f(t) \tag{3.21}$$

for all $f \in L^2(\mathbb{R}, \mathscr{N})$. Therefore

$$(Hf)(x) = xf(x) \tag{3.22}$$

with

$$\operatorname{dom}(H) = \left\{ f : \int_{-\infty}^{\infty} (1 + x^2) \| f(x) \|^2 \, dx < \infty \right\} \tag{3.23}$$

from which the spectral properties of H may be readily deduced. QED

The conclusion of Theorem 3.2 might be taken as suggesting that the Hamiltonian H is unphysical. However, Hamiltonians which are not bounded below are typical in the study of infinite reservoirs in quantum statistical mechanics, so one may also consider that Theorem 3.2 suggests that a system with irreversible dynamics only arises by an interaction of the system with an infinite reservoir.

One does not actually need strict exponential decay to draw these conclusions about the Hamiltonian H.

Theorem 3.3 *Let H be a self-adjoint operator on a Hilbert space $\mathscr{H}$ and let $\phi \in \mathscr{H}$ be a unit vector such that for some constants A and $B > 0$*

$$|\langle e^{iHt}\phi, \phi\rangle| \leqslant A\, e^{-B|t|} \tag{3.24}$$

for all $t \in \mathbb{R}$. Then the spectrum of H equals the whole real line.

Proof By Bochner's theorem there exists a probability measure μ on $\mathbb{R}$ such that

$$\int_{\mathbb{R}} e^{ixt}\mu(dx) = \langle e^{iHt}\phi, \phi\rangle \equiv f(t). \tag{3.25}$$

If $\mathscr{H}_0 = L^2(\mathbb{R}, d\mu)$, $\phi_0(x) = 1$ for all $x \in \mathbb{R}$ and

$$(U_t^0\psi)(x) = e^{itx}\psi(x) \tag{3.26}$$

for all $\psi \in \mathscr{H}_0$ then

$$\langle e^{itH_0}\phi_0, \phi_0\rangle = \langle e^{itH}\phi, \phi\rangle \tag{3.27}$$

for all $t \in \mathbb{R}$. Therefore U_t^0 is a minimal unitary dilation of the positive definite function $f(t)$ and by Theorem 2.1 $\mathscr{H}_0$ can be identified with the subspace

$$\overline{\text{lin}}\{e^{itH}\phi : t \in \mathbb{R}\}. \tag{3.28}$$

The spectrum of H contains the spectrum of its restriction to $\mathscr{H}_0$, which equals the support of the measure μ. Now the Fourier transform $\hat{f}(z)$ of the function $f(t)$ is analytic on the strip

$$\{z : |\text{Im } z| < B\} \tag{3.29}$$

and by Fourier analysis the measure μ is

$$\mu(dx) = \hat{f}(x)dx \tag{3.30}$$

which is absolutely continuous with respect to the Lebesgue measure dx. Since $\hat{f}$ is analytic it vanishes only at isolated points in the strip so the support of μ equals the whole real line. QED

7.4 Absorption of a particle within a region

Let $\mathscr{H} = L^2(\mathbb{R}^3)$ and let $C_c^\infty(\Omega)$ denote the space of infinitely differentiable functions of compact support in an open subset Ω of $\mathbb{R}^3$. The operator H on $C_c^\infty(\mathbb{R}^3)$ defined by

$$(Hf)(x) = -\sum_{n=1}^{3} \frac{\partial^2 f}{\partial x_n^2} \tag{4.1}$$

is essentially self-adjoint. Now let V be a non-negative bounded function on $\mathbb{R}^3$ and let Ω and Ω' be two open sets such that if $x \in \Omega$ then $V(x) \gneq 0$ while if $x \in \Omega'$ then $V(x) = 0$. Suppose further that $\mathbb{R}^3 \setminus (\Omega \cup \Omega')$ has Lebesgue measure zero. If P is the projection onto $L^2(\Omega)$ and P' is the projection onto $L^2(\Omega')$ then $P' = 1 - P$.

In this section we try to find an expression for the time which elapses before a particle evolving under the Hamiltonian H first enters the region Ω. If the initial state is determined by the unit vector $\psi \in \mathscr{H}$ and if the particle is observed not to be in Ω at times $\{t/n, 2t/n, \ldots, t\}$, its state at time t is

$$\psi_n(t) = (P' e^{-iHt/n})^n \psi \tag{4.2}$$

so it seems plausible that the state at time t conditional upon not being in Ω at any time in $[0, t]$ is

$$\psi(t) = \lim_{n\to\infty} (P' e^{-iHt/n})^n \psi. \tag{4.3}$$

The probability that the particle has not entered the region Ω within the time $[0, t]$ is then $\|\psi(t)\|^2$.

It turns out that the above limit has very counter-intuitive properties. Since

$$P' = \lim_{\lambda\to+\infty} e^{-\lambda V t} \tag{4.4}$$

in the strong operator topology, we have

$$\psi(t) = \lim_{n\to+\infty} \lim_{\lambda\to+\infty} \{e^{-iHt/n} e^{-\lambda V t/n}\}^n \psi. \tag{4.5}$$

Now (A8) suggests that we interchange the order of the limits and look at the equation

$$\psi(t) = \lim_{\lambda\to+\infty} e^{(-iH-\lambda V)t} \psi \tag{4.6}$$

instead of Eq. (4.3). We do not claim that there is a rigorous passage from Eq. (4.3) to Eq. (4.6) but prefer to study the latter equation because it fits better into the framework of this book, and does actually have the same pathologies as Eq. (4.3). For fixed λ the state

$$\psi_\lambda(t) = e^{(-iH-\lambda V)t} \psi \tag{4.7}$$

is interpreted as the state at time $t > 0$ conditional upon the particle not having been absorbed up to time t, when there is an absorption rate $\lambda V(x)$ per unit time and volume at the point $x \in \mathbb{R}^3$.

The following three lemmas are devoted to an apparently entirely different problem, but will be used in the proof to Theorem 4.5.

Lemma 4.1 *If $\lambda \in \mathbb{R}^+$, the self-adjoint operators $(H + \lambda V + 1)^{-1}$ decrease monotonically as $\lambda \to +\infty$ and converge strongly to an operator B on $\mathscr{H}$ such that $0 \leqslant B \leqslant 1$.*

Proof The operators $(H + \lambda V)$ all have the same domain $\mathscr{D}$ and their spectra all lie within $[0, \infty)$. Therefore the operators

$$B_\lambda = (H + \lambda V + 1)^{-1} \tag{4.8}$$

are bounded and self-adjoint with $0 \leqslant B_\lambda \leqslant 1$.

If $\lambda_1 \leqslant \lambda_2$ and $\psi \in \mathscr{H}$ and $v_i = B_{\lambda_i}\psi$ then $v_i \in \mathscr{D}$ and

$$\begin{aligned} |\langle B_{\lambda_2}\psi, \psi\rangle|^2 &= |\langle v_2, (H + \lambda_1 V + 1)v_1\rangle|^2 \\ &\leqslant \langle (H + \lambda_1 V + 1)v_2, v_2\rangle\langle (H + \lambda_1 V + 1)v_1, v_1\rangle \\ &\leqslant \langle (H + \lambda_2 V + 1)v_2, v_2\rangle\langle (H + \lambda_1 V + 1)v_1, v_1\rangle \\ &= \langle \psi, B_{\lambda_2}\psi\rangle\langle \psi, B_{\lambda_1}\psi\rangle. \end{aligned} \tag{4.9}$$

Therefore

$$0 \leqslant \langle B_{\lambda_2}\psi, \psi\rangle \leqslant \langle B_{\lambda_1}\psi, \psi\rangle \tag{4.10}$$

so $0 \leqslant B_{\lambda_2} \leqslant B_{\lambda_1}$. Since a monotonically decreasing family of operators which is uniformly norm bounded converges weakly and strongly the proof is finished. QED

Lemma 4.2 *There exists a self-adjoint operator $H' \geqslant 0$ on $L^2(\Omega')$ such that*

$$B\psi = (1 + H')^{-1}\psi \tag{4.11}$$

for all $\psi \in L^2(\Omega')$. Moreover if $\psi \in L^2(\Omega)$ then

$$B\psi = 0. \tag{4.12}$$

The domain of H' contains $C_c^\infty(\Omega')$ and

$$H'\psi = H\psi \tag{4.13}$$

for all $\psi \in C_c^\infty(\Omega')$.

Proof If $\phi \in C_c^\infty(\Omega')$ then $\phi \in \mathrm{dom}(H)$ and $V\phi = 0$. Therefore $\psi = (H + 1)\phi$ satisfies

$$(H + \lambda V + 1)^{-1}\psi = \phi \tag{4.14}$$

for all $\lambda \geqslant 0$, so that

$$B\psi = \phi \tag{4.15}$$

and

$$C_c^\infty(\Omega') \subseteq \text{Range } B. \tag{4.16}$$

Since

$$\langle (H + \lambda V + 1)\phi, \phi \rangle \geqslant \langle (\lambda V + 1)\phi, \phi \rangle \tag{4.17}$$

for all $\phi \in \mathscr{D}$, we can conclude that

$$0 \leqslant B_\lambda \leqslant (\lambda V + 1)^{-1} \tag{4.18}$$

as in the proof of Eq. (4.10). If $\phi \in L^2(\Omega)$ it follows that

$$0 \leqslant \langle B\phi, \phi \rangle \leqslant \lim_{\lambda \to +\infty} \langle (\lambda V + 1)^{-1}\phi, \phi \rangle = 0 \tag{4.19}$$

so

$$B\phi = 0. \tag{4.20}$$

Now for any self-adjoint operator B

$$(\text{Range } B)^\perp = \{\phi \colon B\phi = 0\} \tag{4.21}$$

so by Eqs (4.16) and (4.20)

$$\{\phi \colon B\phi = 0\} = L^2(\Omega). \tag{4.22}$$

From Eq. (4.22) and the fact that $\psi \in C_c^\infty(\Omega')$ it follows that if H' is defined by Eq. (4.11) then

$$(1 + H')\phi = \psi \tag{4.23}$$

which proves Eq. (4.13). QED

Lemma 4.3 *If $\psi \in L^2(\Omega')$ and $t > 0$ then*

$$\lim_{\lambda \to +\infty} e^{-(iH + i\lambda V)t}\psi = e^{iH't}\psi. \tag{4.24}$$

Proof Define the bounded function f on $[0, 1]$ by

$$f\{(1 + x)^{-1}\} = e^{-itx} \tag{4.25}$$

if $x \in \mathbb{R}^+$ and $f(0) = 0$. Then f is continuous except at zero and

$$f(B_\lambda)\psi = e^{-(iH + i\lambda V)t}\psi \tag{4.26}$$

for all $\psi \in L^2(\mathbb{R}^3)$. Moreover

$$f(B)\psi = e^{iH't}\psi \tag{4.27}$$

for all $\psi \in L^2(\Omega')$, while

$$f(B)\psi = 0 \tag{4.28}$$

for all $\psi \in L^2(\Omega)$. Since B_λ converges strongly to B, it follows by (A2) that

$$\lim_{\lambda \to +\infty} f(B_\lambda)\psi = f(B)\psi \tag{4.29}$$

for all $\psi \in L^2(\Omega')$. QED

Lemma 4.4 *Let f be a continuous $\mathscr{H}$-valued function on $S = \{z \in \mathbb{C}: \operatorname{Im} z \leqslant 0\}$ which is analytic in the interior of this region. If $\|f(z)\| \leqslant a$ for all $z \in S$ and $|f(z)| \leqslant \varepsilon$ for all $1 \leqslant z \in \mathbb{R}^+$ then*

$$\|f(-i)\| \leqslant \varepsilon^{1/4} a^{3/4}. \tag{4.30}$$

Proof If $\psi \in \mathscr{H}$ and $\|\psi\| = 1$ we define the complex-valued analytic function g on the disc $D = \{z \in \mathbb{C}: |z| \leqslant 1\}$ by

$$g(z) = \left\langle f\left(i\frac{z-1}{z+1}\right), \psi \right\rangle \tag{4.31}$$

Then $|g(z)| \leqslant a$ for all $z \in D$ and $|g(e^{i\theta})| \leqslant \varepsilon$ if $-(\pi/2) \leqslant \theta \leqslant 0$. Applying the maximum modulus theorem to

$$h(z) = \prod_{r=1}^{4} g(i^r z) \tag{4.32}$$

we find that

$$|\langle f(-i), \psi \rangle|^4 = |g(0)|^4 = |h(0)| \leqslant \varepsilon a^3. \tag{4.33}$$

But ψ is arbitrary so Eq. (4.30) follows. QED

Theorem 4.5 *If $\psi \in L^2(\Omega')$ and $t > 0$ then*

$$\lim_{\lambda \to +\infty} e^{(-iH-\lambda V)t}\psi = e^{-itH'}\psi. \tag{4.34}$$

Proof By Eq. (1.9.7) the $\mathscr{H}$-valued function

$$f_\lambda(z) = e^{(-iH-iz\lambda V)t}\psi \tag{4.35}$$

is analytic in z for all $\lambda > 0$. Moreover if $\operatorname{Im} z \leqslant 0$ then $\{-iH - iz\lambda V\}$ is dissipative so by Lemma 1.2

$$\|f_\lambda(z)\| \leqslant \|\psi\|. \tag{4.36}$$

By Lemma 4.3 if $\varepsilon > 0$ there exists $K(\varepsilon) \in \mathbb{R}^+$ such that if $\lambda \geqslant K(\varepsilon)$ and $z \geqslant 1$ then

$$\|f_\lambda(z) - e^{-iH't}\psi\| < \varepsilon. \tag{4.37}$$

By Lemma 4.4 it follows that if $\lambda \geqslant K(\varepsilon)$ then

$$\|f_\lambda(-i) - e^{-iH't}\psi\| \leqslant \varepsilon^{1/4}2^{3/4} \tag{4.38}$$

which implies Eq. (4.34). QED

The above theorem implies that if the initial state is given by a unit vector $\psi \in L^2(\Omega')$ then

$$\lim_{\lambda \to +\infty} \|e^{(-iH-\lambda V)t}\psi\| = 1 \tag{4.39}$$

so that in the limit $\lambda \to +\infty$ there is no absorption. This is superficially in conflict with the equation

$$\frac{d}{dt}\|\psi_\lambda(t)\|^2 = -2\lambda\langle V\psi_\lambda(t), \psi_\lambda(t)\rangle \tag{4.40}$$

and shows how careful one must be in interpreting operator differential equations.

The explanation for this result is that although the particle is gradually absorbed within Ω it is also partially reflected at the surface of Ω. As $\lambda \to +\infty$ the reflection coefficient at the surface converges to unity, so in the limit the particle never enters the absorbing region.

Our conclusion is that one may calculate the probability of absorption up to time t for any choice of the absorption function V and parameter λ. However, there seems to be no sensible definition of this probability which depends only on the Hamiltonian H and the set Ω.

Notes

1 A Banach space version of Lemma 1.2 was proved by Lumer and Phillips (1961). Contraction semigroups have been used to discuss decay of unstable particles by Allcock (1969), Eckstein and Siegert (1971) and Williams (1971), among many others. A comparison of quantum-mechanical and probabilistic calculations of waiting times was made by Thomas (1974).

2 This theory is due to Sz-Nagy and we can do no better than refer the reader to the very thorough account in Sz-Nagy and Foias (1970). For the case where B_t is a one-parameter semigroup of isometries see Cooper (1947).

3 This theory has been developed under many different names and we refer the reader to Lax and Phillips (1967), Sz-Nagy and Foias (1970) and Lewis and Thomas (1974) for further references and for the more complicated theory when the semigroup B_t is not strict. Theorem 3.2 was proved by Sinai (1964). A proof of Theorem 3.3 and related results may be found in Williams (1971), Sinha (1972), and Horwitz *et al.* (1971).

4 Friedman (1971/2) approached this problem via Eq. (4.3). Lemmas 4.1, 4.2 and 4.3 are small modifications of results in Kato (1966). Davies (1975b) proved Theorem 4.5, and also gave an exact characterisation of H' in quadratic form terms. The explanation of Theorem 4.5 in terms of reflection at the boundary of Ω was given by Allcock (1969). The rather different approach to decay of Ekstein and Siegert (1971) is also studied in Davies (1975b).

8

Quantum and Classical Fields

8.1 Fock space

We turn to the application of operational methods in quantum field theory. Many of the ideas introduced below have been developed in quantum optics, but we present them here for a general boson quantum field, without specifying the type of particle.

Given a single particle Hilbert space $\mathscr{H}$ the boson Fock space $\mathscr{F}$ is

$$\mathscr{F} = \mathbb{C} \oplus \mathscr{H} \oplus (\otimes^2_{\text{sym}} \mathscr{H}) \oplus \cdots \tag{1.1}$$

the nth term $\otimes^n_{\text{sym}} \mathscr{H}$ being the symmetric subspace of the nth tensor product of $\mathscr{H}$, completed in norm so as to be a Hilbert space. If U is a unitary operator on $\mathscr{H}$ there is a unitary operator $U^\sim$ on $\mathscr{F}$ whose restriction to $\otimes^n_{\text{sym}} \mathscr{H}$ is $\otimes^n U$. If S is a single particle Hamiltonian on $\mathscr{H}$ the free time evolution on $\mathscr{F}$ is given by

$$e^{iH_0 t} = (e^{iSt})^\sim. \tag{1.2}$$

The operator H_0 is given on the n-particle subspace by the expression

$$\begin{aligned}(H_0)_n = (S \otimes 1 \otimes 1 \otimes \cdots \otimes 1) + (1 \otimes S \otimes 1 \otimes \cdots \otimes 1) \\ + \cdots + (1 \otimes \cdots \otimes 1 \otimes S).\end{aligned} \tag{1.3}$$

For each $\psi \in \mathscr{H}$ we define a unit vector $\psi^\sim \in \mathscr{F}$, called a (Glauber) coherent state by

$$(\psi^\sim)_n = e^{-\|\psi\|^2/2}(n!)^{-1/2} \otimes^n \psi. \tag{1.4}$$

A direct calculation leads to

$$\langle \phi^\sim, \psi^\sim \rangle = \exp\left\{\langle \phi, \psi \rangle - \frac{\|\phi\|^2}{2} - \frac{\|\psi\|^2}{2}\right\} \tag{1.5}$$

and to

$$U^\sim \phi^\sim = (U\phi)^\sim \tag{1.6}$$

for all unitary operators U on $\mathscr{H}$.

Lemma 1.1 *The linear span of the coherent states is dense in $\mathscr{F}$.*

Proof If L is the closed linear span of the coherent states then

$$\int_0^{2\pi} e^{-in\theta}(e^{i\theta}\psi)^{\sim}\, d\theta = 2\pi\, e^{-\|\psi\|^2/2}(n!)^{-1/2} \otimes^n \psi \tag{1.7}$$

so $\otimes^n \psi \in L$ for all $\psi \in \mathscr{H}$. If $\psi_1, \ldots, \psi_n \in \mathscr{H}$ and $t_1, \ldots, t_n \in \mathbb{C}$ then

$$\otimes^n \left(\sum_{i=1}^n t_i \psi_i \right) \tag{1.8}$$

is a polynomial in $t_1, \ldots, t_n$ such that the coefficient of $\prod_{i=1}^n t_i$ is

$$\sum_\pi \psi_{\pi(1)} \otimes \cdots \otimes \psi_{\pi(n)} \tag{1.9}$$

where the sum is over all permutations π of $1, \ldots, n$. Therefore

$$\sum_\pi \psi_{\pi(1)} \otimes \cdots \otimes \psi_{\pi(n)} = \frac{\partial}{\partial t_1} \cdots \frac{\partial}{\partial t_n} \otimes^n \left(\sum_{i=1}^n t_i \psi_i \right)\Bigg|_{t_1 = \cdots = t_n = 0} \tag{1.10}$$

and lies in L. But such vectors span a dense subspace of $\otimes^n_{\text{sym}} \mathscr{H}$, so L is dense in $\mathscr{F}$. QED

The coherent states allow one to give a very quick definition of the Weyl operators $W(f)$.

Lemma 1.2 *For each $f \in \mathscr{H}$ there is a unique unitary operator $W(f)$ on $\mathscr{F}$ such that*

$$W(f)(\psi^{\sim}) = e^{i2^{-1/2}\operatorname{Re}\langle \psi, f\rangle}(\psi + i2^{-1/2} f)^{\sim} \tag{1.11}$$

for all $\psi \in \mathscr{H}$. The map $f \to W(f)$ is continuous from $\mathscr{H}$ to the set of unitary operators on $\mathscr{F}$ with the strong operator topology.

Proof By direct computation it may be seen that Eq. (1.11) defines a map of the set of coherent states into $\mathscr{F}$ such that

$$\langle W(f)\phi^{\sim}, W(f)\psi^{\sim} \rangle = \langle \phi^{\sim}, \psi^{\sim} \rangle \tag{1.12}$$

for all $\phi, \psi \in \mathscr{H}$. $W(f)$ therefore has a unique extension to an isometric linear map on the linear span L of the coherent states and hence to an isometry on $\mathscr{F}$. The range of $W(f)$ contains L and is closed, so $W(f)$ is unitary. If $\phi \in \mathscr{F}$ is a coherent state then $f \to W(f)\phi$ is norm continuous by inspection. The same holds for $\phi \in L$ and hence for all $\phi \in \mathscr{F}$ by Lemma 1.1. QED

By a computation on the coherent states one finds that

$$W(f)W(g) = W(f + g)\exp\left\{\frac{i\,\mathrm{Im}\langle f, g\rangle}{2}\right\} \tag{1.13}$$

because of which one says that $W(\cdot)$ is a reprcsentation of thc canonical commutation relations (CCRs). If the Fock space vacuum Ω is taken to be $0^{\sim}$, that is

$$\Omega = 1 \oplus 0 \oplus 0 \oplus \cdots \tag{1.14}$$

then

$$W(f)\Omega = (i2^{-1/2}f)^{\sim} \tag{1.15}$$

because of which the coherent states are also called displaced vacuum states. One finds by direct computation that

$$\langle W(f)\Omega, \Omega\rangle = \exp\left\{-\frac{\|f\|^2}{4}\right\} \tag{1.16}$$

a formula which is known to characterise the Fock space representation of the CCRs up to unitary equivalence. If H_0 is a free Hamiltonian on $\mathscr{F}$ the time evolution of the Weyl operators is given by

$$e^{itH_0}W(f)e^{-itH_0} = W(e^{iSt}f) \tag{1.17}$$

this again being proved by direct computation for the coherent states.

We turn now to the question of tensor products of Fock spaces. Our main theorem is the following.

Theorem 1.3 *If $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$ there is a unique unitary map B' of $\mathscr{F}_1 \otimes \mathscr{F}_2$ onto $\mathscr{F}$ such that*

$$B'(\phi^{\sim} \otimes \psi^{\sim}) = (\phi + \psi)^{\sim} \tag{1.18}$$

for all $\phi \in \mathscr{H}_1$ and $\psi \in \mathscr{H}_2$.

Proof If for $i = 1, 2$ we let $L_i = \mathrm{lin}\{\phi^{\sim} : \phi \in \mathscr{H}_i\}$ then L_i is a dense subspace of $\mathscr{F}_i$. If $B: L_1 \times L_2 \to \mathscr{F}$ is defined by

$$B\left(\sum_{i=1}^{n} \alpha_i \phi_i^{\sim}, \sum_{j=1}^{n} \beta_j \psi_j^{\sim}\right) = \sum_{i,j} \alpha_i \beta_j (\phi_i + \psi_j)^{\sim} \tag{1.19}$$

then it is an immediate consequence of Eq. (1.5) that B satisfies Eq. (1.10.3). Therefore there exists an isometric linear map $B' : \mathscr{F}_1 \otimes \mathscr{F}_2 \to \mathscr{F}$ satisfying Eq. (1.18). The range of B' contains all $\phi^{\sim}$ with $\phi \in \mathscr{H}$, so B' is unitary. QED

We shall in future identify $\mathscr{F}_1 \otimes \mathscr{F}_2$ with $\mathscr{F}$ instead of continually referring to the unitary map B'.

8.2 Classical states

If $\mathscr{H}$ is a separable Hilbert space we define a probability measure μ on $\mathscr{H}$ to be a positive countably additive measure with total mass one on the σ-field of Borel sets in $\mathscr{H}$. For every such measure μ we define a state ρ_μ on the Fock space by

$$\rho_\mu = \int_{\mathscr{H}} |\psi^\sim\rangle\langle\psi^\sim|\mu(d\psi). \tag{2.1}$$

States constructed in this manner are called classical states and form a convex subset of the state space V of $\mathscr{F}$.

Lemma 2.1 *The only classical states on $\mathscr{F}$ which are also pure states are those of the form $|\psi^\sim\rangle\langle\psi^\sim|$ for some $\psi \in \mathscr{H}$.*

Proof If μ is not the evaluation at some point of $\mathscr{H}$, there exists a Borel set E in $\mathscr{H}$ such that $0 < \mu(E) < 1$. If the probability measures μ_1, μ_2 on $\mathscr{H}$ are defined by

$$\mu_1(F) = \frac{\mu(F \cap E)}{\mu(E)} \tag{2.2}$$

and

$$\mu_2(F) = \frac{\mu(F \setminus E)}{\mu(\mathscr{H} \setminus E)} \tag{2.3}$$

then $\mu_1 \neq \mu_2$ and

$$\rho_\mu = \mu(E)\rho_{\mu_1} + \{1 - \mu(E)\}\rho_{\mu_2} \tag{2.4}$$

If ρ_μ is pure then $\rho_{\mu_1} = \rho_{\mu_2}$ so

$$\langle \rho_{\mu_1}\phi^\sim, \phi^\sim\rangle = \langle \rho_{\mu_2}\phi^\sim, \phi^\sim\rangle \tag{2.5}$$

for all $\phi \in \mathscr{H}$. By Eq. (1.5) this implies that

$$\int_{\mathscr{H}} e^{-\|\phi-\psi\|^2}\mu_1(d\psi) = \int_{\mathscr{H}} e^{-\|\phi-\psi\|^2}\mu_2(d\psi) \tag{2.6}$$

for all $\phi \in \mathscr{H}$. If $\mathscr{H}_1$ is a finite dimensional subspace of $\mathscr{H}$ and $\mathscr{H}_2$ its orthogonal complement we put $\psi = \psi_1 + \psi_2$ where $\psi_i \in \mathscr{H}_i$ for $i = 1, 2$. We then define the measures ν_i on $\mathscr{H}_1$ by

$$\nu_i(d\psi_1) = \int_{\mathscr{H}_2} e^{-\|\psi_2\|^2}\mu_i(d\psi_1\, d\psi_2) \tag{2.7}$$

so that

$$\int_{\mathscr{H}_1} e^{-\|\phi-\psi_1\|^2}\nu_1(d\psi_1) = \int_{\mathscr{H}_1} e^{-\|\phi-\psi_1\|^2}\nu_2(d\psi_1) \tag{2.8}$$

for all $\phi \in \mathscr{H}_1$. Since $\mathscr{H}_1$ is finite-dimensional we can take Fourier transforms and deduce that $\nu_1 = \nu_2$. Thus

$$\mu_1(F) = \mu_2(F) \tag{2.9}$$

for all cylinder sets F in $\mathscr{H}$. It follows by Theorem 1.11.2 that $\mu_1 = \mu_2$. The contradiction implies that ρ_μ is not pure. QED

If one allows linear combinations of classical states with negative coefficients, however, then one can obtain a close approximation to any state.

Lemma 2.2 *The set L of linear combinations*

$$\sum_{r=1}^{n} \lambda_r |\psi_r^\sim\rangle\langle\psi_r^\sim| \tag{2.10}$$

where $\lambda_r \in \mathbb{R}$ and $\psi_r \in \mathscr{H}$, is dense in V for the trace norm topology.

Proof If $\mathscr{H}_n$ is an increasing sequence of finite-dimensional subspaces in $\mathscr{H}$ whose union is dense in $\mathscr{H}$, then $V_n = \mathscr{T}_s(\mathscr{F}_n)$ is an increasing sequence in $V = \mathscr{T}_s(\mathscr{F})$ whose union is dense in V for the trace norm. It is therefore sufficient to prove this lemma if $\mathscr{H} = \mathbb{C}^n$ with the standard inner product.

If L is not dense in V then by the Hahn–Banach theorem and Eq. (1.5.4) there exists a non-zero self-adjoint operator A on $\mathscr{F}$ such that

$$\langle Az^\sim, z^\sim\rangle = 0 \tag{2.11}$$

for all $z \in \mathbb{C}^n$. If f is the complex-valued function defined by

$$f(z, w) = \langle Az^\sim, (\overline{w})^\sim\rangle e^{\|z\|^2/2+\|w\|^2/2} \tag{2.12}$$

then by Eq. (1.4) f is analytic on $\mathbb{C}^{2n}$. If we define

$$g(z, w) = f(z + iw, z - iw) \tag{2.13}$$

then g is analytic and by Eq. (2.11) $g(z, w) = 0$ if z and w are both real. Examining its power series expansion we see that $g \equiv 0$ so

$$\langle Az^\sim, w^\sim\rangle = 0 \tag{2.14}$$

for all $z, w \in \mathbb{C}^n$. By Lemma 1.1 it follows that $A = 0$. The contradiction implies that L is indeed dense in V. QED

There are two main reasons for the usefulness of the classical states in quantum optics. The first is the fact that it has proved possible to make very good approximations to the states appropriate to many different types of light source by using classical states. The second is that classical states enable one to set up a close relationship between classical and quantum optics.

Both of these points are most simply illustrated in the case of a single mode field, where $\mathscr{H} = \mathbb{C}$ and $\mathscr{F} = l^2(\mathbb{Z}^+)$ is the Hilbert space of a harmonic oscillator. We recall that the Gibbs state at the inverse temperature β is the state ρ_β on $\mathscr{F}$ with matrix

$$(\rho_\beta)_{mn} = (1 - e^{-\beta})e^{-n\beta}\delta_{mn}. \tag{2.15}$$

Theorem 2.3 *The Gibbs state ρ_β is the classical state associated with the Gaussian probability density*

$$\pi^{-1}(e^\beta - 1)\exp\{-(e^\beta - 1)(x^2 + y^2)\} \tag{2.16}$$

on $\mathbb{R}^2$.

Proof If e_m is the mth element of the standard orthonormal basis of $l^2(\mathbb{Z}^+)$ and ρ is the classical state associated with the probability density

$$(2\pi)^{-1}\alpha \exp\left\{-\frac{\alpha(x^2 + y^2)}{2}\right\} \tag{2.17}$$

then

$$\begin{aligned}
\langle \rho e_m, e_n \rangle &= \int_{\mathbb{R}^2} \langle z^\sim, e_n\rangle \langle e_m, z^\sim\rangle \frac{\alpha}{2\pi} e^{-\alpha(x^2+y^2)/2}\, dx\, dy \\
&= \int_{\mathbb{R}^2} e^{-(x^2-y^2)(1+\alpha/2)} z^n \bar{z}^m \alpha n!^{-1/2} m!^{-1/2} (2\pi)^{-1}\, dx\, dy \\
&= \delta_{mn}(n!)^{-1} \int_0^\infty e^{\;\; r^2(1-\alpha/2)} r^{2n+1} \alpha\, dr \\
&= \delta_{mn} \frac{\alpha(1+\alpha/2)^{-n-1}}{2} \\
&= \delta_{mn}(1 - e^{-\beta})e^{-n\beta}
\end{aligned} \tag{2.18}$$

if

$$1 + \frac{\alpha}{2} = e^\beta. \qquad \text{QED} \tag{2.19}$$

We now describe the relevance of the classical states to the relationship between the classical and quantum-mechanical harmonic oscillators. The classical oscillator has phase space $\mathbb{R}^2$ with Hamiltonian equations

$$\frac{dq}{dt} = p; \qquad \frac{dp}{dt} = -\omega^2 q. \tag{2.20}$$

If we put

$$z = \omega^{1/2} q + i\omega^{-1/2} p \tag{2.21}$$

then

$$\frac{dz}{dt} = -i\omega z \tag{2.22}$$

so

$$z(t) = e^{-i\omega t} z(0) \tag{2.23}$$

Hamilton's equations therefore describe a motion on the phase space, which induces a corresponding action on the probability measures on the phase space.

We recall that the Hamiltonian $S = \omega 1$ on the Hilbert space $\mathscr{H} = \mathbb{C}$ induces a free Hamiltonian H_0 on the Fock space by Eq. (1.2).

Theorem 2.4 *The time evolution of the probability measures on phase space under the classical Hamiltonian is consistent with the time evolution of the classical states under the free quantum mechanical Hamiltonian H_0.*

Proof If μ is a probability measure on $\mathbb{C}$ and

$$\rho = \int_{\mathbb{C}} |z^{\sim}\rangle\langle z^{\sim}| \mu(dx\, dy) \tag{2.24}$$

then by Eq. (1.6)

$$\begin{aligned} e^{-itH_0}\rho e^{itH_0} &= \int_{\mathbb{C}} |(e^{-it\omega}z)^{\sim}\rangle\langle(e^{-it\omega}z)^{\sim}| \mu(dx\, dy) \\ &= \int_{\mathbb{C}} |z^{\sim}\rangle\langle z^{\sim}| \mu_t(dx\, dy) \end{aligned} \tag{2.25}$$

where μ_t is the rotated measure, given by

$$\mu_t(E) = \mu(e^{it\omega} E) \tag{2.26}$$

for all Borel sets $E \subseteq \mathbb{C}$. QED

The above "correspondence principle" has extensions to quantum fields with arbitrarily many modes. It may be used both for extending the scope of classical ideas to quantum theory and for re-interpreting quantum-mechanical ideas in classical terms. Although it might be argued that the correspondence principle obviates the need for the existence of both classical and quantum optics, historically each of these fields has led to valuable insights into the subject.

8.3 Evolution of quantum fields with fluctuations

One frequently wants to discuss the time evolution of a quantum-mechanical system whose Hamiltonian $H(\omega)$ depends upon a random element ω. In quantum optics this occurs, for example, if the radiation is passing through a medium with randomly distributed refractive index. If the state at time zero is ρ then the expected state at time t is

$$\bar{\rho}_t = \int e^{-iH(\omega)t}\rho\, e^{iH(\omega)t}\, d\omega. \tag{3.1}$$

The map $\rho \to \bar{\rho}_t$ is then a positive linear map on the state space of the system. Phenomenological models for these positive linear maps have been constructed in quantum optics and are of a few general types which we now describe.

We suppose that V is the state space of a boson Fock space $\mathscr{F}$ and denote by $\rho(\mu)$ the classical state associated with a probability measure μ on the single particle space $\mathscr{H}$. If μ and ν are two probability measures on $\mathscr{H}$ we denote their convolution by $\mu \circ \nu$; if $\mathscr{H}$ is finite-dimensional and μ, ν have densities f, g with respect to Lebesgue measure then $\mu \circ \nu$ has the density $f \circ g$.

Theorem 3.1 *For every probability measure ν on $\mathscr{H}$ there is a unique positive linear map $T: V \to V$ such that*

$$\operatorname{tr}[T(\rho)] = \operatorname{tr}[\rho] \tag{3.2}$$

for all $\rho \in V$ and

$$T\{\rho(\mu)\} = \rho(\mu \circ \nu) \tag{3.3}$$

for every probability measure μ on $\mathscr{H}$.

Proof By Eq. (1.11)

$$W(-i2^{1/2}\phi)|\psi^\sim\rangle\langle\psi^\sim|W(-i2^{1/2}\phi)^* = |(\phi+\psi)^\sim\rangle\langle(\phi+\psi)^\sim| \tag{3.4}$$

for all $\phi, \psi \in \mathscr{H}$. Therefore if we define

$$T(\rho) = \int W(-i2^{1/2}\phi)\rho W(-i2^{1/2}\phi)^* \nu(d\phi) \tag{3.5}$$

T is a positive linear map satisfying Eq. (3.2) and if μ is any probability measure on $\mathscr{H}$

$$T\{\rho(\mu)\} = \int_{\mathscr{H}\times\mathscr{H}} |(\phi + \psi)^{\sim}\rangle\langle(\phi + \psi)^{\sim}|\mu(d\psi)\nu(d\phi)$$

$$= \int_{\mathscr{H}} |\phi^{\sim}\rangle\langle\phi^{\sim}|(\mu \circ \nu)(d\phi)$$

$$= \rho(\mu \circ \nu). \tag{3.6}$$

The uniqueness of T is an immediate consequence of Lemma 2.2. QED

Theorem 3.2 *For every linear contraction B on $\mathscr{H}$ there exists a unique positive linear map $T: V \to V$ such that*

$$\mathrm{tr}[T(\rho)] = \mathrm{tr}[\rho] \tag{3.7}$$

for all $\rho \in V$ and

$$T\{\rho(\mu)\} = \rho(\mu_B) \tag{3.8}$$

for every probability measure μ on $\mathscr{H}$, where μ_B is the probability measure such that

$$\mu_B(E) = \mu(B^{-1}(E)) \tag{3.9}$$

for every Borel set $E \subseteq \mathscr{H}$.

Proof Uniqueness again follows from Lemma 2.2 so we need only prove existence.

By Theorems 7.2.1 and 7.2.2 there exists a Hilbert space $\mathscr{H}_2$ such that if $\mathscr{H} = \mathscr{H}_1$ and $\mathscr{H}_0 = \mathscr{H}_1 \oplus \mathscr{H}_2$ and P is the projection of $\mathscr{H}_0$ onto $\mathscr{H}_1$ then $B\phi_1 = PU\phi_1$ for all $\phi_1 \in \mathscr{H}_1$ and some unitary operator U on $\mathscr{H}_0$. If we put $\phi_2 = (1 - P)U\phi_1$ then $U\phi_1 = B\phi_1 + \phi_2$ so

$$U^{\sim}\phi_1^{\sim} = (U\phi_1)^{\sim} = (B\phi_1)^{\sim} \otimes \phi_2^{\sim} \tag{3.10}$$

by Theorem 1.3, where $U^{\sim}$ is a unitary operator on $\mathscr{F}_0 = \mathscr{F}_1 \otimes \mathscr{F}_2$.

Now the formula $A \to A \otimes 1$ defines a normal positive linear map from $\mathscr{L}(\mathscr{F}_1)$ into $\mathscr{L}(\mathscr{F}_0)$ and by Lemma 2.2.2 this map is the adjoint of a positive linear map $S: V_0 \to V_1$. In the other direction the map $\phi \to \phi \otimes 0^{\sim}$ is an isometric linear map from $\mathscr{F}_1$ into $\mathscr{F}_0$ and so induces a positive isometric linear map $J: V_1 \to V_0$. If we define $T: V_1 \to V_1$ by

$$T(\rho) = S\{U^{\sim}(J\rho)U^{\sim *}\} \tag{3.11}$$

then T is a composition of positive linear trace-preserving maps and therefore one such itself. If $\phi_1 \in \mathscr{H}_1$ then for all $A \in \mathscr{L}_s(\mathscr{F}_1)$

$$\begin{aligned}\operatorname{tr}[AT(|\phi_1^\sim\rangle\langle\phi_1^\sim|)] &= \operatorname{tr}[(A\otimes 1)|(U\phi_1)^\sim\rangle\langle(U\phi_1)^\sim|] \\ &= \operatorname{tr}[(A\otimes 1)|(B\phi_1)^\sim \otimes \phi_2^\sim\rangle\langle(B\phi_1)^\sim \otimes \phi_2^\sim|] \\ &= \langle A(B\phi_1)^\sim, (B\phi_1)^\sim\rangle\langle 1\phi_2^\sim, \phi_2^\sim\rangle \\ &= \operatorname{tr}[A|(B\phi_1)^\sim\rangle\langle(B\phi_1)^\sim|] \qquad (3.12)\end{aligned}$$

Since A is arbitrary

$$T(|\phi_1^\sim\rangle\langle\phi_1^\sim|) = |(B\phi_1)^\sim\rangle\langle(B\phi_1)^\sim| \qquad (3.13)$$

from which Eq. (3.8) follows immediately. QED

We comment that although T takes pure classical states to pure classical states it takes most pure states to mixed states.

The following theorem gives a typical example of the type of time evolution used in quantum optics. For simplicity we assume that there is only one mode, so $\mathscr{H} = \mathbb{C}$. If f is a probability density on $\mathbb{C}$ we denote by $\rho(f)$ the corresponding classical state.

Theorem 3.3 *If $\omega \in \mathbb{R}$ and $a, b \in \mathbb{R}^+$ there exists a unique dynamical semigroup T_t on V such that for every classical state $\rho(f)$*

$$T_t\{\rho(f)\} = \rho(f_t) \qquad (3.14)$$

where

$$f_t(z) = \int_{\mathbb{C}} e^{at} f\{e^{(a-i\omega)t}(z-w)\}(2\pi\alpha_t)^{-1}. \\ \exp\{-|w|^2/2\alpha_t\}du\,dv \qquad (3.15)$$

and $w = u + iv$, $z = x + iy$ while

$$\alpha_t = ba^{-1}\{1 - e^{-2at}\}. \qquad (3.16)$$

Moreover f_t is the solution of the Fokker–Planck equation

$$\frac{\partial f}{\partial t} - a\left(x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + 2f\right) + \omega\left(y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}\right) - b\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right) = 0 \quad (3.17)$$

so a, ω, b may be called the dissipation rate, the rotation rate and the fluctuation rate respectively of the dynamical semigroup.

Proof The proof that f_t satisfies Eq. (3.17) is most easily done by taking Fourier transforms, followed by an explicit computation. The map $f \to f_t$ on the probability densities is the result of applying two maps in succession. The first is induced by the linear contraction $z \to e^{(-a+i\omega)t}z$ on $\mathbb{C}$ for which we use Theorem 3.2, while the second is convolution with the probability density

$$(2\pi\alpha_t)^{-1}\exp\left\{-\frac{|w|^2}{2\alpha_t}\right\} \tag{3.18}$$

for which we use Theorem 3.1. This proves the existence of a positive trace-preserving linear map $T_t: V \to V$ satisfying Eq. (3.14) for all $t \geq 0$. Its uniqueness, strong continuity and the semigroup property all follow from the corresponding results for the classical states by Lemma 2.2. QED

The above theorem associates a certain dynamical semigroup on V with a Markov semigroup on $\mathbb{C}$. Inspection of Eq. (3.17) shows that the relevant Markov process is of the Ornstein–Uhlenbeck type. The correspondence can be used to study the dynamical semigroup.

Theorem 3.4 *For every state* $\rho \in V^{+}$ *with* $\mathrm{tr}[\rho] = 1$ *we have*

$$\lim_{t\to+\infty} T_t(\rho) = \rho_\beta \tag{3.19}$$

in trace norm, where ρ_β *is the Gibbs state associated to the inverse temperature* β *given by*

$$e^\beta - 1 = \frac{a}{2b}. \tag{3.20}$$

Proof It is an immediate consequence of Eq. (3.15) that for any $f \in L^1(\mathbb{C})$

$$\lim_{t\to-\infty} f_t(z) = \frac{a}{2\pi b}\exp\left\{-\frac{a|z|^2}{2b}\right\} \tag{3.21}$$

the limit being in L^1-norm. This proves Eq. (3.19) for any classical state ρ by Theorem 2.3. Now Lemma 2.2 implies that

$$\lim_{t\to+\infty} T_t(\rho) = \rho_\beta\,\mathrm{tr}[\rho] \tag{3.22}$$

for all $\rho \in V$, which implies Eq. (3.19). QED

8.4 The Weyl algebra*

We have developed the above ideas in the context of mappings on the state space, but it is sometimes useful to know the dual results on the Weyl algebra.

For each $f \in \mathcal{H}$ we have defined a unitary operator $W(f)$ on $\mathcal{F}$ and proved that

$$W(f)W(g) = W(f+g)\exp\{i\,\mathrm{Im}\langle f, g\rangle/2\}. \tag{4.1}$$

The set of finite linear combinations $\sum_{i=1}^{n} \lambda_i W(f_i)$ is a *-algebra in $\mathscr{L}(\mathscr{F})$ and its norm closure is a C^*-algebra which we call the Weyl algebra $\mathscr{A}$. This algebra is actually independent of the particular representation of the CCR's used to construct it, but we do not need this fact.

Theorem 4.1 *For every probability measure μ on $\mathscr{H}$ there is a unique positive linear map $S\colon \mathscr{A} \to \mathscr{A}$ such that*

$$S\{W(f)\} = W(f) \int_{\mathscr{H}} e^{i2^{1/2}\operatorname{Re}\langle f, \phi\rangle} \mu(d\phi) \tag{4.2}$$

for all $f \in \mathscr{H}$.

Proof Uniqueness is immediate from the definition of the Weyl algebra, so we need only prove existence. The map T of Theorem 3.1 has an adjoint map S from $\mathscr{L}_s(\mathscr{F})$ into $\mathscr{L}_s(\mathscr{F})$ which is positive and normal, and this is uniquely extensible to a complex linear map $S\colon \mathscr{L}(\mathscr{F}) \to \mathscr{L}(\mathscr{F})$. If $\psi \in \mathscr{H}$ then

$$\begin{aligned}
\langle S\{W(f)\}\psi^{\sim}, \psi^{\sim}\rangle &= \operatorname{tr}[W(f)T(|\psi^{\sim}\rangle\langle\psi^{\sim}|)] \\
&= \int_{\mathscr{H}} \operatorname{tr}[W(f)|(\phi + \psi)^{\sim}\rangle\langle(\phi + \psi)^{\sim}|]\mu(d\phi) \\
&= \int_{\mathscr{H}} e^{i2^{-1/2}\operatorname{Re}\langle f, \phi + \psi\rangle} \\
&\quad \times \langle(\phi + \psi + i2^{-1/2}f)^{\sim}, (\phi + \psi)^{\sim}\rangle \mu(d\phi) \\
&= \int_{\mathscr{H}} \exp\{i2^{-1/2} \operatorname{Re}\langle f, \phi + \psi\rangle \\
&\quad + \langle \phi + \psi + i2^{-1/2}f, \phi + \psi\rangle \\
&= -\tfrac{1}{2}\| \phi + \psi + i2^{-1/2}f\|^2 - \tfrac{1}{2}\|\phi + \psi\|^2\}\mu(d\phi) \\
&= \int_{\mathscr{H}} \exp\{i2^{-1/2}2 \operatorname{Re}\langle f, \phi + \psi\rangle - \tfrac{1}{4}\|f\|^2\}\mu(d\phi) \\
&= \langle W(f)\psi^{\sim}, \psi^{\sim}\rangle \int_{\mathscr{H}} e^{i2^{1/2}\operatorname{Re}\langle f, \phi\rangle}\mu(d\phi).
\end{aligned} \tag{4.3}$$

If B denotes the RHS of Eq. (4.2) we have shown that

$$\operatorname{tr}[S\{W(f)\}\rho] = \operatorname{tr}[B\rho] \tag{4.4}$$

whenever $\rho = |\psi^{\sim}\rangle\langle\psi^{\sim}|$. The same now holds for all $\rho \in V$ by Lemma 2.2 and hence for all $\rho \in \mathscr{T}(\mathscr{F})$. Therefore $S\{W(f)\} = B$. Eq. (4.2) and the norm continuity of S imply that S maps $\mathscr{A}$ into $\mathscr{A}$. QED

Theorem 4.2 *For every linear contraction B on $\mathscr{H}$ there is a unique positive linear map $S: \mathscr{A} \to \mathscr{A}$ such that*

$$S\{W(f)\} = W(B^*f)\exp\tfrac{1}{4}\{\|B^*f\|^2 - \|f\|^2\} \tag{4.5}$$

for all $f \in \mathscr{H}$.

Proof If $S: \mathscr{L}(\mathscr{F}) \to \mathscr{L}(\mathscr{F})$ is the adjoint of the map T of Theorem 3.2 then S is positive and normal, and if $\psi \in \mathscr{H}$ then

$$\begin{aligned}
\langle S\{W(f)\}\psi^\sim, \psi^\sim\rangle &= \mathrm{tr}[W(f)T(|\psi^\sim\rangle\langle\psi^\sim|)] \\
&= \langle W(f)(B\psi)^\sim, (B\psi)^\sim\rangle \\
&= e^{i2^{-1/2}\mathrm{Re}\langle B\psi, f\rangle}\langle(B\psi + i2^{-1/2}f)^\sim, (B\psi)^\sim\rangle \\
&= \exp\{i2^{-1/2}\,\mathrm{Re}\langle B\psi, f\rangle + \langle B\psi + i2^{-1/2}f, B\psi\rangle \\
&\quad - \tfrac{1}{2}\|B\psi + i2^{-1/2}f\|^2 - \tfrac{1}{2}\|B\psi\|^2\} \\
&= \exp\{i2^{1/2}\,\mathrm{Re}\langle B\psi, f\rangle - \tfrac{1}{4}\|f\|^2\} \\
&= \exp\{i2^{1/2}\,\mathrm{Re}\langle \psi, B^*f\rangle - \tfrac{1}{4}\|f\|^2\} \\
&= \langle W(B^*f)\psi^\sim, \psi^\sim\rangle\exp\tfrac{1}{4}\{\|B^*f\|^2 - \|f\|^2\}.
\end{aligned} \tag{4.6}$$

The proof is completed as in Theorem 4.1. QED

8.5 Generalised coherent states

The theory of Glauber coherent states is closely related to the Weyl group. In recent years other classes of coherent states, for example those of Bloch, have been used in the study of large quantum-mechanical systems. These generalised coherent states are usually related to a group other than the Weyl group, $SU(2)$ in the case of the Bloch coherent states. We now study some of the abstract ideas behind these developments.

Let X be a locally compact second countable Hausdorff space provided with a Borel measure dx. If $\mathscr{H}$ is a Hilbert space we define a continuous resolution of the identity to be a norm continuous family of vectors $\psi_x \in \mathscr{H}$ parametrised by $x \in X$ and such that

$$\int_X |\psi_x\rangle\langle\psi_x|\,dx = 1 \tag{5.1}$$

where the integral converges in the strong operator topology.

Theorem 5.1 *If $\{\psi_x\}$ is a continuous resolution of the identity on $\mathscr{H}$ the formula*

$$(J\psi)(x) = \langle\psi, \psi_x\rangle \tag{5.2}$$

defines an isometric embedding of $\mathscr{H}$ onto a closed subspace $J(\mathscr{H})$ of $L^2(X)$ consisting entirely of continuous functions. The projection P of $L^2(X)$ onto $J(\mathscr{H})$ has the continuous integral kernel

$$K(x, y) = \langle \psi_y, \psi_x \rangle. \tag{5.3}$$

If X is compact then $\mathscr{H}$ must be finite-dimensional.

Proof If $\psi \in \mathscr{H}$ then

$$\int_X |\langle \psi, \psi_x \rangle|^2 \, dx = \|\psi\|^2 \tag{5.4}$$

so J is indeed isometric. It is clearly also linear. $x \to (J\psi)(x)$ is a continuous function on X because $x \to \psi_x$ is norm continuous. If $\psi \in \mathscr{H}$ then

$$\begin{aligned}\int_X K(x, y)(J\psi)(y)dy &= \int_X \langle \psi_y, \psi_x \rangle \langle \psi, \psi_y \rangle dy \\ &= \langle \psi, \psi_x \rangle = (J\psi)(x)\end{aligned} \tag{5.5}$$

Therefore

$$\int_X K(x, y)\phi(y)dy = \phi(x) \tag{5.6}$$

for all $\phi \in J(\mathscr{H})$. If $\phi \in J(\mathscr{H})^\perp$ then

$$\int_X \phi(y) \langle \psi_y, \psi \rangle dy = 0 \tag{5.7}$$

for all $\psi \in \mathscr{H}$, so

$$\int_X K(x, y)\phi(y)dy = 0. \tag{5.8}$$

Equations (5.6) and (5.8) imply that $K(x, y)$ is indeed the kernel of the projection P.

The kernel K is a continuous function because $x \to \psi_x$ is norm continuous. If X is compact then K is a Hilbert–Schmidt kernel. But a projection P can only be a Hilbert–Schmidt operator if it is of finite rank. Therefore $\mathscr{H}$ is finite-dimensional. QED

Since P is a projection K satisfies the equation

$$\int_X K(x, y)K(y, z)dy = K(x, z) \tag{5.9}$$

by virtue of which it is called a reproducing kernel.

In chapter 3 section 4 we construct a POV measure on phase space which could be considered as a simultaneous observable for position and momentum. We do this again in a more general context and show how the POV measure is related to a projection-valued measure on a larger Hilbert space. This is a particular case of Naimark's theorem, which we shall present in chapter 9 section 3.

Theorem 5.2 *The formula*

$$A(E) = \int_E |\psi_x\rangle\langle\psi_x| dx \tag{5.10}$$

defines an observable on X with values in $\mathscr{L}(\mathscr{H})$. If J is the isometric embedding of $\mathscr{H}$ into $L^2(X)$ and the projection-valued measure $P(\cdot)$ on $L^2(X)$ is defined by

$$\{P(E)\psi\}(x) = \chi_E(x)\psi(x) \tag{5.11}$$

for all Borel sets $E \subseteq X$ and all $\psi \in L^2(X)$ then

$$A(E) = J^*P(E)J \tag{5.12}$$

for all Borel sets $E \subseteq X$.

Proof If $\phi, \psi \in \mathscr{H}$ then

$$\begin{aligned}\langle J^*P(E)J\phi, \psi\rangle &= \int_X \chi_E(x)(J\phi)(x)\overline{(J\psi)(x)}dx \\ &= \int_E \langle\phi, \psi_x\rangle\langle\psi_x, \psi\rangle dx \\ &= \langle A(E)\phi, \psi\rangle. \end{aligned} \tag{5.13}$$

Since ϕ and ψ are arbitrary Eq. (5.12) follows. QED

We now divert our attention to an apparently unrelated subject. Let G be a locally compact second countable group and U_g a representation of G on a Hilbert space $\mathscr{H}$. Let $\psi \in \mathscr{H}$ be any unit vector and define the generalised coherent states $\psi_g \in \mathscr{H}$ by

$$\psi_g = U_g\psi. \tag{5.14}$$

Then for any probability measure μ on G define the classical state $\rho(\mu)$ as

$$\rho(\mu) = \int_G |\psi_g\rangle\langle\psi_g| \mu(dg). \tag{5.15}$$

We now generalise the ideas of section 3. A random walk on G is defined as a one-parameter family $\{\mu_t\}_{t\geq 0}$ of probability measures on G such that

(*i*) $\mu_0 = \delta_e$ where δ_e is the Dirac measure at the identity e of G; (5.16) (*ii*) $\mu_{s+t} = \mu_s \circ \mu_t$ for all s, $t \geqslant 0$ where $\circ$ denotes the convolution of measures; (5.17)

(*iii*) if W is any open set containing e then

$$\lim_{t\to 0} \mu_t(W) = 1. \tag{5.18}$$

The dynamical semigroup T_t of the following theorem is said to be subordinated to the random walk μ_t.

Theorem 5.3 *Given a random walk $\{\mu_t\}$ on G and a representation U of G on a Hilbert space $\mathscr{H}$ there exists a dynamical semigroup T_t on the state space V of $\mathscr{H}$ such that*

$$\rho(\mu_t \circ \mu) = T_t\{\rho(\mu)\} \tag{5.19}$$

for all $t \geqslant 0$ and all probability measures μ on G.

Proof If we define $T_t: V \to V$ by

$$T_t(\rho) = \int_G U_g \rho U_g^* \mu_t(dg) \tag{5.20}$$

then Eqs (5.17) and (5.18) imply that T_t is a dynamical semigroup. Moreover

$$\begin{aligned} T_t\{\rho(\mu)\} &= \int_G \int_G U_g U_h |\psi\rangle\langle\psi| U_h^* U_g^* \mu(dh)\mu_t(dg) \\ &= \int_G U_g |\psi\rangle\langle\psi| U_g^* (\mu_t \circ \mu)(dg) \\ &= \rho(\mu_t \circ \mu). \qquad \text{QED} \end{aligned} \tag{5.21}$$

A very simple example of the above theorem is obtained by taking μ_t to be the Gaussian random walk on the real line. If H is a bounded self-adjoint operator on $\mathscr{H}$ then the formula

$$T_t(\rho) = (2\pi t)^{-1/2} \int_{-\infty}^{\infty} e^{-s^2/2t} e^{-iHs} \rho\, e^{iHs}\, ds \tag{5.22}$$

defines a dynamical semigroup T_t whose infinitesimal generator Z is given by

$$Z(\rho) = -\tfrac{1}{2}[H, (H, \rho)] \tag{5.23}$$

as a straightforward expansion of e^{iHs} shows.

We mention that Theorem 5.3 can be extended to multiplier representations of groups, that is maps $U: G \to \mathscr{L}(\mathscr{H})$ such that

(*i*) $U_g U_g^* = U_g^* U_g = 1$ for all $g \in G$; (5.24)

(*ii*) $U_g U_h = \lambda(g, h) U_{gh}$ for all $g, h \in G$ (5.25) where $\lambda(g, h)$ is a scalar of absolute value one. Theorem 3.1 may be obtained from Theorem 5.3 by taking G to be the Weyl group, or by taking G to be the additive group $\mathbb{R}^2$ with a multiplier representation. The latter course is computationally rather simpler.

Many of the continuous resolutions of the identity in quantum theory arise in the following way. If G is a locally compact second countable unimodular group with Haar measure dg an irreducible representation U of G on a Hilbert space $\mathscr{H}$ is called square-integrable if there is a constant d, called the formal dimension of the representation, such that

$$\int_G |\langle U_g\psi, \phi\rangle|^2 \, dg = d^{-1}\|\phi\|^2\|\psi\|^2 \tag{5.26}$$

for all $\phi, \psi \in \mathscr{H}$. In such a situation if $\psi \in \mathscr{H}$ and $\|\psi\| = d^{1/2}$ and if we define $\psi_g = U_g\psi$ then $\{\psi_g\}$ is a continuous resolution of the identity in $\mathscr{H}$. For compact groups the situation is particularly simple.

Theorem 5.4 *Every irreducible representation U of a compact metrisable group G has formal dimension equal to its actual dimension, if the Haar measure is normalised to have mass one.*

Proof If we define

$$A = \int_G |U_g\psi\rangle\langle U_g\psi| \, dg \tag{5.27}$$

then by the invariance of the Haar measure

$$U_h A = A U_h \tag{5.28}$$

for all $h \in G$, so $A = \lambda 1$ for some constant λ. If d is the dimension of the representation then

$$\begin{aligned} \lambda d &= \mathrm{tr}[A] \\ &= \int_G \mathrm{tr}[|U_g\psi\rangle\langle U_g\psi|] dg \\ &= \|\psi\|^2. \end{aligned} \tag{5.29}$$

Therefore

$$\begin{aligned} \int_G |\langle U_g\psi, \phi\rangle|^2 \, dg &= \langle A\phi, \phi\rangle \\ &= d^{-1}\|\phi\|^2\|\psi\|^2. \qquad \text{QED} \end{aligned} \tag{5.30}$$

In the above construction of a continuous resolution of the identity, the choice of the vector ψ was arbitrary. In physical problems it is often sensible to choose ψ to be the ground state of some Hamiltonian. It then frequently happens that some of the group parameters are not relevant to the value of ψ_g. If $U_h\psi = \psi$ for all h in the subgroup H of G then $\psi_g = \psi_{gh}$ for all $h \in H$ so the vectors ψ_g may be parametrised by the points of the coset space G/H. This is the explanation of the fact that the Bloch coherent states are generally parametrised by the points of the unit sphere rather than the points of $SU(2)$.

Notes

1 Coherent states were introduced into quantum theory by Schrodinger but started being used widely only in the 1960s (Bargmann, 1961; Glauber, 1963; Klauder and Sudarshan, 1968; Segal, 1962). Lemma 1.1 is due to von Neumann and is a special case of a more general completeness theorem for the coherent states (Bargmann *et al.*, 1971). Our account of Fock space was inspired by that of Klauder (1970).

2 The classical states are those which in the language of Glauber (1963) have a P-representation for a non-negative P. Lemma 2.2 was proved by Klauder (1966) and Rocca (1966). Theorem 2.3 was proved by Glauber (1963) although the idea is also present in Wigner (1932). Theorem 2.4 is a rephrasing of standard facts about the harmonic oscillator.

3 The relationship between dynamical semigroups and Markov semigroups can be derived for the harmonic oscillator in many ways. McKenna and Frisch (1966) studied it using the Husimi transform (Husimi, 1940) while Ford *et al.* (1965) derived a Langevin equation from the consideration of the dynamics of an infinite chain of harmonic oscillators, a method refined by Davies (1972a, 1973) and Lewis and Thomas (1974, 1975).

4 The Weyl algebra started being widely used in quantum field theory after the publication of papers of Segal (1961, 1962) and Araki (1960). There are in fact several slightly different Weyl algebras, ours being that of Manuceau (1968) and Slawny (1972). Theorem 4.2 may be found in Davies (1972a) and Nelson (1973), while the analogous results for the CARs are in (Schrader and Uhlenbrock, 1975). The proof that the Weyl algebra is independent of the representation may be found in Slawny (1972).

5 The theory of reproducing kernel Hilbert spaces was unified by Aronsajn (1950). Generalised coherent states were introduced into mathematical physics in (Arrechi *et al.*, 1972; Aslaken and Klauder, 1969; Barut and Girardello (1971); Lu, 1971; Perelomov, 1972, 1975; Radcliffe, 1971), our account being close to that of Perelomov (1972). The general theory of square-integrable representations may be found in Dixmier (1969a). The theory of subordinated dynamical semigroups was developed by Davies (1972b) in the abelian case, and in greater generality by Kossakowski (1972a). A complete classification of all random walks on a general Lie group was given by Hunt (1956). Applications of generalised coherent states may be found in Hepp and Lieb (1973a) and Lieb (1973).

9

Hamiltonian Formalism

9.1 Fundamental considerations

Since almost all known laws of physics are invariant under time reversal and time translation, Theorem 2.3.4 suggests that any closed (i.e. isolated) system will possess a Hamiltonian H such that the state ρ_t at time t is computed from the state ρ at time 0 by the formula

$$\rho_t = e^{-itH}\rho\, e^{itH}. \tag{1.1}$$

The question arises whether a dynamical semigroup on the state space of a Hilbert space can in principle be related to a Hamiltonian on a larger Hilbert space, the larger Hilbert space arising from the quantisation of the part of the external world which affects the evolution of the open system through the phenomenological equations.

For technical reasons we prefer to consider this problem in the Heisenberg formulation. We shall have to use some results from the theory of C^*-algebras, although these will be applied only to the algebra $\mathscr{L}(\mathscr{H})$. If $\mathscr{H}$ is the space of the system then by a larger Hilbert space we mean the Hilbert space $\mathscr{H} \otimes \mathscr{K}$, where $\mathscr{K}$ is a Hilbert space representing the external world. It is easily seen that the formula $A \to A \otimes 1$ defines a positive normal linear map of $\mathscr{L}(\mathscr{H})$ into $\mathscr{L}(\mathscr{H} \otimes \mathscr{K})$. If we are given the state of the external world then we can define a map going in the opposite direction.

Lemma 1.1 *If ρ is a state on $\mathscr{K}$ then the formula*

$$\operatorname{tr}[E_\rho(A)\sigma] = \operatorname{tr}[A(\rho \otimes \sigma)] \tag{1.2}$$

where $A \in \mathscr{L}(\mathscr{H} \otimes \mathscr{K})$ and $\sigma \in \mathscr{T}(\mathscr{H})$, defines a normal positive linear map $E_\rho\colon \mathscr{L}(\mathscr{H} \otimes \mathscr{K}) \to \mathscr{L}(\mathscr{H})$ such that

$$E_\rho(B \otimes 1) = B \tag{1.3}$$

for all $B \in \mathscr{L}(\mathscr{H})$.

Proof The map $\sigma \to \sigma \otimes \rho$ is a positive linear map from $\mathscr{T}(\mathscr{H})$ to $\mathscr{T}(\mathscr{H} \otimes \mathscr{K})$, and its adjoint E_ρ has the stated properties by Lemma 2.2.2

QED

If one is given the initial state ρ of the external world and a Hamiltonian H on $\mathscr{H} \otimes \mathscr{K}$ which describes the time evolution of the world and system, including their interaction with each other, then the time evolution of the system alone is determined by the family of positive linear maps T_t on $\mathscr{T}(\mathscr{H})$ defined for $t \geqslant 0$ by

$$T_t(A) = E_\rho\{e^{iHt}(A \otimes 1)e^{-iHt}\}. \tag{1.4}$$

The family $\{T_t\}_{t \geqslant 0}$ will not generally be a semigroup.

In the converse direction given a family $\{T_t\}_{t \geqslant 0}$ of positive linear maps on $\mathscr{L}(\mathscr{H})$ one can ask under what conditions a representation of the type of Eq. (1.4) exists. It turns out that a new condition, of complete positivity, is relevant.

9.2 Complete positivity*

If $\mathscr{A}$ and $\mathscr{B}$ are C^*-algebras, the algebra $\mathscr{A}_n \simeq \mathscr{A} \otimes \mathscr{L}(\mathbb{C}^n)$ of $n \times n$ matrices with entries in $\mathscr{A}$ is also a C^* algebra in a natural way. Given a linear map $T\colon \mathscr{A} \to \mathscr{B}$ we can define a linear map $T_n\colon \mathscr{A}_n \to \mathscr{B}_n$ by acting with T on each of the coefficients of an element of $\mathscr{A}_n$. We say that T is completely positive if T_n is positive for all $n \geqslant 1$.

It has been argued that only the completely positive linear maps can have any physical significance, for the following reason. If $\mathscr{H}$ is the Hilbert space of a system localised in some box and there exists a particle with n degrees of freedom so far away from the box that there is no interaction between the two then the Hilbert space for the system plus particle is $\mathscr{H} \otimes \mathbb{C}^n$. An operation on the system which does not affect the distant particle is described by a positive linear map T_n on $\mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ such that

$$T_n(A \otimes B) = T(A) \otimes B \tag{2.1}$$

for some positive linear map T on $\mathscr{L}(\mathscr{H})$ and all $A \in \mathscr{L}(\mathscr{H})$ and $B \in \mathscr{L}(\mathbb{C}^n)$. But $\mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ is isomorphic to the $n \times n$ matrix algebra over $\mathscr{L}(\mathscr{H})$ and T_n is the map described above. Therefore T_n is positive, and T is completely positive.

Whether or not one finds this argument convincing, it is certainly the case that completely positive maps are of great importance. A mathematical, rather than operational, justification for their introduction may be based upon Theorem 2.3. We remark that the condition of complete positivity could have been imposed on operations, instruments and all subsequent results in this volume. We have refrained from doing so because the slight but wearisome extra details necessitated would have been completely unrewarded until this chapter.

Theorem 2.1 *If $\mathscr{A}$ is a C*-algebra with identity and $T\colon \mathscr{A} \to \mathscr{L}(\mathscr{H})$ a linear map, then T is completely positive if and only if it has the form*

$$T(A) = V^*\rho(A)V \tag{2.2}$$

for some representation ρ of $\mathscr{A}$ on a Hilbert space $\mathscr{K}$ and some bounded linear map $V\colon \mathscr{H} \to \mathscr{K}$. If $\mathscr{A}$ and $\mathscr{H}$ are separable then $\mathscr{K}$ can be taken to be separable. If $\mathscr{A}$ is a von Neumann algebra and T is normal then ρ can be taken to be normal.

Proof Suppose that such a representation of T exists. If $A \in \mathscr{A}_n^+$ then there exists $B \in \mathscr{A}_n$ with $A = B^*B$. If $x_1, \ldots, x_n \in \mathscr{H}$ and $x = (x_1, \ldots, x_n)$ then

$$\begin{aligned}
\langle T_n(A)x, x\rangle &= \langle T_n(B^*B)x, x\rangle \\
&= \sum_{i,j,k} \langle T\{(B_{ji})^*B_{jk}\}x_k, x_i\rangle \\
&= \sum_{i,j,k} \langle \rho\{(B_{ji})^*B_{jk}\}Vx_k, Vx_i\rangle \\
&= \sum_{i,j,k} \langle \rho(B_{jk})Vx_k, \rho(B_{ji})Vx_i\rangle \\
&= \sum_j \left\| \sum_k \rho(B_{jk})Vx_k \right\|^2 \\
&\geqslant 0
\end{aligned} \tag{2.3}$$

so $T_n(A) \geqslant 0$ and T_n is positive.

Conversely suppose T is completely positive and let $\mathscr{A} \odot \mathscr{H}$ denote the algebraic tensor product of $\mathscr{A}$ and $\mathscr{H}$. If

$$\phi = \sum_{i=1}^n A_i \otimes x_i, \qquad \psi = \sum_{i=1}^n B_i \otimes y_i \tag{2.4}$$

lie in $\mathscr{A} \odot \mathscr{H}$ let x and y be the sequences

$$x = (x_1, \ldots, x_n), \qquad y = (y_1, \ldots, y_n) \tag{2.5}$$

and let $A', B' \in \mathscr{A}_n$ be the $n \times n$ matrices

$$(A')_{ij} = \delta_{1j}A_i, \qquad (B')_{ij} = \delta_{1j}B_i. \tag{2.6}$$

If we define the sesquilinear form $\langle \cdot, \cdot \rangle$ on $\mathscr{A} \odot \mathscr{H}$ by

$$\langle \phi, \psi\rangle = \langle T_n(B'^*A')x, y\rangle \tag{2.7}$$

then

$$\begin{aligned}
\langle \phi, \phi\rangle &= \langle T_n(A'^*A')x, x\rangle \\
&\geqslant 0
\end{aligned} \tag{2.8}$$

by the positivity of T_n, so $\langle \cdot, \cdot \rangle$ is an inner product except that the subspace

$$\mathcal{N} = \{\phi \in \mathcal{A} \odot \mathcal{H} : \langle \phi, \phi \rangle = 0\} \tag{2.9}$$

need not be zero. We define $\mathcal{K}$ to be the completion of the quotient space $(\mathcal{A} \odot \mathcal{H})/\mathcal{N}$ for the induced inner product.

If $X \in \mathcal{A}$ we define $X'' \in \mathcal{A}_n$ by

$$(X'')_{ij} = \delta_{ij} X \tag{2.10}$$

and $\xi \in \mathcal{A} \odot \mathcal{H}$ by

$$\xi = \sum_{i=1}^{n} (XA_i) \otimes x_i. \tag{2.11}$$

Then

$$\|\xi\|^2 = \langle T_n(A'^* X''^* X'' A')x, x \rangle. \tag{2.12}$$

Since

$$0 \leqslant A'^* X''^* X'' A' \leqslant \|X''\|^2 A'^* A' \tag{2.13}$$

as elements of $\mathcal{A}_n$, and T_n is positive, we have

$$\begin{aligned} \|\xi\|^2 &\leqslant \|X''\|^2 \langle T_n(A'^* A')x, x \rangle \\ &= \|X\|^2 \|\phi\|^2. \end{aligned} \tag{2.14}$$

Therefore the map

$$\sum_{i=1}^{n} A_i \otimes x_i \to \sum_{i=1}^{n} (XA_i) \otimes x_i \tag{2.15}$$

induces a bounded linear map $\rho(X)$ on $\mathcal{K}$. It is immediate that ρ is a representation of $\mathcal{A}$ on $\mathcal{K}$. Since

$$\begin{aligned} \|1 \otimes x\|^2 &= \langle T(1)x, x \rangle \\ &\leqslant \|T\| \, \|x\|^2 \end{aligned} \tag{2.16}$$

for all $x \in \mathcal{H}$, the map $x \to 1 \otimes x$ induces a bounded linear map $V: \mathcal{H} \to \mathcal{K}$. If $x \in \mathcal{H}$ and $X \in \mathcal{A}$ then

$$\begin{aligned} \langle \rho(X)Vx, Vx \rangle &= \langle X \otimes x, 1 \otimes x \rangle \\ &= \langle T(X)x, x \rangle. \end{aligned} \tag{2.17}$$

Since x is arbitrary it follows that

$$T(X) = V^* \rho(X) V. \tag{2.18}$$

If $\{X_n\}$ is a countable dense set in $\mathcal{A}$ and $\{\psi_m\}$ is a countable dense set in $\mathcal{H}$, then $\mathcal{K}$ is generated by the countable set $\{X_n \otimes \psi_m\}$ and so is separable.

If $\mathscr{A}$ is a von Neumann algebra and T is normal let X_r be a monotonically increasing generalised sequence in $\mathscr{A}$ converging to $X \in \mathscr{A}$ and let $A, B \in \mathscr{A}$. Then

$$\begin{aligned} T(A^*X_rB) = \; & \frac{1}{4}T\{(A+B)^*X_r(A+B)\} \\ & -\frac{1}{4}T\{(A-B)^*X_r(A-B)\} \\ & -\frac{i}{4}T\{(A+iB)^*X_r(A+iB)\} \\ & +\frac{i}{4}T\{(A-iB)^*X_r(A-iB)\}. \end{aligned} \tag{2.19}$$

Since

$$(A+i^sB)^*X_r(A+i^sB) \uparrow (A+i^sB)^*X(A+i^sB) \tag{2.20}$$

for all integers s we conclude that

$$\lim_{r\to\infty} T(A^*X_rB) = T(A^*XB) \tag{2.21}$$

in the weak operator topology.

If ϕ is defined by Eq. (2.4) and $X_r \uparrow X$ then

$$\begin{aligned} \lim_{r\to\infty} \langle \rho(X_r)\phi, \phi\rangle &= \lim_{r\to\infty} \sum_{i,j=1}^{n} \langle (X_rA_i) \otimes x_i, A_j \otimes x_j\rangle \\ &= \lim_{r\to\infty} \sum_{i,j=1}^{n} \langle T(A_j^*X_rA_i)x_i, x_j\rangle \\ &= \sum_{i,j=1}^{n} \langle T(A_j^*XA_i)x_i, x_j\rangle \\ &= \langle \rho(X)\phi, \phi\rangle. \end{aligned} \tag{2.22}$$

Since such ϕ are dense in $\mathscr{K}$, $\rho(X_n) \uparrow \rho(X)$ and ρ is normal. QED

Lemma 2.2 *Let ρ be a normal representation of $\mathscr{L}(\mathscr{H})$ on the Hilbert space $\mathscr{K}$. Then there exists a direct sum decomposition*

$$\mathscr{K} = \sum_n \oplus \mathscr{K}_n \tag{2.23}$$

where the subspaces $\mathscr{K}_n$ are invariant under ρ and the restriction of ρ to each subspace $\mathscr{K}_n$ is unitarily equivalent to the standard representation of $\mathscr{L}(\mathscr{H})$ on $\mathscr{H}$. If $\mathscr{K}$ is separable then the index set is countable.

Proof If e is a unit vector in $\mathscr{H}$ then the projection $P = \rho(|e\rangle\langle e|)$ is non-zero. For if U_n are unitary operators in $\mathscr{L}(\mathscr{H})$ such that $e_n = U_n e$ form a maximal orthonormal set in $\mathscr{H}$ and if $P = 0$ then

$$\begin{aligned}\rho(|e_n\rangle\langle e_n|) &= \rho(U_n)P\rho(U_n^*)\\ &= 0\end{aligned} \tag{2.24}$$

Since ρ is normal

$$\begin{aligned}\rho(1) &= \sum_n \rho(|e_n\rangle\langle e_n|)\\ &= 0\end{aligned} \tag{2.25}$$

which contradicts the requirement that $\rho(1) = 1$ for every representation.

If $\psi \in \mathscr{K}$ is a unit vector such that $P\psi = \psi$ then

$$\begin{aligned}\langle \rho(X)\psi, \psi\rangle &= \langle \rho(X)P\psi, P\psi\rangle\\ &= \langle \rho\{|e\rangle\langle e|X|e\rangle\langle e|\}\psi, \psi\rangle\\ &= \langle Xe, e\rangle.\end{aligned} \tag{2.26}$$

By the GNS theorem the standard representation of $\mathscr{L}(\mathscr{H})$ on $\mathscr{H}$ is unitarily equivalent to the representation ρ restricted to the subspace

$$\mathscr{K}_1 = \{\rho(X)\psi : X \in \mathscr{K}\}^- \tag{2.27}$$

of $\mathscr{K}$. Thus we may decompose the representation ρ into the sum of a representation unitarily equivalent to the standard representation and a normal representation on $\mathscr{K}_1^\perp$.

By Zorn's lemma there is a maximal class of orthogonal subspaces $\mathscr{K}_n$ of $\mathscr{K}$ with unitary maps $U_n : \mathscr{H} \to \mathscr{K}_n$ such that $\mathscr{K}_n$ are invariant with respect to ρ and

$$X = U_n^* \rho(X) U_n \tag{2.28}$$

for all $X \in \mathscr{L}(\mathscr{H})$. Applying the argument of the last paragraph to the representation ρ restricted to $(\bigcup_n \mathscr{K}_n)^\perp$, we find by maximality that

$$\mathscr{K} = \sum_n \oplus \mathscr{K}_n. \qquad \text{QED} \tag{2.29}$$

Theorem 2.3 *A normal positive linear map T on $\mathscr{L}(\mathscr{H})$ is completely positive if and only if there exist bounded operators A_n on $\mathscr{H}$ such that*

$$T(X) = \sum_{n \in N} A_n^* X A_n \tag{2.30}$$

for all $X \in \mathscr{L}(\mathscr{H})$. If $\mathscr{H}$ is separable then the indexing set N may be taken to be countable.

Proof Continuing with the notation of the last proof, and assuming that T is completely positive, the projections P_n of $\mathscr{K}$ onto $\mathscr{K}_n$ commute with the representation ρ so

$$\begin{aligned} T(X) &= \sum_n V^* P_n \rho(X) P_n V \\ &= \sum_n V^* P_n U_n X U_n^* P_n V \\ &= \sum_n A_n^* X A_n \end{aligned} \tag{2.31}$$

if $A_n = U_n^* P_n V$. The converse, that every map T of the form of Eq. (2.30) is completely positive, is trivial.

If $\{\phi_m\}$ is a countable dense set in $\mathscr{H}$ then

$$\infty > \langle T(|\phi_m\rangle\langle\phi_m|)\phi_r, \phi_r\rangle = \sum_n |\langle A_n \phi_r, \phi_m\rangle|^2 \tag{2.32}$$

so

$$N_{r,m} = \{n : \langle A_n \phi_r, \phi_m\rangle \neq 0\} \tag{2.33}$$

is countable. Moreover

$$\begin{aligned} N &\equiv \bigcup_{r,m} N_{r,m} \\ &= \{n : A_n \neq 0\} \end{aligned} \tag{2.34}$$

so N is countable. QED

We finally comment that this representation of T is by no means unique. The operators A_n must clearly satisfy

$$0 \leqslant \sum_n A_n^* A_n \leqslant c1 \tag{2.35}$$

in order to define a completely positive linear map T by Eq. (2.30), but they are otherwise unrestricted.

9.3 Projection-valued measures

In chapter 3 section 1 we presented arguments for defining an observable as a normalised POV measure. From this point of view a projection-valued measure is just a special kind of observable. We show here that it is possible to reverse one's viewpoint and regard an observable as the restriction of a projection-valued measure on a larger Hilbert space. A particular case of this was obtained in Theorem 8.5.2.

Lemma 3.1 *If $\mathscr{A}$ is a separable commutative C^*-algebra with identity then every positive linear map $T: \mathscr{A} \to \mathscr{L}(\mathscr{H})$ is completely positive.*

Proof There exists a compact metrisable space Ω such that $\mathscr{A} \simeq C(\Omega)$ and by Lemma 3.1.2 a POV measure T on Ω which corresponds to the linear map $\bar{T}$. Let $\mathscr{B}_n(\Omega)$ be the C^*-algebra of all bounded Borel-measurable $n \times n$ matrix-valued functions on Ω and define

$$T_n^{\sim} : \mathscr{B}_n(\Omega) \to \mathscr{L}(\mathscr{H})_n \tag{3.1}$$

by

$$(T_n^{\sim} f)_{ij} = \int_\Omega f_{ij}(\omega) T(d\omega). \tag{3.2}$$

Since $\mathscr{B}_n(\Omega) \supseteq \mathscr{A}_n$ it is sufficient to show that the bounded linear map $T_n^{\sim}$ is positive for all n. If $f \in \mathscr{B}_n(\Omega)^+$ then f can be uniformly approximated by expressions of the form

$$g(\omega) = \sum_{r=1}^{R} \chi_{E_r}(\omega) B^{(r)*} B^{(r)} \tag{3.3}$$

where $B^{(r)}$ are $n \times n$ matrices. Now if $\psi = (\psi_1, \ldots, \psi_n)$ where $\psi_i \in \mathscr{H}$ then

$$\begin{aligned} \langle T_n^{\sim}(g)\psi, \psi \rangle &= \sum_{i,j} \langle T^{\sim}(g_{ij})\psi_j, \psi_i \rangle \\ &= \sum_{i,j,k,r} \overline{B_{ki}^{(r)}} B_{kj}^{(r)} \langle T(E_r)\psi_j, \psi_i \rangle \\ &= \sum_{k,r} \left\langle T(E_r)\left(\sum_j B_{kj}^{(r)}\psi_j\right), \left(\sum_i B_{ki}^{(r)}\psi_i\right)\right\rangle \\ &\geqslant 0 \end{aligned} \tag{3.4}$$

so $T_n^{\sim}(g)$ is indeed positive. QED

Theorem 3.2 *Let X be a compact metrisable space and $A(\cdot)$ an observable on the Borel σ-field of X with values in $\mathscr{L}(\mathscr{H})$. Then there exists a projection-valued measure $P(\cdot)$ on a Hilbert space $\mathscr{K} \supseteq \mathscr{H}$ such that if P is the projection of $\mathscr{K}$ onto $\mathscr{H}$ then $A(E)$ is the restriction of $PP(E)P$ to $\mathscr{H}$ for all Borel sets $E \subseteq X$. If $\mathscr{H}$ is separable then $\mathscr{K}$ may be taken to be separable.*

Proof By Theorem 2.1 and Lemma 3.1 there is a representation ρ of $C(\Omega)$ into $\mathscr{L}(\mathscr{K})$ and a bounded linear map $V: \mathscr{H} \to \mathscr{K}$ such that

$$A(f) = V^* \rho(f) V \tag{3.5}$$

for all $f \in C(\Omega)$. Since $A(1) = 1$ we have

$$V^*V = 1 \tag{3.6}$$

so V is an isometry, which enables us to identify $\mathscr{H}$ as a subspace of $\mathscr{K}$. The representation ρ of $C(\Omega)$ corresponds to a projection-valued measure on $\mathscr{K}$ by Lemma 3.1.2. QED

9.4 Finite-dimensional dynamical semigroups

If $\mathscr{H}$ is a finite-dimensional Hilbert space we define a completely positive normalised (CPN) semigroup to be a continuous one-parameter semigroup T_t of completely positive linear maps on $\mathscr{L}(\mathscr{H})$ such that $T_t(1) = 1$ for all $t \geqslant 0$. Such a semigroup induces a CPN semigroup $T_{n,t}$ on $\mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ such that

$$T_{n,t}(A \otimes B) = T_t(A) \otimes B \tag{4.1}$$

for all $A \in \mathscr{L}(\mathscr{H})$ and $B \in \mathscr{L}(\mathbb{C}^n)$. The infinitesimal generator Z_n of $T_{n,t}$ satisfies an equation similar to Eq. (4.1).

Lemma 4.1 *If $\mathscr{H}$ is finite-dimensional and T_t is a CPN semigroup on $\mathscr{L}(\mathscr{H})$ then*

$$Z_n(X^*X) \geqslant Z_n(X^*)X + X^*Z_n(X) \tag{4.2}$$

for all $X \in \mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ and all $n \geqslant 1$.

Proof If V_t and ρ_t are related to T_t as in Theorem 2.1, then $V_t^*V_t = 1$ so $0 \leqslant V_tV_t^* \leqslant 1$.

Therefore if $X \in \mathscr{L}(\mathscr{H})$

$$T_t(X^*)T_t(X) = V_t^*\rho_t(X)^*V_tV_t^*\rho_t(X)V_t \leqslant V_t^*\rho_t(X)^*\rho_t(X)V_t = T_t(X^*X). \tag{4.3}$$

Therefore

$$t^{-1}\{T_t(X^*X) - X^*X\} \geqslant t^{-1}\{T_t(X^*) - X^*\}T_t(X) + t^{-1}X^*\{T_t(X) - X\} \tag{4.4}$$

and letting $t \to 0$

$$Z(X^*X) \geqslant Z(X^*)X + X^*Z(X). \tag{4.5}$$

The same argument applied to the semigroup $T_{n,t}$ yields Eq. (4.2). QED

Theorem 4.2 *If $\mathscr{H}$ is finite-dimensional and T_t is a CPN semigroup on $\mathscr{L}(\mathscr{H})$ then there exists a bounded operator K on $\mathscr{H}$ and a completely positive linear map $\mathscr{J}$ on $\mathscr{L}(\mathscr{H})$ such that*

$$Z(X) = KX + XK^* + \mathscr{J}(X) \tag{4.6}$$

for all $X \in \mathscr{L}(\mathscr{H})$. Moreover

$$\mathscr{J}(1) + K + K^* = 0. \tag{4.7}$$

Proof Let dU be the Haar measure normalised to have total mass unity on the compact group G of all unitary operators on $\mathscr{H}$, and define

$$K = \int_G Z(U^*)U \, dU. \tag{4.8}$$

Then it is clear that $\|K\| \leqslant \|Z\|$. If V is a unitary operator on $\mathscr{H}$ and W is a unitary operator on $\mathbb{C}^n$ then

$$(K \otimes 1)(V \otimes W) = \left(\int_G Z(U^*)UV\,dU\right) \otimes W = \left(\int_G Z(VU^*)U\,dU\right) \otimes W$$

$$= \int_G Z_n\{(V \otimes W)(U^* \otimes 1)\}(U \otimes 1)dU. \tag{4.9}$$

Since the linear span of such $V \otimes W$ equals $\mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ we find that

$$(K \otimes 1)B = \int_G Z_n\{B(U^* \otimes 1)\}(U \otimes 1)dU \tag{4.10}$$

for all $B \in \mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$. Upon taking adjoints we obtain

$$B(K^* \otimes 1) = \int_G (U^* \otimes 1)Z_n\{(U \otimes 1)B\}dU \tag{4.11}$$

for all $B \in \mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$. If we define $\mathscr{J}$ on $\mathscr{L}(\mathscr{H})$ by

$$\mathscr{J}(X) = Z(X) - KX - XK^* \tag{4.12}$$

and $\mathscr{J}_n$ is the extension to $\mathscr{L}(\mathscr{H} \otimes \mathbb{C}^n)$ defined by

$$\mathscr{J}_n(X) = Z_n(X) - (K \otimes 1)X - X(K^* \otimes 1) \tag{4.13}$$

then

$$\begin{aligned}
\mathscr{J}_n(X^*X) = Z_n(X^*X) &- \int_G Z_n\{X^*(U^* \otimes 1)\}(U \otimes 1)X\,dU \\
&- \int_G X^*(U^* \otimes 1)Z_n\{(U \otimes 1)X\}\,dU \\
= &\int_G [Z_n\{X^*(U^* \otimes 1)(U \otimes 1)X\} - Z_n\{X^*(U^* \otimes 1)\}(U \otimes 1)X \\
&- X^*(U^* \otimes 1)Z_n\{(U \otimes 1)X\}]dU \\
\geqslant 0 &
\end{aligned} \tag{4.14}$$

by Eq. (4.2). Therefore $\mathscr{J}_n$ is positive so $\mathscr{J}$ is completely positive. Equation (4.7) is a consequence of the fact that $Z(1) = 0$. QED

The representation of Z in Eq. (4.6) is by no means unique. For example, if $\alpha \geqslant 0$ we can replace K and $\mathscr{J}$ by K' and $\mathscr{J}'$ where

$$\mathscr{J}'(X) = \mathscr{J}(X) + 2\alpha X, \qquad K' = K - \alpha 1 \tag{4.15}$$

for all $X \in \mathscr{L}(\mathscr{H})$.

Using Theorem 2.3 we may rewrite the generator Z in the form

$$Z(X) = (iH - \tfrac{1}{2}R)X + X(-iH - \tfrac{1}{2}R) + \sum_{n=1}^{\infty} A_n^* X A_n \tag{4.16}$$

where

$$R \equiv \sum_{n=1}^{\infty} A_n^* A_n \tag{4.17}$$

and H are both self-adjoint operators on $\mathscr{H}$. This proves that the CPN semigroups obtained in this section are a subclass of the semigroups constructed in chapter 5 section 5.

Theorem 4.3 *Let $\mathscr{H}$ be finite-dimensional and let T_t be a* CPN *semigroup on $\mathscr{L}(\mathscr{H})$. Then there exists a Hilbert space $\mathscr{K}$, a state $\rho = |\Omega\rangle\langle\Omega|$ on $\mathscr{K}$ and a strongly continuous one-parameter semigroup V_t of isometries on $\mathscr{H} \otimes \mathscr{K}$ such that*

$$T_t(A) = E_\rho\{V_t^*(A \otimes 1)V_t\} \tag{4.18}$$

for all $A \in \mathscr{L}(\mathscr{H})$ and all $t \geqslant 0$.

Proof This is a modification of the proof of Theorem 5.3.6, whose notation we follow. We let B_t be the contraction semigroup on $\mathscr{H}$ with generator $Y = iH - \frac{1}{2}R$, and $\mathscr{J}$ the PMV measure on $X = \mathbb{Z}$ defined by

$$\mathscr{J}(E, \rho) = \sum_{n \in E} A_n \rho A_n^*. \tag{4.19}$$

If $\mathscr{E}^t$ is the QSP on $\mathscr{T}(\mathscr{H})$, then $\mathscr{E}^t$ is related to the CPN semigroup by

$$\mathrm{tr}[\mathscr{E}^t(X_t, \rho)A] = \mathrm{tr}[\rho T_t(A)] \tag{4.20}$$

for all $\rho \in \mathscr{T}(\mathscr{H})$, $A \in \mathscr{L}(\mathscr{H})$ and $t \geqslant 0$ because of the relations between the two generators given by Eq. (5.5.7) and Eq. (4.16).

By Eqs (5.3.33) and (5.3.35)

$$\mathscr{E}^t(A_t^n, \rho) = \int_{0 < t_1 < \cdots < t_n \leqslant t} S_{t-t_n} \mathscr{J}_{\mathbb{Z}} S_{t_n - t_{n-1}} \mathscr{J}_{\mathbb{Z}} \cdots S_{t_1} \rho \, dt_1 \cdots dt_n. \tag{4.21}$$

Therefore if $\rho = |\psi\rangle\langle\psi|$

$$\begin{aligned} \|\psi\|^2 &= \sum_{n=0}^{\infty} \mathrm{tr}[\mathscr{E}^t(A_t^n, \rho)] \\ &= \|B_t\psi\|^2 + \sum_{m_1=1}^{\infty} \int_{t_1=0}^{t} \|B_{t-t_1} A_{m_1} B_{t_1}\psi\|^2 \, dt_1 + \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \int_{t_2=0}^{t} \int_{t_1=0}^{t_2} \\ &\quad \times \|B_{t-t_2} A_{m_2} B_{t_2-t_1} A_{m_1} B_{t_1}\psi\|^2 \, dt_1 \, dt_2 + \cdots \end{aligned} \tag{4.22}$$

which may also be verified directly by differentiation of the right-hand side.

We define Y_∞^n to be the space of finite sequences $\omega = \{m_n, t_n, \ldots, m_1, t_1\}$ such that $m_n \in \mathbb{Z}^+$ and $0 < t_1 < \cdots < t_n < \infty$, and put on it the measure which is the product of the counting measures and the Lebesgue measures. On $Y_\infty = \bigcup_{n=0}^{\infty} Y_\infty^n$ we put the sum of these measures where Y_∞^0 consists of the single point z, with measure 1. If

$$\mathscr{H}' = L^2(Y_\infty, \mathscr{H}) \simeq L^2(Y_\infty) \otimes \mathscr{H} \tag{4.23}$$

we shall define a one-parameter semigroup of isometries V_t on $\mathscr{H}'$. If $t \geqslant 0$ and $\omega = (m_n, t_n, \ldots, m_1, t_1)$ let r be the largest integer for which $t_r \leqslant t$ and define $\omega_t = (m_n, t_n - t, \ldots, m_{r+1}, t_{r+1} - t)$. If $\psi \in \mathscr{H}'$ then we define

$$(V_t\psi)(\omega) = B_{t_1} A_{m_1} B_{t_2 - t_1} \cdots A_{m_r} B_{t - t_r} \psi(\omega_t). \tag{4.24}$$

By Eq. (4.22), V_t is an isometry. The fact that $\{V_t\}_{t \geqslant 0}$ is a strongly continuous semigroup is shown by a direct computation.

If $\Omega \in L^2(Y_\infty)$ is defined by $\Omega(\omega) = \delta_z(\omega)$ for all $\omega \in Y_\infty$ and if $\psi \in \mathscr{H}$ then $\psi \otimes \Omega$ is the element of $L^2(Y_\infty; \mathscr{H})$ such that for all $\omega \in Y_\infty$

$$(\psi \otimes \Omega)(\omega) = \delta_z(\omega)\psi. \tag{4.25}$$

If $M \in \mathscr{L}(\mathscr{H})$ then $M \otimes 1$ is the element of $\mathscr{L}(L^2(Y_\infty, \mathscr{H}))$ such that

$$\{(M \otimes 1)\psi\}(\omega) = \{\psi(\omega)\} \tag{4.26}$$

for all $\psi \in L^2(Y_\infty, \mathscr{H})$. Therefore if $\psi \in \mathscr{H}$ and $M \in \mathscr{L}(\mathscr{H})$

$$\begin{aligned}
\langle E_{|\Omega\rangle\langle\Omega|}\{V_t^*(M \otimes 1)V_t\}\psi, \psi\rangle &= \langle (M \otimes 1)V_t(\psi \otimes \Omega), V_t(\psi \otimes \Omega)\rangle \\
&= \int_{Y_\infty} \langle M V_t(\psi \otimes \Omega)(\omega), V_t(\psi \otimes \Omega)(\omega)\rangle d\omega \\
&= \sum_{n=0}^{\infty} \sum_{m_1=1}^{\infty} \cdots \sum_{m_n=1}^{\infty} \int_{0<t_1<\cdots<t_n<t} \\
&\quad \times \langle M B_{t_1} A_{m_1} B_{t_2 - t_1} A_{m_2} \cdots A_{m_n} B_{t - t_n}\psi, \\
&\quad B_{t_1} A_{m_1}, \ldots, B_{t - t_n}\psi\rangle dt_1 \cdots dt_n \\
&= \operatorname{tr}[M\mathscr{E}^t(X_t, |\psi\rangle\langle\psi|)] \\
&= \operatorname{tr}[T_t(M)|\psi\rangle\langle\psi|].
\end{aligned} \tag{4.27}$$

Since ψ is arbitrary Eq. (4.18) follows. QED

Theorem 4.3 solves the problem posed at the beginning of the chapter, except that V_t is a one-parameter semigroup of isometries instead of a one-parameter unitary group. However, we may pass to a one-parameter unitary group on a yet larger Hilbert space by using the theory of chapter 7 section 2. As it stands the space $\mathscr{K}$ may be regarded as representing the outgoing states of the measuring apparatus or external world.

Notes

1,2 The case for the consideration of completely positive linear maps was made by Kraus (1971) and Lindblad (1976). Theorem 2.1 was proved by Stinespring (1955) while Theorem 2.3 was deduced by Kraus (1971) using Lemma 2.2, which may be found in Dixmier (1969b).

3 Lemma 3.1 was proved by Stinespring (1955). Theorem 3.2 is due to Naimark (1940, 1943) but our proof follows Stinespring (1955).

4 Theorem 4.2 was proved independently by Lindblad (1976) and Gorini *et al.* (1976). In fact Lindblad proved a similar result for a norm continuous CPN semigroup on any Hilbert space. A few other special infinite-dimensional cases were studied by Davies (1972b) and the general infinite-dimensional case by Evans (1975). The proof of Theorem 4.3 is due to Davies (1972b). The Dicke maser model as studied by Hepp and Lieb (1973b) is an example of Theorem 4.1, although this is not mentioned explicitly.

10

General Theory of Open Systems

10.1 Systems interacting with reservoirs

In the previous chapter we discussed under what circumstances a given dynamical semigroup on a Hilbert space might be associated with a Hamiltonian evolution on a larger Hilbert space. There are two physical criticisms of this type of investigation. The first is that the assumption that an open system has a Markovian, or memory-free, evolution is motivated mainly by simplicity and is definitely incorrect for some models although it might be a good approximation for many others. The second is that a more correct approach from the fundamental point of view would be to start with a quantum-mechanical model for the system and reservoir, write down a model Hamiltonian for the interaction between the two, and then investigate the evolution of the system as determined by the Hamiltonian.

Let us be a bit more specific about how this programme is carried out. One supposes that one can make a quantum mechanical model of an infinite reservoir. This may consist of a representation of the CCRs or CARs with an infinite number of degrees of freedom on a Hilbert space $\mathscr{F}$, Fock space in the case of a simple reservoir. The system is represented by a Hilbert space $\mathscr{H}$, and the system plus reservoir by $\mathscr{H} \otimes \mathscr{F}$. The Hamiltonian on $\mathscr{H} \otimes \mathscr{F}$ is

$$H = H_{\mathscr{H}} \otimes 1 + 1 \otimes H_{\mathscr{F}} + \lambda H_I \tag{1.1}$$

where $H_{\mathscr{H}}$ is the system Hamiltonian, $H_{\mathscr{F}}$ is the reservoir Hamiltonian and H_I represents the interaction between system and reservoir. We assume for simplicity that H_I is bounded, so that H is a self-adjoint operator in an obvious sense.

We assume that the reservoir is initially in its equilibrium state ρ_β. For a finite volume reservoir one may put

$$\rho_\beta = \frac{e^{-\beta H_{\mathscr{F}}}}{\mathrm{tr}[e^{-\beta H_{\mathscr{F}}}]} \tag{1.2}$$

but for an infinite volume reservoir $H_{\mathscr{F}}$ does not generally have discrete spectrum, so the trace is not finite. The relevant state is then

$$\rho_\beta = |\Omega_\beta\rangle\langle\Omega_\beta| \tag{1.3}$$

for a non-Fock representation of the CCRs or CARs constructed by the GNS procedure from the infinite volume correlation functions. We also suppose that the state of the system at time zero is ρ and that there are no correlations between the system and reservoir at time zero. The total state at time t is then given by

$$e^{-iHt}(\rho \otimes \rho_\beta)e^{iHt}. \tag{1.4}$$

To find the state of the system at time t we take the partial trace with respect to the reservoir variables. If $\mathscr{B} = \mathscr{T}_s(\mathscr{H} \otimes \mathscr{F})$ and $\mathscr{B}_0 = \mathscr{T}_s(\mathscr{H})$ the partial trace is a bounded linear map $P_0 : \mathscr{B} \to \mathscr{B}_0$ characterised by

$$\mathrm{tr}[(P_0 X)A] = \mathrm{tr}[X(A \otimes 1)] \tag{1.5}$$

where $X \in \mathscr{T}_s(\mathscr{H} \otimes \mathscr{F})$ and $A \in \mathscr{L}_s(\mathscr{H})$. One may alternatively characterise P_0 by the property that if $\{\phi_n\}$ is an orthonormal basis of $\mathscr{F}$ and $\psi = \sum_{n=1}^{\infty} \psi_n \otimes \phi_n$ then

$$P_0(|\psi\rangle\langle\psi|) = \sum_{n=1}^{\infty} |\psi_n\rangle\langle\psi_n|. \tag{1.6}$$

If we identify $\rho \in \mathscr{B}_0$ with $\rho \otimes \rho_\beta \in \mathscr{B}$ then P_0 is a projection of $\mathscr{B}$ onto its subspace $\mathscr{B}_0$. If we define the operator Z on $\mathscr{B}$ by

$$Z(\rho) = -i[H_{\mathscr{H}} \otimes 1 + 1 \otimes H_{\mathscr{F}},\ \rho] \tag{1.7}$$

and the bounded operator A on $\mathscr{B}$ by

$$A(\rho) = -i[H_I, \rho] \tag{1.8}$$

then the invariance of ρ_β with respect to $e^{iH_{\mathscr{F}}t}$ implies that

$$ZP_0 = P_0 Z. \tag{1.9}$$

The state of the system at time $t \geqslant 0$ is given by

$$\rho_t = P_0\{e^{(Z + \lambda A)t}\rho\} \tag{1.10}$$

for all $\rho \in \mathscr{B}_0$.

10.2 A second example

One may obtain formally similar equations when considering the dynamics of a closed system. If $\mathscr{H} = \mathbb{C}^n$, where n is large but finite, then the state space $\mathscr{B}$ is the set of $n \times n$ self-adjoint matrices. Let $P_0 : \mathscr{B} \to \mathscr{B}_0$ be the projection onto the subspace $\mathscr{B}_0$ of real diagonal matrices defined by

$$(P_0 A)_{mn} = A_{mn}\delta_{mn} \tag{2.1}$$

and let the Hamiltonian for the system be $H = H_0 + \lambda H_I$, where H_0 is diagonal but the interaction term H_I is not. If $\rho \in \mathscr{B}_0$, one is frequently interested not in the state at time t but only in the diagonal part of the state

$$\rho_t = P_0(e^{-iHt}\rho\, e^{iHt}). \tag{2.2}$$

If we define the operators Z and A on $\mathscr{B}$ by

$$Z(\rho) = -i[H_0, \rho] \tag{2.3}$$

and

$$A(\rho) = -i[H_I, \rho] \tag{2.4}$$

then

$$ZP_0 = P_0 Z = 0 \tag{2.5}$$

but A does not commute with P_0. The state ρ_t is then given by

$$\rho_t = P_0(e^{(Z+\lambda A)t}\rho) \tag{2.6}$$

as before.

10.3 Stochastic evolution equations

In mathematical physics one often wants to consider an evolution equation

$$f'(t) = \{Z_0 + A(\omega)\}f(t) \tag{3.1}$$

where $f(t)$ lies in a Banach space $\mathscr{B}_0$, for example a state space, and the operator $A(\omega)$ on $\mathscr{B}_0$ depends on a random variable ω. If the equation is stationary, in other words the evolution is unaltered by change of the time origin, then this can be formalised as follows.

We suppose that the random variable ω is a point in a probability space, defined as a set Ω with a σ-field $\mathscr{F}$ and a probability measure $d\omega$. We suppose that the real line has an action on Ω which leaves the probability measure invariant, and that the group of isometries V_t on $L^1(\Omega)$ defined by

$$(V_t f)(\omega) = f(t\omega) \tag{3.2}$$

is strongly continuous with infinitesimal generator W.

We define $\mathscr{B}$ as the Banach space $L^1(\Omega, \mathscr{B}_0)$ of all integrable $\mathscr{B}_0$-valued functions on Ω, and identify $\mathscr{B}_0$ with the constant functions in $\mathscr{B}$. We let $P_0: \mathscr{B} \to \mathscr{B}_0$ be the expectation map

$$P_0 f = \int_\Omega f(\omega)d\omega \tag{3.3}$$

and let $\mathscr{B}_1$ be the set of functions with zero expectation, so that $\mathscr{B} = \mathscr{B}_0 \oplus \mathscr{B}_1$. We also suppose that Z_0 is the generator of a strongly continuous one parameter group of isometries Y_t on $\mathscr{B}_0$.

Lemma 3.1 *There exists a strongly continuous one parameter group of isometries U_t on $\mathscr{B}$ with generator Z such that if $g \in \mathrm{dom}(W)$ and $b \in \mathrm{dom}(Z_0)$ then the element $g \otimes b$ of $\mathscr{B}$ defined by*

$$(g \otimes b)(\omega) = g(\omega)b \tag{3.4}$$

lies in $\mathrm{dom}(Z)$ *and*

$$Z(g \otimes b) = (Yg) \otimes b + g \otimes (Z_0 b). \tag{3.5}$$

Moreover U_t commutes with P_0 for all $t \in \mathbb{R}$.

Proof We define U_t by

$$(U_t f)(\omega) = Y_t\{f(t\omega)\} \tag{3.6}$$

so that U_t is a group of isometries which commutes with P_0. The subspace

$$\mathscr{D} = \mathrm{lin}\{g \otimes b : g \in \mathrm{dom}(W) \quad \text{and} \quad b \in \mathrm{dom}(Z_0)\} \tag{3.7}$$

is dense in $\mathscr{B}$ and since

$$U_t(g \otimes b) = (V_t g) \otimes (Y_t b) \tag{3.8}$$

it follows that $t \to U_t f$ is continuous for all $f \in \mathscr{D}$. Therefore U_t is a strongly continuous group. A simple calculation now verifies Eq. (3.5) and establishes that $\mathscr{D} \subseteq \mathrm{dom}(Z)$. QED

We now suppose that the operator $A(\omega)$ is a measurable function of ω and that $\|A(\omega)\| \leqslant K$ for some K independent of ω. The formula

$$(Af)(\omega) = A(\omega)f(\omega) \tag{3.9}$$

defines a bounded operator A on $\mathscr{B}$, so $(Z + A)$ is the generator of a strongly continuous one parameter group on $\mathscr{B}$. If $f \in \mathscr{B}_0$ then the expected value of f at time $t \geqslant 0$ is rigorously defined as

$$\bar{f}(t) = P_0\{e^{(Z+A)t} f\}. \tag{3.10}$$

We should mention that we have deliberately made technical assumptions above so that the problem can be posed in a functional analytic manner, and that specifically probabilistic methods enable one to study a wider class of random evolution equations.

This example differs from the example of section 1 in the replacement of an infinite quantum-mechanical reservoir by a stochastic process, and for this reason is less fundamental. In the present example further analysis depends upon the assumption of ergodic or mixing properties for the stochastic process, which in the model of section 1 have to be proved from the microscopic equations. The difficulty of proving ergodicity may be gauged by the fact that after a century of effort no satisfactory general theory has yet emerged.

10.4 Derivation of the master equation

It is advantageous to abstract the essential features of the above examples. We let P_0 be a projection on a Banach space $\mathscr{B}$ and put $\mathscr{B}_0 = P_0\mathscr{B}$, $P_1 = 1 - P_0$ and $\mathscr{B}_1 = P_1\mathscr{B}$ so that

$$\mathscr{B} = \mathscr{B}_0 \oplus \mathscr{B}_1. \tag{4.1}$$

We suppose that Z is the infinitesimal generator of a strongly continuous one-parameter group of isometries U_t on $\mathscr{B}$ satisfying

$$U_t P_0 = P_0 U_t \tag{4.2}$$

for all $t \in \mathbb{R}$. We also suppose that A is a bounded perturbation of Z and study the one parameter group $e^{(Z+\lambda A)t}$ on $\mathscr{B}$. If $\rho \in \mathscr{B}_0$ then the state ρ_t at time t is defined as

$$\rho_t = P_0\, e^{(Z+\lambda A)t}\rho. \tag{4.3}$$

We introduce the notation $A_{ij} = P_i A P_j$ and $Z_i = ZP_i = P_i Z$. Then we define

$$U_t^\lambda = e^{\{Z+\lambda(A_{00}+A_{11})\}t}. \tag{4.4}$$

Since $(A_{00} + A_{11})$ commutes with P_0 it follows by Eq. (1.9.7) that U_t^λ satisfies

$$U_t^\lambda P_0 = P_0 U_t^\lambda. \tag{4.5}$$

Theorem 4.1 *If $\rho \in \mathscr{B}_0$ then ρ_t is the solution of the integral equation*

$$\rho_t = U_t^\lambda \rho + \lambda^2 \int_{s=0}^{t} \int_{u=0}^{s} U_{t-s}^\lambda A_{01} U_{s-u}^\lambda A_{10} \rho_u \, du\, ds \tag{4.6}$$

for all $t \geqslant 0$.

Proof From Eqs (1.9.6) and (4.5) we obtain

$$\rho_t = U_t^\lambda \rho + \lambda \int_{s=0}^{t} U_{t-s}^\lambda A_{01}(P_1\, e^{(Z+\lambda A)s}\rho)ds \tag{4.7}$$

and

$$P_1 e^{(Z+\lambda A)s}\rho = \lambda \int_{u=0}^{s} U_{s-u}^\lambda A_{10}\rho_u\, du. \tag{4.8}$$

Substituting Eq. (4.8) into Eq. (4.7) yields Eq. (4.6). QED

The importance of Eq. (4.6) is that it is an operator equation on $\mathscr{B}_0$ instead of $\mathscr{B}$. It is usually written in the differentiated form

$$\rho_t' = (Z_0 + \lambda A_{00})\rho_t + \lambda^2 \int_{u=0}^{t} A_{01} U_{t-u}^{\lambda} A_{10} \rho_u \, du \tag{4.9}$$

when it is called the (generalised) master equation of the open system $\mathscr{B}_0$. Note however that Z is an unbounded operator so the master equation requires more care in interpretation than Eq. (4.6). One of the most interesting features of the master equation is that the last term is an integral depending on the value of ρ at earlier times, and so appears to describe a "memory effect". This anthropocentric terminology should be treated with caution because the dynamics associated with Eq. (4.9) is still reversible. One can obtain an equation similar to Eq. (4.9) for $t < 0$, where the "memory" is of the future, being an integral over $\{s: t \leqslant s \leqslant 0\}$.

The study of the master equation has been pursued in an enormous variety of situations in quantum statistical mechanics. In the next section we outline a few methods of finding properties of solutions of the master equation and approximate versions of it.

10.5 Solutions of the master equation

One of the obvious methods of studying the master equation is by taking Laplace transforms. If we define

$$\hat{\rho}(z) = \int_0^{\infty} \rho(t) e^{-tz} \, dt \tag{5.1}$$

where $\operatorname{Re} z > 0$ then we obtain

$$-\rho_0 + z\hat{\rho}(z) = (Z_0 + \lambda A_{00})\hat{\rho}(z) + \lambda^2 A_{01}(z - Z_1 - \lambda A_{11})^{-1} A_{10}\hat{\rho}(z) \tag{5.2}$$

so

$$\hat{\rho}(z) = \{z - Z_0 - \lambda A_{00} - \lambda^2 A_{01}(z - Z_1 - \lambda A_{11})^{-1} A_{10}\}^{-1} \rho_0 \tag{5.3}$$

which can then in principle be used to find ρ_t. While occasionally tractable the difficulty with this method lies in the abstractness of the expressions involved. The A_{ij} are operators on the space $\mathscr{B}$, which is itself a space of trace-class operators, and are hence sometimes called superoperators. The solution of Eq. (5.3) is dependent on a detailed examination of the spectral properties of these superoperators, which presents formidable problems.

A second possibility is to simplify the master equation by dropping memory effects. This can be justified in a non-rigorous fashion by reference to the relative length of various relaxation times. Equation (4.9) is first changed to

$$\rho_t' = (Z_0 + \lambda A_{00})\rho_t + \lambda^2 \int_{u=0}^{t} A_{01}\, e^{Z_1 u} A_{10}\, e^{-Z_0 u}(e^{Z_0 t}\rho)du \tag{5.4}$$

and then to

$$\rho_t' = (Z_0 + \lambda A_{00} + \lambda^2 K)\rho_t \tag{5.5}$$

where

$$K = \int_0^\infty A_{01}\, e^{Z_1 u} A_{10}\, e^{-Z_0 u}\, du. \tag{5.6}$$

Equation (5.5) has the solution

$$\rho_t = e^{(Z_0 + \lambda A_{00} + \lambda^2 K)t}\rho \tag{5.7}$$

whose properties can be easily determined. Although frequently very successful, a criticism of this method is that the approximations are generally uncontrolled and that they may destroy some qualitative effects of the memory term.

One of the important questions about the master equation is whether the limit of ρ_t exists as $t \to +\infty$, and if so what that limit is. The second question can be approached by the use of the resolvent formalism. If $\rho(t) \to \sigma$ as $t \to +\infty$ then

$$\sigma = \lim_{z \uparrow 0} z\hat{\rho}(z). \tag{5.8}$$

Much of the work in this approach therefore centres on an analysis of the singularity of $\hat{\rho}(z)$ at 0.

Similarly the behaviour as $t \to +\infty$ of the approximate solution Eq. (5.7) can often be found by a spectral analysis of the operator $(Z_0 + \lambda A_{00} + \lambda^2 K)$. This is, however, inconclusive because one would expect the approximation to become less accurate as time progresses.

The only cases where one can solve the master equation rigorously seem to be those which reduce to single particle problems. A Hamiltonian which is quadratic in the creation and annihilation operators is equivalent via a Bogoliubov transformation to a free Hamiltonian, which commutes with the number operator on Fock space and is determined by its restriction to the single particle space. In the Friedrichs' model the Hamiltonian need not be quadratic but it still commutes with a number operator, and one can study the evolution in the single particle subspace without approximation.

It is possible to make substantial progress in the rigorous study of the master equation in the weak coupling limit $\lambda \to 0$. Because the diffusive behaviour becomes slower as $\lambda \to 0$ one needs to keep $\tau = \lambda^2 t$ constant as

$\lambda \to 0$ in order to retain the diffusion terms. In the example of section 1 if the interaction term in the Hamiltonian of Eq. (1.1) is

$$H_I = \sum_{r=1}^{n} A_r \otimes B_r \tag{5.9}$$

then the analysis depends crucially upon certain properties of the multi-time correlation functions

$$f(t_1, \ldots, t_n) = \langle B_{r(1)}(t_1), \ldots, B_{r(n)}(t_n)\Omega, \Omega\rangle \tag{5.10}$$

where

$$B_r(t) = e^{iH_{\mathscr{F}}t} B_r\, e^{-iH_{\mathscr{F}}t} \tag{5.11}$$

One actually needs estimates of the decay of the truncated n-point functions as the times converge independently to infinity, which are uniform in n as $n \to \infty$. Such estimates are straightforward when the reservoir Hamiltonian $H_{\mathscr{F}}$ is free, but are very difficult to prove in more general situations. Using these estimates it is possible to prove that Eq. (5.7) is asymptotically exact as $\lambda \to 0$ for times of order λ^{-2}. More precisely one can show that for all $a > 0$

$$\lim_{\lambda \to 0} \left\{ \sup_{0 \leqslant \lambda^2 t \leqslant a} \|\rho_t - e^{(Z_0 + \lambda A_{00} + \lambda^2 K)t}\rho\| \right\} = 0. \tag{5.12}$$

One of the interesting questions about the weak coupling limit is whether

$$\lim_{\lambda \to 0} \lim_{t \to \infty} \rho_{\lambda, t} = \lim_{\tau \to \infty} \lim_{\lambda \to 0} \rho_{\lambda, \lambda^{-2}\tau} \tag{5.13}$$

The method described above enables one to prove in many cases that the limit on the R.H.S. exists and is the Gibbs state of the system at the temperature of the reservoir. However, one cannot interchange the order of the limits using Eq. (5.12) because the limit there is not uniform in time. Indeed there are analogous problems in probability theory where it is known that the two limits exist and are not equal. In such circumstances which side of Eq. (5.13) one studies depends upon one's interests and not upon arguments that one is more physically correct than the other. The situation is precisely the same as with Poincaré cycles; they are so infrequent as to be of no practical interest, and can be eliminated mathematically by taking one's limits in the proper order.

Notes

1,2 Although initiated by Pauli in 1928, the theory of master equations began to flourish only in the 1950s with a series of papers (Montroll, 1960; Nakajima, 1958; Prigogine and Resibois, 1961; Van Hove, 1955, 1957; Zwanzig, 1960, 1964) which were followed by a great variety of applications. The original master equations were the approximations obtained by dropping memory effects, while the generalised master equations included these. For recent surveys of the subject see Peier (1972) and Haake (1973). Examples of systems which do not decay to their equilibrium states exponentially fast were given by Emch (1966) and Radin (1970).

In the example of section 2 it is possible to use either the trace norm or the Hilbert–Schmidt norm (Emch, 1964; Emch and Sewell, 1968), an advantage of the latter being that Hilbert spaces are more familiar than Banach spaces. In the example of section 1, however, one cannot let $\mathscr{B}$ be the space of Hilbert–Schmidt operators unless one takes into account that P_0 then becomes an unbounded operator.

3 Stochastic differential equations of this type were studied by Papanicolaou and Kohler (1974), Papanicolaou and Varadhan (1973) and Martin and Emch (1975). The connection with master equations was made explicit by Davies (1975a). See also Lewis and Thomas (1974, 1975).

4,5 Derivations of the generalised master equation may be found in Haake (1973) and Peier (1972), where the non-rigorous methods of solution are also discussed, and at a more abstract level in Davies (1974a, 1976) and Presutti *et al.* (1972). The decay to equilibrium for quadratic Hamiltonians was studied by Ford *et al.* (1965), Sewell (1974) and Davies (1972a), (1973), and for the Friedrichs' model by Davidson and Kozak (1973), Grecos and Prigogine (1972) and Davies (1974c). The exact solution of the master equation in the weak coupling limit was obtained by Pulè (1974), Davies (1974a, 1976) and Davies and Eckmann (1975). Similar results for a singular coupling limit were obtained by Ford *et al.* (1965), Davies (1972a), Hepp and Lieb (1973b), Lewis and Thomas (1974, 1975) and Gorini and Kossakowski (1976). An example where the limits $\lambda \to 0$ and $t \to \infty$ may not be interchanged was given by Veutsel and Friedlin (1970). Other mathematically rigorous literature on the return to equilibrium may be found in Bongaarts *et al.* (1973), Haag *et al.* (1974), Lanford and Robinson (1972) and Robinson (1973).

Appendix

1 Let $\rho_n \in \mathscr{T}_s(\mathscr{H})^+$ be a sequence which converges in the weak operator topology to $\rho \in \mathscr{L}_s(\mathscr{H})^+$. If there exists a constant K such that $\mathrm{tr}[\rho_n] \leqslant K$ for all n then $\rho \in \mathscr{T}_s(\mathscr{H})^+$ and

$$\mathrm{tr}[\rho] \leqslant \liminf_{n \to \infty} \mathrm{tr}[\rho_n].$$

Moreover ρ_n converges to ρ in trace norm if and only if

$$\lim_{n \to \infty} \mathrm{tr}[\rho_n] = \mathrm{tr}[\rho].$$

(Dell'Antonio, 1967; Davies, 1969)

2 Let $\{A_n\}_{n=0}^{\infty}$ be self-adjoint operators on $\mathscr{H}$ with $\mathrm{sp}(A_n) \subseteq [a, b]$ for all n and $A_n \xrightarrow{s} A$. Let f be a bounded function on $[a, b]$ which is continuous except on the closed set $S \subseteq [a, b]$ and let K be the range of the spectral projection of A associated with the set $[a, b] \setminus S$. Then

$$\lim_{n \to \infty} f(A_n)\psi = f(A)\psi$$

for all $\psi \in K$.

(Dunford and Schwartz, 1963, p. 922)

3 Let $P_{(\cdot)}$ and U_t be a system of imprimitivity on $\mathscr{H}$ for the action of the real line on itself by translation. Then there is a Hilbert space $\mathscr{N}$ and an isomorphism $\mathscr{H} \simeq L^2(\mathbb{R}, \mathscr{N})$ under which

$$(P_E f)(x) = \chi_E(x) f(x)$$

and

$$(U_t f)(x) = f(x - t)$$

for all $f \in L^2(\mathbb{R}, \mathscr{N})$, $t \in \mathbb{R}$ and Borel sets $E \subseteq \mathbb{R}$.

(Mackey, 1949)

4 If $\mathscr{S}$ is a commuting family of continuous affine maps of a compact convex set into itself then there exists a point $x \in X$ such that $Sx = x$ for all $S \in \mathscr{S}$.

(Dunford and Schwartz, 1958)

5 If X is a compact convex set in a locally convex topological linear space then X is the smallest closed convex set containing all its extreme points.

(Dunford and Schwartz, 1958; Alfsen, 1971)

6 Let X be a compact convex set in a locally convex topological linear space and S a closed subset of X containing all the extreme points of X. Then for every point $x_0 \in X$ there exists a probability measure μ on S such that

$$f(x_0) = \int_S f(x)\mu(dx)$$

for all continuous affine functionals f on X. x_0 is called the barycentre of μ.

(Alfsen, 1971)

7 If W is the infinitesimal generator of a strongly continuous one parameter semigroup of contractions on a Banach space $\mathscr{B}$ then the spectrum of W is contained in

$$\{z \in \mathbb{C} : \operatorname{Re} z \leqslant 0\}.$$

(Dunford and Schwartz, 1958; Yosida, 1965)

8 If Z is the generator of a strongly continuous one-parameter semigroup of contractions S_t on the Banach space $\mathscr{B}$, A is a bounded operator on $\mathscr{B}$, and T_t is the semigroup with infinitesimal generator $(Z + A)$ then

$$T_t f = \lim_{n \to \infty} (S_{t/n}\, e^{At/n})^n f$$

in norm for all $f \in \mathscr{B}$ and all $t \geqslant 0$.

(Trotter, 1959; Nelson, 1964; Chernoff, 1968)

References

Alfsen, E. M. (1971). *Compact convex sets and boundary integrals.* Springer, Berlin.

Alguard, M. J. and Drake, C. W. (1973). Stark beats in Lyman series emission. *Phys. Rev.* **A8**, 27–36.

Allcock, G. R. (1969). The time of arrival in quantum mechanics. *Ann. Phys.* **53**, 253–348.

Amrein, W. O. (1969). Localizability for particles of mass zero. *Helv. Phys. Acta* **42**, 149–190.

Andra, H. J. (1970). Zero field quantum beats subsequent to beam foil excitation. *Phys. Rev. Lett.* **25**, 325–327.

Andra, H. J. (1974). Fine structure, hyperfine structure and Lamb shift measurements by the beam foil technique. *Physica Scripta* **9**, 257–280.

Araki, H. (1960). Hamiltonian formalism and the canonical commutation relations in quantum field theory. *J. Math. Phys.* **1**, 492–504.

Aronsajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68**, 337–404.

Arrechi, F. T., Courtens, E., Gilmore, R. and Thomas, H. (1972). Atomic coherent states in quantum optics. *Phys. Rev.* **A6**, 2211–2237.

Aslaken, E. W. and Klauder, J. R. (1969). Continuous representation theory using the affine group. *J. Math. Phys.* **10**, 2267–2275.

Bargmann, V. (1961). On a Hilbert space of analytic functions and an associated integral transform I. *Comm. Pure Applied Math.* **14**, 187–214.

Bargmann, V., Butera, P., Girardello, L. and Klauder, J. R. (1971). On the completeness of the coherent states. *Reports Math. Phys.* **2**, 221–228.

Barut, A. O. and Girardello, L. (1971). New "coherent" states associated with non-compact groups. *Commun. Math. Phys.* **21**, 41–55.

Bashkin, S. (1971). Recent advanced in beam foil spectroscopy. In *Atomic Physics* **2** (P. G. H. Sandars, ed.). Plenum Press, London.

Birkhoff, G. and von Neumann, J. (1936). The logic of quantum mechanics. *Ann. Math.* **37**, 823–843.

Bongaarts, P. J. M., Fannes, M. and Verbeure, A. (1973). A remark on ergodicity, dissipativity, return to equilibrium. *Physica* **68**, 587–594.

Bratteli, O. and Robinson, D. W. (1975). Unbounded derivations of C*-algebras. *Commun. Math. Phys.* **42**, 253–268.

Carruthers, P. and Nieto, M. M. (1968). Phase and angle variables in quantum mechanics. *Rev. Mod. Phys.* **40**, 411–440.

Chernoff, P. R. (1968). Note on product formulas for operator semigroups. *J. Funct. Anal.* **2**, 238–242.

Chung, K. L. (1967). *Markov chains with stationary transition probabilities.* 2nd edn. Springer, Berlin.

Cohen, L. (1966). Can quantum mechanics be formulated as a classical probability theory? *Phil. of Science* **33**, 317–322..

Cooper, J. L. B. (1947). One parameter semi-groups of isometric operators on Hilbert space. *Ann. Math.* **48**, 827–842.

Dahn, G. (1968). An attempt at an axiomatic formulation of quantum mechanics and more general theories IV. *Commun. Math. Phys.* **9**, 192–211.

Dahn, G. (1973). Two equivalent criteria for modularity of the lattice of all physical decision effects. *Commun. Math. Phys.* **30**, 69–78.

Davidson, R. and Kozak, J. J. (1973). On the relaxation to quantum statistical equilibrium of the Wigner–Weisskopf atom in a one-dimensional radiation field V. Exact solution for infinite systems. *J. Math. Phys.* **14**, 423–431.

Davies, E. B. (1969). Quantum stochastic processes. *Commun. Math. Phys.* **15**, 277–304.

Davies, E. B. (1970a). On the repeated measurement of continuous observables in quantum mechanics. *J. Funct. Anal.* **6**, 318–346.

Davies, E. B. (1970b). Quantum stochastic processes II. *Commun. Math. Phys.* **19**, 83–105.

Davies, E. B. (1971). Quantum stochastic processes III. *Commun. Math. Phys.* **22**, 51–70.

Davies, E. B. (1972a). Diffusion for weakly coupled quantum oscillators. *Commun. Math. Phys.* **27**, 309–325.

Davies, E. B. (1972b). Some contraction semigroups in quantum probability. *Z. Wahrschein* **23**, 261–273.

Davies, E. B. (1973). The harmonic oscillator in a heat bath. *Commun. Math. Phys.* **33**, 171–186.

Davies, E. B. (1974a). Markovian master equations. *Commun. Math. Phys.* **39**, 91–110.

Davies, E. B. (1974b). Symmetries of compact convex sets. *Quart. J. Math.* Oxford (2) **25**, 323–328.

Davies, E. B. (1974c). Dynamics of a multi-level Wigner–Weisskopf atom. *J. Math. Phys.* **15**, 2036–2041.

Davies, E. B. (1976). Markovian master equations II, *Math. Annalen.* **219**, 147–158.

Davies, E. B. (1975a). Markovian master equations III, *Ann. Inst. Henri Poincaré.* **11**, 265–273.

Davies, E. B. (1975b). A model for absorption or decay. *Helv. Phys. Acta.* **48**, 365–382.

Davies, E. B. and Eckmann, J.-P. (1975). Time decay for fermion systems with persistent vacuum. *Helv. Phys. Acta.* **48**, 731–742.

Davies, E. B. and Lewis, J. T. (1970). An operational approach to quantum probability. *Commun. Math. Phys.* **17**, 239–260.

de Korvin, A. (1971). Complete sets of expectations on a von Neumann algebra. *Quart. J. Math.* (Oxford) (2) **22**, 135–142.

Dell'Antonio, G. F. (1967). On the limits of sequences of normal states. *Commun. Pure Applied Math.* **20**, 413–429.

Dixmier, J. (1969a). *Les C*-algèbres et leurs representations.* 2nd edn. Gauthier-Villars, Paris.

Dixmier, J. (1969b). *Les algèbres d'operateurs dans l'espace hilbertien.* 2nd edn. Gauthier-Villars, Paris.

Dunford, N. and Schwartz, J. T. (1958). *Linear operators.* Part 1. Interscience, New York.

Dunford, N. and Schwartz, J. T. (1963). *Linear operators.* Part 2. Interscience, New York.

Eckstein, H. and Siegert, A. J. F. (1971). On a reinterpretation of decay experiments. *Ann. Phys.* **68**, 509–520.

Edwards, C. M. (1970). The operational approach to algebraic quantum theory I. *Commun. Math. Phys.* **16**, 207–230.

Edwards, C. M. (1971). Classes of operations in quantum theory. *Commun. Math. Phys.* **20**, 26–56.

Emch, G. G. (1964). Coarse graining in Liouville space and master equation. *Helv. Phys. Acta* **37**, 532–534.

Emch, G. G. (1966). Non-Markovian model for the approach to equilibrium. *J. Math. Phys.* **7**, 1198–1206.

Emch, G. G. and Sewell, G. L. (1968). Nonequilibrium statistical mechanics of open systems. *J. Math. Phys.* **9**, 946–958.

Evans, D. (1975). D. Phil. thesis. Oxford.

Ford, G. W., Kac, M. and Mazur, P. (1965). Statistical mechanics of assemblies of coupled oscillators. *J. Math. Phys.* **6**, 504-515.

Ford, G. W. and Lewis, J. T. (1974). Measured correlations and the Wigner distribution. Preprint.

Friedman, C. N. (1971/72). Semigroup product formulas, compression and continual observation in quantum mechanics. *Indiana Univ. Math. J.* **21**, 1001–1011.

Frigerio, A. (1974). Quasi-local observables and the problem of measurement in quantum mechanics. *Ann. Inst. Henri Poincaré* A **21**, 259–270.

Furry, W. H. (1966). Some aspects of the quantum theory of measurement. *Lectures in theoretical Physics* **8a**, 1–64.

Gerjouy, E. (1973). Is the principle of superposition really necessary? In *Contemporary research in the foundations and philosophy of quantum theory*. (C. A. Hooker, ed.), pp. 114–142. D. Reidel Publ. Co., Holland.

Glauber, R. J. (1963). Coherent and incoherent states of the radiation field. *Phys. Rev.* **131**, 2766–2788.

Glauber, R. J. (1968). Photon statistics. In *Fundamental problems in statistical mechanics* II. (E. G. D. Cohen, ed.). North Holland Publ. Co., Amsterdam.

Goldberger, M. L. and Watson, K. M. (1964). Measurement of time correlations for quantum-mechanical systems. *Phys. Rev.* (2) **134**, B 919–928.

Gordon, J. P. and Louisell, W. H. (1966). Simultaneous measurement of non-commuting observables. In *Physics of quantum electronics*, pp. 833–840. (P. L. Kelley *et al.*, eds). McGraw-Hill, New York.

Gorini, V. and Kossakowski, A. (1976). N-level system in contact with a singular reservoir. *J. Math. Phys.* in press.

Gorini, V., Kossakowski, A. and Sudarshan, E. C. G. (1976). Completely positive dynamical semigroups on N-level systems. *J. Math. Phys.* **17**, 821–825.

Grecos, A. P. and Prigogine, I. (1972). Kinetic and ergodic properties of quantum systems—the Friedrichs' model. *Physica* **59**, 77–96.

Gudder, S. (1968). Joint distributions of observables. *J. Math. Mech.* **18**, 325–335.

Haag, R. and Kastler, D. (1964). An algebraic approach to quantum field theory. *J. Math. Phys.* **5**, 848–861.

Haag, R., Kastler, D. and Trych-Polymeyer, E. B. (1974). Stability and equilibrium states. *Commun. Math. Phys.* **38**, 173–193.

Haake, F. (1973). Statistical treatment of open systems by generalized master equations. *Springer tracts in modern physics* **66**.

Halmos, P. R. (1950). *Measure theory*. van Nostrand, England.

Hartkamper, A. and Neumann, H. (eds) (1974). *Springer lecture notes in physics* **29**.

Hellwig, K. E. and Kraus, K. (1969). Pure operations and measurements. *Commun. Math. Phys.* **11**, 214–220.

Hellwig, K. E. and Kraus, K. (1970). Operations and measurements II. *Commun. Math. Phys.* **16**, 142–147.

Hepp, K. (1972). Quantum theory of measurement and macroscopic observables. *Helv. Phys. Acta* **45**, 237–248.

Hepp, K. and Lieb, E. H. (1973a). Equilibrium statistical mechanics of matter interacting with the quantized radiation field. *Phys. Rev.* **A8**, 2517–2525.

Hepp, K. and Lieb, E. H. (1973b). Phase transitions in reservoir-driven open systems with applications to lasers and superconductors. *Helv. Phys. Acta* **46**, 573–603.

Hille, E. and Phillips, R. S. (1957). *Functional analysis and semigroups.* 2nd edn. Coll. Publ. Amer. Math. Soc., Providence, R. I.

Holevo, A. S. (1972). An analogue of the theory of statistical decisions in non-commutative probability theory. *Trans. Moscow Math. Soc.* **26**, 133–149.

Holevo, A. S. (1973). Statistical decision theory for quantum systems. *J. Multivariate Anal.* **3**, 337–394.

Horwitz, L. P., LaVita, J. A. and Marchand, J. P. (1971). The inverse decay problem. *J. Math. Phys.* **12**, 2537–2543.

Hunt, G. A. (1956). Semigroups of measures on Lie groups. *Trans. Amer. Math. Soc.* **81**, 264–293.

Hunziker, W. (1972). A note on symmetry operations in quantum mechanics. *Helv. Phys. Acta* **45**, 233–236.

Husimi, K. (1940). Some formal properties of the density matrix. *Proc. Phys. Math. Soc. Japan* (3) **22**, 264–314.

Ingarden, R. S. (1973). Generalized irreversible thermodynamics and its application to x lasers. *Acta Phys. Polon.* A **43**, 1–35.

Jauch, J. M. and Piron, C. (1967). Generalized localizability. *Helv. Phys. Acta* **40**, 559–570.

Jones, V. (1976). Quantum mechanics over fields of non-zero characteristic. *Lett. Math. Phys.* **1**, 99–103.

Kadison, R. V. (1951). Isometries of operator algebras. *Ann. Math.* **54**, 325–338.

Kadison, R. V. (1965). Transformations of states in operator theory and dynamics. *Topology* **3** suppl. 2, 177–198.

Kano, Y. (1964). Probability distribution functions relating to blackbody radiation. *J. Phys. Soc. Japan* **19**, 1555–1560.

Kato, T. (1966). *Perturbation theory for linear operators.* Springer, Berlin.

Kelley, J. L. (1955). *General topology.* van Nostrand, England.

Klauder, J. R. (1966). Improved version of optical equivalence theorem. *Phys. Rev. Lett.* **16**, 534–536.

Klauder, J. R. (1970). Exponential Hilbert space, Fock space revisited. *J. Math. Phys.* **11**, 609–630.

Klauder, J. R. and Sudarshan, E. C. G. (1968). *Fundamentals of quantum optics.* Benjamin, U.S.A.

Kossakowski, A. (1972a). On quantum statistical mechanics of non-hamiltonian systems. *Rep. Math. Phys.* **3**, 247–274.

Kossakowski, A. (1972b). On necessary and sufficient conditions for a generator of a quantum dynamical semigroup. *Bull. Acad. Polon. Sci. Ser. Math. Astr. et Phys.* **20**, 1021–1025.

Kraus, K. (1971). General state changes in quantum theory. *Ann. Phys.* **64**, 311–335.

Kuryshkin, V. V. (1972). La mécanique quantique avec une fonction non-negative de distribution dans l'espace de phases. *Ann. Inst. Henri Poincaré* A **17**, 81–95.

Kuryshkin, V. V. (1973). Some problems of quantum mechanics possessing a non-negative phase-space distribution function. *Inter. J. Theor. Phys.* **7**, 451–466.

Lanford III, O. E. and Robinson, D. W. (1972). Approach to equilibrium of free quantum systems. *Commun. Math. Phys.* **24**, 193–210.

Lax, P. D. and Phillips, R. S. (1967). *Scattering theory.* Academic Press, London and New York.

Lerner, E. C., Huang, H. W. and Walters, G. E. (1970). Some mathematical properties of oscillator phase operators. *J. Math. Phys.* **11**, 1679–1684.

Lewis, J. T. and Thomas, L. C. (1974). A characterization of regular solutions of a linear stochastic differential equation. *Z. Wahrschein.* **30**, 45–55.

Lewis, J. T. and Thomas, L. C. (1975). On the existence of a class of stationary quantum stochastic processes. *Ann. Inst. Henri Poincaré* **22**A, 241–248.

Lieb, E. H. (1973). The classical limit of quantum spin systems. *Commun. Math. Phys.* **31**, 327–340.

Lindblad, G. (1976). On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* in press.

Lu, E. Y. C. (1971). New coherent states of the electromagnetic field. *Lett. al Nuovo Cim.* **2**, 1241–1244.

Ludwig, G. (1967). Attempt of an axiomatic foundation of quantum mechanics and more general theories II. *Commun. Math. Phys.* **4**, 331–348.

Ludwig, G. (1968). Attempt of an axiomatic foundation of quantum mechanics and more general theories III. *Commun. Math. Phys.* **9**, 1–12.

Ludwig, G. (1970). Deutung des Begriffs physikalische Theorie. *Springer lecture notes in physics* **4**.

Lumer, G. and Phillips, R. S. (1961). Dissipative operators in a Banach space. *Pacific J. Math.* **11**, 679–698.

Macek, J. (1970). Theory of atomic lifetime measurements. *Phys. Rev.* **A1**, 618–627.

McKenna, J. and Frisch, H. L. (1966). Quantum mechanical microscopic Brownian motion. *Phys. Rev.* (2) **145**, 93–110.

Mackey, G. W. (1949). A theorem of Stone and von Neumann. *Duke Math. J.* **16**, 313–326.

Mackey, G. W. (1963). *Mathematical foundations of quantum mechanics.* Benjamin, U.S.A.

Manuceau, J. (1968). C*-algèbre de relations de commutation. *Ann. Inst. Henri Poincaré* A**8**, 139–161.

Martin, P. and Emch, G. G. (1975). A rigorous model sustaining van Hove's phenomenon. *Helv. Phys. Acta.* **48**, 59–78.

Mathews, P. M. and Eswaran, K. (1974). Simultaneous uncertainties of the cosine and sine operators. *Nuovo Cim.* **19B**, 99–104.

Mehra, J. and Sudarshan, E. C. G. (1972). Some reflections on the nature of entropy and the second law of thermodynamics. *Nuovo Cim.* **11B**, 215–256.

Mielnik, B. (1969). Theory of filters. *Commun. Math. Phys.* **15**, 1–46.

Montroll, E. W. (1960). Nonequilibrium statistical mechanics. *Lectures in Theor. Phys.* **3**, 221–325.

Moyal, J. E. (1949). Quantum mechanics as a statistical theory. *Proc. Camb. Phil. Soc.* **45**, 99–124.

Naimark, M. A. (1940). Spectral functions of a symmetric operator. *Izv. Akad. Nauk. SSSR Ser. Math.* **4**, 277–318.

Naimark, M. A. (1943). On a representation of additive operator set functions. *C.R. (Doklady) Acad. Sci. URSS* **41**, 359–361.

Nakajima, S. (1958). On quantum theory of transport phenomena. *Prog. Theor. Phys.* **20**, 948–959.

Nakamura, M. and Turumaru, T. (1954). Expectations in an operator algebra. *Tohoku Math. J.* **6**, 182–188.

Nelson, E. (1964). Feynman integrals and the Schrodinger equation. *J. Math. Phys.* **5**, 332–343.

Nelson, E. (1973). The free Markov field. *J. Funct. Anal.* **12**, 211–227.

Newton, T. D. and Wigner, E. P. (1949). Localized states for elementary systems. *Rev. Mod. Phys.* **21**, 400–406.

Papanicolaou, G. C. and Kohler, W. (1974). Asymptotic theory of mixing stochastic ordinary differential equations. *Commun. Pure Applied Math.* **27**, 641–668.

Papanicolaou, G. C. and Varadhan, S. R. S. (1973). A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations. *Commun. Pure Applied Math.* **26**, 497–524.

Park, J. L. and Margenau, H. (1968). Simultaneous measurability in quantum theory. *Inter. J. Theor. Phys.* **1**, 211–283.

Peier, W. (1972). Time evolution of open systems. I. Master equations. *Physica* **57**, 565–584.

Perelomov, A. M. (1972). Coherent states for arbitrary Lie groups. *Commun. Math. Phys.* **26**, 222–236.

Perelomov, A. M. (1975). Coherent states and symmetric spaces. *Commun. Math. Phys.* **44**, 197–210.

Poulsen, N. K. S. (1970). Regularity aspects of the theory of infinite-dimensional representations of Lie groups. Ph.D. thesis. M.I.T.

Presutti, E., Scacciatelli, E., Sewell, G. L. and Wanderlingh, F. (1972). Studies in the C^*-algebraic theory of non-equilibrium statistical mechanics: dynamics of open and mechanically driven systems. *J. Math. Phys.* **13**, 1085–1098.

Prigogine, I. and Resibois, P. (1961). On the kinetics of the approach to equilibrium. *Physica* **27**, 629–646.

Pulè, J. (1974). The Bloch equations. *Commun. Math. Phys.* **38**, 241–256.

Radcliffe, J. M. (1971). Some properties of coherent spin states. *J. Phys.* A**4**, 313–323.

Radin, C. (1970). Approach to equilibrium in a simple model. *J. Math. Phys.* **11**, 2945–2955.

Reed, M. and Simon, B. (1972). *Methods of modern mathematical physics.* Vol. 1, *Functional analysis.* Academic Press, London and New York.

Robinson, D. W. (1973). Return to equilibrium. *Commun. Math. Phys.* **31**, 171–189.

Rocca, F. (1966). La P-representation dans la formalisme de la convolution gauche. *Comptes Rendus Acad. Sci. Paris* **262**, A 547–549.

Sakai, S. (1971). C^*-algebras and W^*-algebras. Springer, Berlin.

Schatten, R. (1960). Norm ideals of completely continuous operators. Springer, Berlin.

Schrader, R. and Uhlenbrock, D. A. (1975). Markov structures on Clifford algebras. *J. Funct. Anal.* **18**, 369–413.

Segal, I. (1961). Foundations of the theory of infinite-dimensional dynamical systems II. *Canad. J. Math.* **13**, 1–18.

Segal, I. (1962). Mathematical characterization of the physical vacuum for a linear Bose–Einstein field. *Ill. J. Math.* **6**, 500–523.

Sewell, G. L. (1974). Relaxation, amplification and the KMS conditions. *Ann. Phys.* **85**, 336–377.

She, C. Y. and Heffner, H. (1966). Simultaneous measurement of non-commuting observables. *Phys. Rev.* **152**, 1103–1110.

Sinai, Ja G. (1964). Dynamical systems with countably multiple Lebesgue spectrum. *Transl. Amer. Math. Soc.* (2) **39**, 83–110.

Sinha, K. (1972). On the decay of an unstable particle. *Helv. Phys. Acta* **45**, 619–628.

Slawny, J. (1972). On factor representations and the C^*-algebra of canonical commutation relations. *Commun. Math. Phys.* **24**, 151–170.

Srinivas, M. D. (1973). Foundations of a quantum probability theory. *J. Math. Phys.* **16**, 1672–1685.

Steinmann, O. (1968). Particle localization in field theory. *Commun. Math. Phys.* **7**, 112–137.

Stinespring, W. F. (1955). Positive functions on C^*-algebras. *Proc. Amer. Math. Soc.* **6**, 211–216.

Størmer, E. (1963). Positive linear maps of operator algebras. *Acta Math.* **110**, 233–278.

Størmer, E. (1972). On projection maps on von Neumann algebras. *Math. Scand.* **30**, 46–50.

Sudarshan, E. C. G., Mathews, P. M. and Rao, J. (1961). Stochastic dynamics of quantum mechanical systems. *Phys. Rev.* **121**, 920–924.

Sz-Nagy, B. and Foias, C. (1970). *Harmonic analysis of operators on Hilbert space.* North-Holland, Amsterdam, Holland.

Thomas, E. (1970). L'integration par rapport à une measure de Radon vectorielle. *Ann. Inst. Fourier* (Grenoble) **20**, 2, 55–191.

Thomas, L. C. (1974). On the eventual absorption of a quantum particle. *Ann. Inst. Henri Poincaré* A **21**, 81–87.

Tomiyama, J. (1957). On the projection of norm one in W^*-algebras. *Proc. Japan Acad.* **33**, 608–612.

Trotter, H. F. (1959). On the product of semigroups of operators. *Proc. Amer. Math. Soc.* **10**, 545–551.

Umegaki, H. (1954). Conditional expectation in an operator algebra I. *Tohoku Math. J.* **6**, 177–181.

Urbanik, K. (1961). Joint distributions of observables in quantum mechanics. *Studia Math.* **21**, 117–133.

Van Hove, L. (1955). Quantum-mechanical perturbations giving rise to a statistical transport equation. *Physica* **21**, 517–540.

Van Hove, L. (1957). The approach to equilibrium in quantum statistics. *Physica* **23**, 441–480.

Varadarajan, V. S. (1962). Probability in physics and a theorem on simultaneous observability. *Commun. Pure Applied Math.* **15**, 189–217.

Varadarajan, V. S. (1968). *Geometry of quantum theory.* Vol. 1. van Nostrand, England.

Varadarajan, V. S. (1970). *Geometry of quantum theory.* Vol. 2. van Nostrand, England.

Veutsel, A. D. and Freidlin, M. I. (1970). On small random perturbations of dynamical systems. *Russian Math. Surveys* **25**, 1–55.

von Neumann, J. (1955). *Mathematical foundations of quantum mechanics.* Princeton University Press.

Wightman, A. S. (1962). On the localizability of quantum mechanical systems. *Rev. Mod. Phys.* **34**, 845–872.

Wigner, E. P. (1932). On the quantum correction for thermodynamic equilibrium. *Phys. Rev.* (2) **40**, 749–759.

Wigner, E. P. (1959). *Group theory.* Academic Press, London and New York.

Williams, D. N. (1971). Difficulty with a kinematic concept of unstable particles: the Sz-Nagy extension and the Mathews Salam–Zwanziger representation. *Commun. Math. Phys.* **21**, 314–333.

Wright, J. D. M. (1973). Products of positive vector measures. *Quart. J. Math.* (Oxford) (2) **24**, 189–206.

Yosida, K. (1965). Functional analysis. Springer, Berlin.

Zwanzig, R. (1960). Statistical mechanics of irreversibility. *Lectures in Theor. Phys.* **3**, 106–141.

Zwanzig, R. (1964). On the identity of three generalized master equations. *Physica* **30**, 1109–1123.

Index

absorption, 112
abstract state space, 30
analogy, 33
antiunitary map, 21
approximate observable, 39

Baire σ-field, 11
barycentre, 159
beam-foil spectroscopy, 85
Bloch coherent states, 129, 134
Borel σ-field, 11
boson, 80, 117

C*-algebra, 7, 31
canonical commutation relations (CCR), 119
causality, 80
classical state, 120
closed operator, 7
closed system, 26
cocycle, 26
coherent state, 117
complete positivity, 136
completely positive normalised (CPN) semigroup, 143
composite measurement, 15, 38
composition of instruments, 51
conditional expectation, 58
cone, 2
conjugate linear, 21
continuous measure, 13, 59
continuous resolution of identity, 129
contraction semigroup, 103
convex set, 2
convolution, 40, 64, 124
correlation functions, 69, 156
correspondence principle, 123
coset, 61
cosine operator, 44
covariant, 61
cyclic representation, 7
cylinder set, 13

degenerate, 1, 22
dilation, 105
discrete measure, 13, 59
dissipative operator, 103
domain, 7
duality, 4
dynamical semigroup, 27

effect, 19, 31
equilibrium state, 97
expectation, 35
external world, 26, 135
extreme point, 2, 31

fermion, 79
Fock space, 79, 117
Fokker-Planck, 126
formal dimension, 133
free evolution, 117
Friedrichs' model, 155

Gaussian probability density, 122
Gelfand-Naimark-Segal (GNS) theorem, 7
generalised coherent states, 129
generalised master equation, 154, 157
generalised sequence, 5
Gibbs state, 122
Glauber coherent states, 117
G-space, 61

harmonic oscillator, 122
Heisenberg, 43

infinitesimal generator, 8
instrument, 49

interaction rate, 71, 75
invariant order ideal, 86
irreducible process, 86

joint distribution, 16

Langevin equation, 134
lattice, 2, 31
localisability of photon. 68

marginal observable, 38
Markov semigroup, 27, 100
master equation, 154
mean, 35, 40
measure, 11
memory effect, 154
metrisability, 12
minimal unitary dilation, 105
multiplier representation, 132

Naimark's theorem, 131
normal, 6, 18
null set, 89

observable, 36
one parameter semigroup, 8
open system, 26, 149
operation, 17
order ideal, 54
ordering, 2
Ornstein-Uhlenbeck process, 127
orthocomplementation, 19, 31
orthogonal complement, 1

partial trace, 150
phase space, 42
Poincaré cycle, 156
positive definite, 105
positive linear functional, 2
positive linear map, 2
positive-map-valued (PMV) measure, 49
positive-operator-valued (POV) measure, 16, 35
projection-valued measure, 35, 141
process, 86
pure, 2, 21

quantum optics, 69
quantum stochastic process (QSP), 70
question, 19, 31

random walk, 131
rank, 1
recurrent process, 90, 96
reduction of wave packet, 58, 68
repeated measurement, 34
representation, 7, 61
reproducing kernel, 130
reservoir, 149

sample space, 70, 87
second countable, 11
self-adjoint, 1, 8
separable, 12
simple process, 90
simultaneous measurement, 39, 46
sine operator, 44
singular coupling limit, 157
skeleton, 88
spectrum, 1, 7
square integrable representation, 133
stability subgroup, 61
state, 4
state space, 4
Stern-Gerlach, 15
stochastic differential equation, 151, 157
strict, 108
strong convergence, 5, 49
strong repeatability, 54
subordinated semigroup, 132, 134
substantially finite-dimensional, 96
superoperator, 154
superposition principle, 101
support, 1
symmetric operator, 8
symmetry, 33
system of imprimitivity, 62, 68

tensor product, 10
trace class, 2
trace norm, 3
transient process, 96
transition process, 98
transitive, 61
transmission probability, 18

ultraweak convergence, 5
unstable particle, 104, 116

vacuum, 119
value space, 70

variance, 35, 40
von Neumann algebra, 7, 31

weak convergence, 5
weak coupling limit, 155

Weyl algebra, 127, 134
Weyl operator, 44, 118
Wigner distribution, 44
Wigner-Weisskopf atom, 27