

# Quantum Statistical Mechanics

William C. Schieve and Lawrence P. Horwitz



CAMBRIDGE

CAMBRIDGE

[www.cambridge.org/9780521841467](http://www.cambridge.org/9780521841467)

This page intentionally left blank

## QUANTUM STATISTICAL MECHANICS

Many-body theory stands at the foundation of modern quantum statistical mechanics. It is introduced here to graduate students in physics, chemistry, engineering and biology. The book provides a contemporary understanding of irreversibility, particularly in quantum systems. It explains entropy production in quantum kinetic theory and in the master equation formulation of non-equilibrium statistical mechanics.

The first half of the book focuses on the foundations of non-equilibrium statistical mechanics with emphasis on quantum mechanics. The second half of the book contains alternative views of quantum statistical mechanics, and topics of current interest for advanced graduate level study and research.

Uniquely among textbooks on modern quantum statistical mechanics, this work contains a discussion of the fundamental Gleason theorem, presents quantum entanglements in application to quantum computation and the difficulties arising from decoherence, and derives the relativistic generalization of the Boltzmann equation. Applications of statistical mechanics to reservoir ballistic transport are developed.

WILLIAM C. SCHIEVE is Professor Emeritus in the Physics Department and Center for Complex Quantum Systems at the University of Texas, Austin. His research interests lie in non-equilibrium statistical mechanics and its applications to areas such as quantum optics, relativistic statistical mechanics, dynamical models in biophysics, and chaos theory.

LAWRENCE P. HORWITZ is Professor of Physics Emeritus in the School of Physics at Tel Aviv University. He is also Professor of Physics at Bar Ilan University, and Research Director in Theoretical Physics at Ariel University. His research interests lie in the theory of unstable systems, foundations of quantum theory, quantum field theory, particle physics, relativistic mechanics and general relativity, quantum and classical dynamical systems, and chaos theory.



# QUANTUM STATISTICAL MECHANICS

## Perspectives

Professor Emeritus WILLIAM C. SCHIEVE

*Physics Department, University of Texas at Austin, Texas*

*Center for Complex Quantum Systems*

Professor Emeritus LAWRENCE P. HORWITZ

*School of Physics Tel Aviv University Ramat Aviv, Israel*

*Department of Physics Bar Ilan University Ramat Gan, Israel*



CAMBRIDGE  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521841467](http://www.cambridge.org/9780521841467)

© W. Schieve and L. Horwitz 2009

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2009

ISBN-13 978-0-511-50122-7 eBook (Adobe Reader)

ISBN-13 978-0-521-84146-7 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

# Contents

<i>Preface</i>	<i>page xi</i>
<b>1 Foundations of quantum statistical mechanics</b>	<b>1</b>
1.1 The density operator and probability	1
1.2 The Gleason theorem and consequences	6
1.3 Calculation of averages of observables	9
Appendix 1A: Gleason theorem	12
References	18
<b>2 Elementary examples</b>	<b>19</b>
2.1 Introduction	19
2.2 Harmonic oscillator	19
2.3 Spin one-half and two-level atoms	27
Appendix 2A: the Fokker–Planck equation	34
References	35
<b>3 Quantum statistical master equation</b>	<b>37</b>
3.1 Reduced observables	37
3.2 The Pauli equation	39
3.3 The weak coupling master equation for open systems	42
3.4 Pauli equation: time scaling	46
3.5 Reservoir states: rigorous results and models	53
3.6 The completely positive evolution	54
Appendix 3A: Chapman–Kolmogorov master equation	57
References	59
<b>4 Quantum kinetic equations</b>	<b>61</b>
4.1 Introduction	61
4.2 Reduced density matrices and the B.B.G.Y.K. hierarchy	61
4.3 Derivation of the quantum Boltzmann equation	63
4.4 Phase space quantum Boltzmann equation	66
4.5 Memory of initial correlations	76

4.6	Quantum Vlasov equation	79
	Appendix 4A: phase space distribution functions	80
	References	83
<b>5</b>	<b>Quantum irreversibility</b>	85
5.1	Quantum reversibility	85
5.2	Master equations and irreversibility	87
5.3	Time irreversibility of the generalized master and Pauli equations	87
5.4	Irreversibility of the quantum operator Boltzmann equation	89
5.5	Reversibility of the quantum Vlasov equation	90
5.6	Completely positive dynamical semigroup: a model	92
	Appendix 5A: the quantum time reversal operator	94
	References	96
<b>6</b>	<b>Entropy and dissipation: the microscopic theory</b>	98
6.1	Introduction	98
6.2	Macroscopic non-equilibrium thermodynamics	98
6.3	Dissipation and the quantum Boltzmann equation	105
6.4	Negative probability and the quantum $\mathfrak{H}$ theorem	111
6.5	Entropy and master equations	113
	Appendix 6A: quantum recurrence	120
	References	121
<b>7</b>	<b>Global equilibrium: thermostatics and the microcanonical ensemble</b>	123
7.1	Boltzmann's thermostatic entropy	124
7.2	Thermostatics	125
7.3	Canonical and grand canonical distribution of Gibbs	126
7.4	Equilibrium fluctuations	129
7.5	Negative probability in equilibrium	131
7.6	Non-interacting fermions and bosons	132
7.7	Equilibrium limit theorems	136
	References	139
<b>8</b>	<b>Bose–Einstein ideal gas condensation</b>	141
8.1	Introduction	141
8.2	Continuum box model of condensation	142
8.3	Harmonic oscillator trap and condensation	145
8.4	$^4\text{He}$ : the $\lambda$ transition	148
8.5	Fluctuations: comparison of the grand canonical and canonical ensemble	150
8.6	A master equation view of Bose condensation	152
	Appendix 8A: exact treatment of condensate traps	155
	References	158



<b>9</b>	<b>Scaling, renormalization and the Ising model</b>	159
9.1	Introduction	159
9.2	Mean field theory and critical indices	160
9.3	Scaling	167
9.4	Renormalization	169
9.5	Renormalization and scaling	172
9.6	Two-dimensional Ising model renormalization	174
	References	177
<b>10</b>	<b>Relativistic covariant statistical mechanics of many particles</b>	178
10.1	Introduction	178
10.2	Quantum many-particle dynamics: the event picture	180
10.3	Two-event Boltzmann equation	183
10.4	Some results of the quantum event Boltzmann equation	187
10.5	Relativistic quantum equilibrium event ensembles	191
	References	197
<b>11</b>	<b>Quantum optics and damping</b>	199
11.1	Introduction	199
11.2	Atomic damping: atomic master equation	199
11.3	Cavity damping: the micromaser: detection	206
11.4	Detection master equation for the cavity field	207
	Appendix 11A: the field quantization and interaction	214
	References	219
<b>12</b>	<b>Entanglements</b>	221
12.1	Introduction	221
12.2	Entanglements: foundations	221
12.3	Entanglements: Q bits	224
12.4	Entanglement consequences: quantum teleportation, the Bob and Alice story	226
12.5	Entanglement consequences: dense coding	228
12.6	Entanglement consequences: quantum computation	228
12.7	Decoherence: entanglement destruction	231
12.8	Decoherence correction (error correction)	235
	Appendix 12A: entanglement and the Schmidt decomposition	236
	References	238
<b>13</b>	<b>Quantum measurement and irreversibility</b>	240
13.1	Introduction	240
13.2	Ideal quantum measurement	241
13.3	Irreversibility: measurement master equations	243
13.4	An open system master equation model for measurement	246

13.5	Stochastic energy based collapse	248
	References	251
<b>14</b>	<b>Quantum Langevin equation and quantum Brownian motion</b>	<b>253</b>
14.1	Introduction	253
14.2	Quantum Langevin equation	254
14.3	Quantum Langevin equation with measurement	260
	References	262
<b>15</b>	<b>Linear response: fluctuation and dissipation theorems</b>	<b>264</b>
15.1	Introduction	264
15.2	Quantum linear response in the steady state	266
15.3	Linear response, time dependent	269
15.4	Fluctuation and dissipative theorems	272
15.5	Comments and comparisons	277
	References	279
<b>16</b>	<b>Time-dependent quantum Green's functions</b>	<b>281</b>
16.1	Introduction	281
16.2	One- and two-time quantum Green's functions and their properties	282
16.3	Analytic properties of Green's functions	284
16.4	Connection to linear response theory	288
16.5	Green's function hierarchy truncation	289
16.6	Keldysh time-loop path perturbation theory	297
	References	302
<b>17</b>	<b>Decay scattering</b>	<b>303</b>
17.1	Basic notions and the Wigner–Weisskopf theory	303
17.2	Wigner–Weisskopf method: pole approximation	306
17.3	Wigner–Weisskopf method and Lee–Friedrichs model with a single channel	312
17.4	Wigner–Weisskopf and multichannel decay	318
17.5	Wigner–Weisskopf method with many-channel decay: the Lee–Friedrichs model	321
17.6	Gel'fand triple	332
17.7	Lax–Phillips theory	335
17.8	Application to the Stark model	354
	References	362
<b>18</b>	<b>Quantum statistical mechanics, extended</b>	<b>365</b>
18.1	Intrinsic theory of irreversibility	365
18.2	Complex Liouvillian eigenvalue method: introduction	366
18.3	Operators and states with diagonal singularity	367
18.4	Super operators and time evolution	369

18.5	Subdynamics and analytic continuation	371
18.6	The Pauli equation revisited	375
	References	378
<b>19</b>	<b>Quantum transport with tunneling and reservoir ballistic transport</b>	<b>379</b>
19.1	Introduction	379
19.2	Pauli equation and boundary interaction	380
19.3	Ballistic transport	383
19.4	Green's function closed-time path theory to transport	385
	References	389
<b>20</b>	<b>Black hole thermodynamics</b>	<b>390</b>
20.1	Introduction to black holes	390
20.2	Equilibrium thermodynamic analogies: the first law	394
20.3	The second law of thermodynamics and black holes	397
20.4	Extended entropy principle for black holes	399
20.5	Acausal evolution: extended irreversible dynamics in black holes	401
	References	402
<b>A</b>	<b>Problems</b>	<b>404</b>
A.1	Comments on the problems	404
A.2	"Foundations" problems	404
A.3	Kinetic dynamics problems	407
A.4	Equilibrium and phase transition problems	409
	References	410
	<i>Index</i>	411



# Preface

This book had its origin in a graduate course in statistical mechanics given by Professor W. C. Schieve in the Ilya Prigogine Center for Statistical Mechanics at the University of Texas in Austin.

The emphasis is *quantum* non-equilibrium statistical mechanics, which makes the content rather unique and advanced in comparison to other texts. This was motivated by work taking place at the Austin Center, particularly the interaction with Radu Balescu of the Free University of Brussels (where Professor Schieve spent a good deal of time on various occasions). Two Ph.D. candidate theses at Austin, those of Kenneth Hawker and John Middleton, are basic to Chapters 3 and 4, where the master equations and quantum kinetic equations are discussed. The theme there is the dominant and fundamental one of quantum irreversibility. The particular emphasis throughout this book is that of open systems, i.e. quantum systems in interaction with reservoirs and not isolated. A particularly influential work is the book of Professor A. McLennan of Lehigh University, under whose influence Professor Schieve first learned non-equilibrium statistical mechanics.

An account of relatively recent developments, based on the addition in the Schrödinger equation of stochastic fluctuations of the wave function, is given in Chapter 13. These methods have been developed to account for the collapse of the wave function in the process of measurement, but they are deeply connected as well with models for irreversible evolution.

The first six chapters of the present work set forth the theme of our book, particularly extending the entropy principle that was first introduced by Boltzmann, classically. These, with equilibrium quantum applications (Chapters 7, 8, 9 and possibly also Chapters 14 and 15), represent a one-semester advanced course on the subject.

As frequently pointed out in the text, quantum mechanics introduces special problems to statistical mechanics. Even in Chapter 1, written by the coauthor of this work, Professor Lawrence P. Horwitz of Tel Aviv, the idea of a *density operator* is required which is *not* a probability distribution, as in the classical case. The idea of the density operator lies at the very foundations of the quantum theory, providing a description of a quantum state in the most general way. Statistical mechanics requires this full generality. We give a proof of the Gleason theorem, stating that in a Hilbert space of three or more real dimensions, a general quantum state has a representation as a density operator, based on an elegant construction of C. Piron. This structure gives the quantum  $\mathfrak{H}$  theorem, a content which is essentially different from the classical one. This makes the subject surely interesting and important, but difficult.

Quantum entanglements are quite like magic, so to speak. It is necessary and important to see these modern developments; they are described in Chapter 15. This is one chapter that might be used in the extension of the course to a second semester. One- and two-time Green's functions, introduced by Kadanoff and Baym, might be included in the extended treatment, since they are popular but difficult. This is included in Chapter 16 with an application in Chapter 19.

An extension to special relativity is described in Chapter 10. This is a new derivation of a many-body *covariant* kinetic theory. The Boltzmann-like kinetic equation outlined here was derived in collaboration by the authors. The covariant picture is an event dynamics controlled by an abstract time variable first introduced by both Feynman and Stueckelberg and obtains a covariant scalar many-body wave function parameterized by the new time variable. The results of this event picture are outlined in Chapter 10.

Another arena of activity utilizing quantum kinetic equations for open systems is the extensive development in quantum optics. This has been a personal interest of one of the authors (WCS). This interest was a result of a Humboldt Foundation grant to the Max Planck Institute in Munich and later to Ulm, under the direction of Professors Herbert Walther, Marlon Scully and Wolfgang Schleich. The particular area of interest is described in the results outlined in Chapter 11. This material can be included as an introduction to quantum optics in an extended two-semester course.

The idea of spontaneous decay in a quantum system goes back to Gamov in quantum mechanics. This irreversible process seems intrinsic, introducing the notion of the Gel'fand triplet and rigged Hilbert spaces states. The coauthor (LPH) has made personal contributions to this fundamental change in the wave function picture. It is very appropriate to include an extensive discussion of this, which is the content of Chapter 17, describing, among other things, the Wigner–Weisskopf method and the Lax–Phillips approach to enlarging the scope of quantum wave

functions. All of this requires a more advanced mathematical approach than the earlier discussions in this book. However, it is necessary that a well-grounded student of quantum mechanics know these things, as well as acquire the mathematical tools, and therefore it is very appropriate here in a discussion of quantum statistical mechanics.

Chapter 18 is in many ways an extension of Chapter 17. It is an outline of what has been called extended statistical mechanics. Ilya Prigogine and his colleagues in Brussels and Austin, in the past few years, have attempted to formulate many-body dynamics which is intrinsically irreversible. In the classical case this may be termed the complex Liouville eigenvalue method. As an example, the Pauli equation is derived again by these nonperturbative methods. This is not an open-system dynamics but rather, like the previous Chapter 17 discussion, one of closed isolated dynamics. This effort is not finished, and the interested student may look upon this as an introductory challenge.

The final chapter of this book is in many ways a diversion, a topic for personal pleasure. The remarkable objects of our universe known as black holes apparently exist in abundance. These super macroscopic objects obey a simple equilibrium thermodynamics, as first pointed out by Bekenstein and Hawking. Remarkably, the area of a black hole has a similarity to thermodynamic entropy. More remarkable, the  $S$ -matrix quantum field theoretic calculation of Hawking showed that the baryon emission of a black hole follows a Planck formula. Hawking introduced a superscattering operator which is analogous to the extended dynamical theory of Chapter 18.

To complete these comments, we would like to thank Florence Schieve for support and encouragement over these last years of effort on this work. She not only gave passive help but also typed into the computer several drafts of the book as well as communicating with the coauthor and the editorial staff of the publisher. The second coauthor wishes also to thank his wife Ruth for her patience, understanding, and support during the writing of some difficult chapters.

We also acknowledge the help of Annie Harding of the Center here in Austin. Three colleagues at the University of Texas—Tomio Petrosky, George Sudarshan and Arno Bohm—also made valuable technical comments. WCS also thanks the graduate students who, over many years of graduate classes, made enlightened comments on early manuscripts.

We recognize the singular role of Ilya Prigogine in creating an environment in Brussels and Austin in which the study of non-equilibrium statistical mechanics was our primary goal and enthusiasm.

Finally, WCS thanks the Alexander von Humboldt Foundation for making possible extended visits to the Max Planck Institute of Quantum Optics in Garching and later in Ulm. LPH thanks the Center for Statistical Mechanics and Complex

Systems at the University of Texas at Austin for making possible many visits over the years that formed the basis for his collaboration with Professor Schieve, and the Institute for Advanced Study at Princeton, particularly Professor Stephen L. Adler, for hospitality during a series of visits in which, among other things, he learned of the theory of stochastic evolution, and which brought him into proximity with the University of Texas at Austin.



# 1

## Foundations of quantum statistical mechanics

### 1.1 The density operator and probability

Statistical mechanics is concerned with the construction of methods for computing the expected value of observables important for characterizing the properties of physical systems, generally containing many degrees of freedom. Starting with a formally complete detailed description for these many degrees of freedom, probability theory is used to obtain effective procedures. Quantum statistical mechanics makes use of *two* types of probability theory. One of these is the set of natural probabilities associated with the quantum theory which emerges from its structure as a Hilbert space. For example, the Born probability is associated with the square of a wave function. The second is the essentially classical probability associated with an ensemble of separate systems, each with an *a priori* probability assigned by the frequency of occurrence in the ensemble. The quantity which describes both types of probability in an efficient, convenient way is the density operator.

As an example which illustrates many of the basic ideas, consider a beam of particles with spin  $\frac{1}{2}$ . We shall repeat the resulting definitions later in complete generality.

The spin states of these particles are represented by two-dimensional spinors which we denote by the Dirac kets  $|\sigma_z\rangle$  for  $\sigma_z = \pm 1$ , corresponding to the  $z$  component of the spin  $\sigma$  of the particle. If we perform a filtering measurement to select a particle of spin  $\sigma'$  with spin  $\sigma'_z = \pm 1$  in the  $z$  direction, the outcome of the measurement on a beam of particles with spin  $\sigma_z$  is

$$|\langle \sigma'_z | \sigma_z \rangle|^2 = \delta_{\sigma'_z, \sigma_z}.$$

This result can be written as

$$|\langle \sigma'_z | \sigma_z \rangle|^2 = \text{Tr} P_z(\sigma') P_z(\sigma),$$

where the projection operator  $P_z(\sigma) = |\sigma_z\rangle\langle\sigma_z|$  represents the state of the beam of particles with spin  $\sigma$  of definite value  $\sigma_z$ , and the projection operator  $P_z(\sigma')$  represents the experimental question of which value,  $\pm 1$ , this set of particles has.

If we measure instead a different component of spin and, for example, ask for the fraction of particles in the ensemble with spin in the  $\pm x$  direction, the measurement is represented by a projection operator  $P_x(\sigma) = |\sigma_x\rangle\langle\sigma_x|$ , with  $\sigma_x = \pm 1$ . In terms of the eigenvectors of  $\sigma_z$ ,

$$|\sigma_x = \pm 1\rangle = \frac{1}{\sqrt{2}} (|+1\rangle \pm |-1\rangle).$$

It is true (for any of the values of  $\sigma_x$  and  $\sigma_z$ ) that

$$|\langle\sigma_x | \sigma_z\rangle|^2 = \frac{1}{2}.$$

We can write this result as

$$|\langle\sigma_x | \sigma_z\rangle|^2 = \text{Tr}(P_x(\sigma) P_z(\sigma)).$$

Let us now consider a beam of spin  $\frac{1}{2}$  particles with a fraction  $\gamma_+$  with spin up and  $\gamma_-$  with spin down in the  $z$  direction ( $\gamma_+ + \gamma_- = 1$ ). The probability to find spin up as the outcome of the experiment is

$$\begin{aligned} P_+ &= |\langle\sigma'_z = +1 | \sigma_z = +1\rangle|^2 \gamma_+ + |\langle\sigma'_z = +1 | \sigma_z = -1\rangle|^2 \gamma_- \\ &= \gamma_+, \end{aligned}$$

since the second term vanishes. If  $\gamma_+ = \frac{1}{2}$ , the result is indistinguishable from the probability to find a spin  $\pm \frac{1}{2}$  in the  $x$  direction in a beam of particles with definite spin in the  $z$  direction.

We can write the result of the second example as

$$\begin{aligned} P_+ &= \gamma_+ \text{Tr}(P(\sigma'_z = +1) P(\sigma_z = +1)) + \gamma_- \text{Tr}(P(\sigma'_z = +1) P(\sigma_z = -1)) \\ &= \text{Tr}(\rho P(\sigma'_z = +1)) \end{aligned}$$

for

$$\rho \equiv \gamma_+ P(\sigma_z = +1) + \gamma_- P(\sigma_z = -1).$$

The operator  $\rho$  is called the *density operator*, representing a state consisting of a *mixture* of components with spin up and spin down in the ensemble of possibilities. We see that, with a slight generalization of the procedure used above with  $\rho_z \rightarrow \rho_0$ , no matter what direction 0 we test in the experiment, the outcome  $P_0$  (a linear combination of  $\gamma_+$ ,  $\gamma_-$  with coefficients less than unity) can never reach unity if  $\gamma_+$  or  $\gamma_-$  is not unity. In the first example, where we have a beam with definite

$\sigma_z$ , the state is represented by a *vector*, and the measurement of the spin in the  $z$ -direction can yield probability one. For a general choice of  $\gamma_{\pm}$ , there is *no* vector that can represent the state. In the first case the state is called *pure*, and it can be represented by a projection into a one-dimensional subspace (in the previous example,  $P_{\sigma_z} = |\sigma_z\rangle\langle\sigma_z|$ ). This is equivalent to specifying the vector, up to a phase, corresponding to the one-dimensional subspace. In the second case, it is called *mixed* and does not correspond to a vector in the Hilbert space.

It is clear from the discussion of these examples that the *a priori* probabilities  $\gamma_{\pm}$  are essentially classical, reflecting the composition of the beam that was prepared in the macroscopic laboratory.

Although a density operator  $\rho$  of the type that we have defined in this example appears to be a somewhat artificial construction, it is actually a fundamental structure in quantum statistical mechanics (Dirac, 1958). It enables one to study a complex system in the framework of an ensemble and in fact occurs on the most fundamental level of the axioms of the quantum theory.

It was shown by Birkhoff and von Neumann (1936) that both quantum mechanics and classical mechanics can be formulated as the description of a set of questions for which the answer, as a result of experiment, is “yes” or “no.” Such a set, which includes the empty set  $\phi$  (questions that are absurd, e.g. the statement that the system does not exist) and the trivial set  $I$  (the set of all sets, e.g. the statement that the system exists), and is closed with respect to intersections and unions, is called a lattice. A lattice that satisfies the distributive law

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c),$$

where  $\cup$  represents the union and  $\cap$  the intersection, is called Boolean. These operations have the physical meaning of “or” (the symbol  $\cup$ ), in which one or the other of the propositions is true, and “and” (the symbol  $\cap$ ), for which both must be true for the answer of the compound measurement to be “yes.” An example of such a lattice may be constructed in terms of two-dimensional closed regions on a piece of paper. This is discussed again in the appendix to this chapter.

Both classical and quantum theories may be associated with lattices in terms, respectively, of the occupancy of cells in phase space or states in the subspaces of the Hilbert space. The questions  $a$  correspond, in the first case, to the phase space cells (with answer corresponding to occupancy) and in the second to the projection operators  $P_{\alpha}$  associated with a subspace  $M_{\alpha}$ , with the answer corresponding to the values  $\pm 1$  which a projection operator can have. These values correspond to evaluating the projection operator on vectors which lie within or outside the subspace.

Birkhoff and von Neumann asserted that the fundamental difference between classical and quantum mechanics is that the lattices corresponding to classical

mechanics are Boolean, and those corresponding to quantum mechanics *are not*. The non-Boolean structure of the quantum lattice is associated with the lack of commutativity of the projection operators associated with different subspaces:

$$a \cap (b \cup c) \neq (a \cap b) \cup (a \cap c). \quad (1.1)$$

This is a fundamental difference between classical and quantum statistics.

Let us illustrate this point by a simple example, again using the spin  $\frac{1}{2}$  system. Each of the Pauli spin matrices has eigenvalues  $\pm 1$  and is therefore associated with a set of projection operators of the form

$$P_i = \frac{1}{2} (1 \pm \sigma_i)$$

for  $i = x, y, z$ . Let us consider three closed linear subspaces associated with the projections into the subspaces with the  $\sigma_i$  positive, i.e. with the  $P_i$  defined as above with positive signs. We call these subspaces  $M_x, M_y, M_z$ ; they correspond to propositions which are not compatible, i.e. the corresponding projection operators do not commute. We shall show explicitly, for this simple example, that

$$M_z \cap (M_x \cup M_y) \neq (M_z \cap M_x) \cup (M_z \cap M_y),$$

that is, this set of propositions is not Boolean. The construction is interesting in that it illustrates the special structure of the topology of Hilbert spaces as well as the notion of the non-Boolean lattice.

We start by constructing the union of the manifolds  $M_x$  and  $M_y$  by their joint linear span. Taking the standard definition of the Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

the projection operators into the subspaces with positive eigenvalues are

$$\begin{aligned} P_x &= \frac{1}{2} (1 + \sigma_x) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ P_y &= \frac{1}{2} (1 + \sigma_y) = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \\ P_z &= \frac{1}{2} (1 + \sigma_z) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

The corresponding eigenvectors are given by projecting a generic vector  $v$  into the respective subspaces. For

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

using the result just given,

$$P_x v = \frac{1}{2} (v_1 + v_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

so that  $M_x$  is represented by the linear span of the normalized eigenvector:

$$v_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Similarly,

$$P_y v = \frac{1}{2} (v_1 - i v_2) \begin{pmatrix} 1 \\ i \end{pmatrix},$$

so that the corresponding (normalized) eigenvector is

$$v_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

Finally,

$$P_z v = v_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

so the corresponding eigenvector is

$$v_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The union of the subspaces  $M_x$  and  $M_y$  is the closed linear span of vectors in both subspaces. By taking the combination  $v_x + i v_y$ , it is easy to see that the vector  $v_z$  (and hence the subspace  $M_z$ ) is contained in  $M_x \cup M_y$ . To construct the distributed operation

$$(M_z \cap M_x) \cup (M_z \cap M_y),$$

we must use the construction for which the projection operator corresponding to the intersection of two noncompatible subspaces is generated by an alternating succession of projections into the two subspaces (Jauch, 1968). The products  $P_z P_x$  and  $P_z P_y$  are, it so happens, idempotents up to coefficients less than one, i.e.

$$P_z P_x = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

$$(P_z P_x)^2 = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

and

$$P_z P_y = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}$$

$$(P_z P_y)^2 = \frac{1}{4} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix},$$

which implies that both  $(P_z P_x)^n$  and  $(P_z P_y)^n$  go to zero as  $n \rightarrow \infty$ . Therefore,

$$M_z \cap M_x = M_z \cap M_y = 0.$$

Clearly,

$$M_z \cap (M_x \cup M_y) \neq (M_z \cap M_x) \cup (M_z \cap M_y).$$

Although  $P_z P_x$  and  $P_z P_y$  are not zero (the two corresponding vectors are not orthogonal), the closed subspace that is common is empty. One can think of this geometrically in terms of two lines that have some projection on the other, but the intersection of the two lines is just a point of zero measure. Physically, this implies that we cannot have a definite statement of the joint values of  $\sigma_z$  and  $\sigma_x$  or  $\sigma_y$ . The noncommutativity of the associated projections is essential; if they were commutative, the product of projections would be a projection, and the products would not converge to zero. It is clear from this example that compatible subspaces would satisfy Boolean distributivity.

We shall later discuss the Wigner function, which appears to provide joint distributions over noncommutative variables such as  $q$  and  $p$ ; however, these functions are not probabilities, since, although they are the coefficients of what might be called the Weyl basis for the operator algebra of the quantum theory which appear in expectation values, they are not positive (Wigner, 1936).

## 1.2 The Gleason theorem and consequences

The axioms of quantum mechanics are implicitly developed in the fundamental work of Dirac (1958). Let us focus here on probability. Given  $P_i$  ( $i = 1, \dots$ ), a sequence of projections  $P_i P_k = 0$  for  $i \neq k$ , then the probability measure  $w$

$$w : P \rightarrow [0, 1]$$

satisfies

$$\begin{aligned}
 (a) \quad & \cup_i w(P_i) = w\left(\sum_i P_i\right) \\
 (b) \quad & w(\phi) = 0, \quad w(I) = 1 \\
 & (\phi \text{ is the zero projection}) \\
 (c) \quad & w(P) = w(F) = 1 \rightarrow w(P \cap F) = 1
 \end{aligned} \tag{1.2}$$

Piron (1976) added another axiom, namely that partially ordered (by inclusion) sets of the non-Boolean lattice of the quantum theory form Boolean sublattices, and with this he was able to show a converse result, i.e. that such partially ordered lattices can be embedded in a Hilbert space (or a family of Hilbert spaces if there are superselection rules), thus inducing the full structure of the quantum theory.

Along with the sets of “yes-no” questions that form the basic elements  $a$  of the quantum lattice, one may assume a function  $w(a)$  with values between zero and unity, with the interpretation of a probability measure, which has the so-called sigma additivity property

$$w(a \cup b) = w(a) + w(b) \tag{1.3}$$

when  $a$  and  $b$  have no intersection, i.e.  $a \cap b = \phi$ . This idea is consistent with the notion of probability for the “yes” answer for  $a$  and  $b$ . Gleason (1957) showed that for any Hilbert space of three or more real dimensions, there is a density operator, self-adjoint and positive,  $\rho$ , such that

$$w(a) = \text{Tr} \rho P_a, \tag{1.4}$$

where  $P_a$  is the projection operator into a subspace corresponding to the question  $a$ . This existence theorem is one of the most powerful and important theorems in the foundations of the statistical quantum theory. The function  $w(a)$  is called a *state*, a notion completely consistent with Dirac’s definition of a state in the quantum theory, i.e. for any  $a$ , this function provides the probability of its truth and therefore corresponds to maximum knowledge.

The original proof of Gleason is rather long and involved, but Piron has given a simple and elegant proof, which is given here in an appendix to this chapter for the mature student.

The density operator (often called “density matrix”) has the properties

$$\begin{aligned}
 \text{Tr} \rho &= 1 \\
 \text{Tr} \rho^2 &\leq 1.
 \end{aligned} \tag{1.5}$$

The first follows from the fact that the sum over all disjoint  $a$  of  $w(a)$  is the total probability measure on the set of all questions (and the sum over all disjoint  $P_a$  is

the unit operator). The second follows from the first; all eigenvalues of  $\rho$  are real and positive with values less than or equal to unity. With these properties, one can prove that the spectrum of  $\rho$  must be completely discrete.

Mackey (1963) has given a converse theorem. If the function  $w(a)$  can reach the value unity on a one-dimensional subspace of the Hilbert space, the corresponding density operator is just a projection into this one-dimensional subspace and can be put into correspondence (up to a phase) with the vector of the Hilbert space generating this one-dimensional subspace. Such a state is called *pure*. A state which cannot reach the value of unity on any one-dimensional subspace is called *mixed*.

The proof is very simple. Let  $P_0$  be the projection onto a one-dimensional subspace generated by the vector  $\phi_0$ , and let us use the representation, taking into account the discrete spectrum of  $\rho$ ,

$$\rho = \sum_i \gamma_i |\psi_i\rangle \langle \psi_i|. \quad (1.6)$$

Here we use the Dirac ket  $|\psi_i\rangle$  to signify an element of the Hilbert space. Then if  $\text{Tr} \rho P_0 = 1$ , it follows that

$$\text{Tr} \rho (1 - P_0) = 0,$$

or

$$\text{Tr} \sum_i \gamma_i \langle \psi_i | (1 - P_0) | \psi_i \rangle = \text{Tr} \sum_i \gamma_i \|(1 - P_0) | \psi_i \rangle\|^2 = 0,$$

where  $\| \chi \|^2$  is defined as  $\langle \chi | \chi \rangle$ , the norm of the vector  $|\chi\rangle$ . Since the  $\gamma_i$  are positive, this implies that

$$(1 - P_0) | \psi_i \rangle = 0$$

for all of the  $|\psi_i\rangle$ , i.e.,

$$|\psi_i\rangle = \lambda_i |\phi_0\rangle$$

for all  $i$ . Substituting into Eq. (1.6), we see that in this case we must have

$$\rho = \sum_i \gamma_i |\lambda_i|^2 |\phi_0\rangle \langle \phi_0|.$$

Furthermore, if the  $|\psi_i\rangle$  and  $|\phi_0\rangle$  are normalized,  $|\lambda_i|^2 = 1$ . Then, by Eq. (1.5) and Eq. (1.6) (for the  $|\psi_i\rangle$  orthogonal), one sees that the sum of the  $\gamma_i$  is unity; hence

$$\rho = |\phi_0\rangle \langle \phi_0|,$$

which is the projection operator into the subspace generated by  $|\phi_0\rangle$ . This theorem therefore identifies the pure states with vectors of the Hilbert space, and it is for this reason that one often calls the vectors of the Hilbert space “states.” Every vector in the Hilbert space corresponds to a pure state.



If  $w_1$  and  $w_2$  are two different states, then

$$w = \lambda_1 w_1 + \lambda_2 w_2$$

with  $\lambda_1 + \lambda_2 = 1$  and with  $\lambda_1, \lambda_2$  positive also is a state; the set of states form a convex set (Jauch, 1968). Such a state is called a *mixture*. A state which cannot be represented in terms of two others is called *pure*; the pure states are the extremal subset of a convex set. These definitions are, of course, consistent with Mackey's result.

### 1.3 Calculation of averages of observables

Let us now consider an observable represented by a self-adjoint operator  $A$  on the Hilbert space with a spectrum of discrete eigenvalues  $a_k$ . Such an operator can be represented as a sum over projections into its eigenstates, i.e.

$$A = \sum_k a_k P_k, \quad (1.7)$$

where, if  $P_k = |\phi_k\rangle\langle\phi_k|$  and the  $|\phi_k\rangle$  form a normalized orthogonal set, we clearly have

$$A |\phi_k\rangle = a_k |\phi_k\rangle.$$

The expectation of this operator in some pure state represented by  $|\psi_i\rangle$  is then

$$\begin{aligned} \langle\psi_i| A |\psi_i\rangle &= \sum_k a_k \langle\psi_i| P_k |\psi_i\rangle \\ &= \sum_k a_k |\langle\psi_i | \phi_k\rangle|^2, \end{aligned} \quad (1.8)$$

with the usual quantum interpretation that  $|\langle\psi_i | \phi_k\rangle|^2$  is the quantum mechanical probability that a system in the state described by  $|\phi_k\rangle$  is found in the state  $|\psi_i\rangle$ . The weighting of the eigenvalues of  $A$  by this probability then gives the expected value of this observable in the state described by  $|\psi_i\rangle$ . Suppose now that we prepare a system which contains subsystems in the states  $|\psi_i\rangle$  according to the *a priori* probability distribution  $\gamma_i$ . This can be arranged by preparing a system with the number of subsystems in each state  $|\psi_i\rangle$  proportional to the  $\gamma_i$ . This is an ensemble. We emphasize here that this step, as in our previous example, is entirely *classical*. We build an ensemble of subsystems with *a priori* probabilities based on their frequency of occurrence, a completely classical notion of probability, i.e. the frequency interpretation.

The overall expectation of the value of the observable  $A$  is then given by the sum over all of the expected values in each of the quantum states, with coefficients

equal to the classical probabilities of the occurrence of each quantum state in the ensemble, i.e.

$$\langle A \rangle = \sum_i \gamma_i \langle \psi_i | A | \psi_i \rangle.$$

This result is obtained directly by computing

$$\langle A \rangle = \text{Tr} \rho A, \quad (1.9)$$

where

$$\rho = \sum_i \gamma_i |\psi_i\rangle \langle \psi_i|. \quad (1.10)$$

Viewing this in a slightly different way, we see that

$$\langle A \rangle = \sum_k a_k \text{Tr} (\rho P_k), \quad (1.11)$$

where

$$\begin{aligned} \text{Tr} (\rho P_k) &= \sum_i \gamma_i \langle \psi_i | P_k | \psi_i \rangle \\ &= \sum_i \gamma_i |\langle \psi_i | \phi_k \rangle|^2 \end{aligned} \quad (1.12)$$

is the probability of finding the system in the subspace associated with  $P_k$ . This probability is composed of two types of expectation: the quantum probability to find the  $P_k$  in each state  $\psi_i$ , and the classical probability for the occurrence of the state  $\psi_i$  (determined by the relative number of subsystems in that state).

The results that we have given can easily be extended to the most general case of an observable with both discrete and continuous spectra without change in the formal structure, although as we shall see later, there are special technical aspects that arise in the continuous case (for example, in scattering theory). To see this, we use the spectral representation theory of von Neumann. It was shown by von Neumann (1955) that every self-adjoint operator  $A$ , corresponding to a physical observable, has a spectral representation of the form

$$A = \int a \, dE(a), \quad (1.13)$$

where  $a$  takes on a continuous set of values (the real line), and the self-adjoint set of operators  $E(a)$  is called a “spectral family.” It satisfies the property

$$E(a) E(b) = E(\min(a, b)), \quad (1.14)$$

with  $E(-\infty) = 0$  and  $E(\infty) = I$ . It easily follows from these properties that

$$dE(a) dE(b) = \begin{cases} dE(a), & \text{if } a = b; \\ 0, & \text{otherwise} \end{cases} \quad (1.15)$$

where  $a$  and  $b$  now refer to names given to infinitesimal intervals along the line (i.e. for  $\Delta a$  small,  $dE(a) = E(a + \Delta a) - E(a)$ ). The integral Eq. (1.13) is considered to be of Stieltjes–Lebesgue type, in the sense that if the weight function  $\langle \psi | dE(a) | \psi \rangle = d \|E(a) | \psi \rangle\|^2$  has a jump discontinuity at some point  $a_0$ , the integral is evaluated as the difference between the values of  $\|E(a) | \psi \rangle\|^2$  above and below the point  $a_0$ . If, in particular,  $d \|E(a) | \psi \rangle\|^2$  is zero in the neighborhood of the point  $a_0$  (except at the point itself), so that the jump is isolated, one obtains a contribution to any expectation value of  $A$  just from the point  $a = a_0$  (in this neighborhood). The coefficient, since  $E(a)^2 = E(a)$ , is  $\langle \psi | E(a_0 + \varepsilon) - E(a_0 - \varepsilon) | \psi \rangle$ , where  $\varepsilon$  is infinitesimal. The operator  $E(a_0 + \varepsilon) - E(a_0 - \varepsilon)$  may then be identified with one of the discrete projection operators appearing in Eq. (1.7). Hence, the representation Eq. (1.11) includes both discrete and continuous spectra. In Eq. (1.8) one then uses

$$\langle \psi_i | A | \psi_i \rangle = \int ad \|E(a) | \psi_i \rangle\|^2,$$

and Eq. (1.9) remains valid quite generally.

We now turn to time evolution, which is the central issue of this book. The quantum states  $\psi_i$  from which the density operator is constructed evolve under Schrödinger evolution as

$$i\hbar \frac{\partial}{\partial t} |\psi_i\rangle = H |\psi_i\rangle. \quad (1.16)$$

It follows simply that for  $\rho$  of the form of Eq. (1.10), acting with the time derivative on both factors  $|\psi_i\rangle$  and  $\langle \psi_i|$ , using Eq. (1.16) and its conjugate, we see that

$$\frac{d\rho}{dt} = i\hbar (\rho H - H\rho) = i\hbar [\rho, H], \quad (1.17)$$

a time evolution similar to the evolution of a Heisenberg operator but with opposite sign.

Eq. (1.17) forms the basis for the description of the dynamical evolution of a system in statistical mechanics, the analog of the classical Liouville equation (Tolman, 1938). Since the Schrödinger equation is reversible in time, this evolution is reversible (Farquahar, 1964). Under such an evolution, a pure state remains pure, and a mixed state does not change its character (this follows from the fact that the change in time of  $\text{Tr}\rho^2$ , given by  $2i\hbar\text{Tr}(\rho[\rho, H])$ , vanishes). We shall discuss in later chapters evolution given by, for instance, master equations, the Pauli equation and the Lindblad equation, describing *irreversible* processes. Such equations

can describe the evolution of a density matrix for a pure state into a density matrix corresponding to a mixed state. (For this more general evolution,  $\text{Tr}(\rho\dot{\rho})$  does not vanish.)

Although, as we have previously emphasized, the density operator might appear to be a somewhat artificial construction, combining both classical and quantum probability notions to achieve an overall expectation value, it actually arises on the most fundamental level of the quantum theory. Methods for the construction and study of this operator and its time evolution are the essential goal of the techniques of statistical mechanics; the theory is constructed on this basic foundation.

Good general references to the topics of this chapter are the books of Tolman (1938), Dirac (1958), Farquahar (1964), Landau and Lifshitz (1970), Balescu (1975), Dvurecenskij (1993), and Huang (1987). Extensive pertinent references are given at the ends of later chapters.

### **Appendix 1A: Gleason theorem**

The Gleason theorem (Gleason, 1957) is concerned with the calculation of the probability  $w$  of obtaining the answer “yes” as a result of carrying out an experiment which is an ideal measurement of the first kind on a system in some given state. In working out the proof of this theorem, we shall follow closely the presentation given by C. Piron (1976).

To study and prove the result, we shall need some definitions already implicit in previous sections.

The logical propositions of the quantum theory correspond to equivalence classes of questions  $\{\beta\}$  which are realized in terms of measurements. A question  $\beta$  is called a measurement of the first kind if, every time the answer is “yes,” the proposition  $b$ , corresponding to the equivalence class defined by  $\{\beta\}$ , is true immediately after the measurement. (Measurement will be taken up again in Chapter 13.)

A question  $\beta$  is said to be ideal if every proposition  $b$  defined by such a  $\beta$ , which is true beforehand, is again true afterwards when the response of the system is “yes.”

We shall assume that the probability  $w$  is the same for every question  $\beta$  defining the proposition  $b$ , for  $\beta$  (or  $\beta^{\sim}$ , its complement) is an ideal measurement of the first kind. We may then denote this probability by  $w(p, b)$ , where  $p$  is the initial state in which the experiment is carried out, and  $b$  is the proposition defined by the equivalence class  $\{\beta\}$ .

The Gleason theorem applies to the construction of the function  $w$  in the framework of a Hilbert space, on which the operators of the quantum theory are

represented. The closed subspaces of a Hilbert space, with their associated projection operators, form a set subject to the operations of intersection and union, and contain the empty set and the set of all subsets, i.e. a structure called a lattice, isomorphic to the lattice of propositions (Birkhoff and von Neumann, 1936; Birkhoff, 1961; Piron, 1976), as mentioned earlier. For an irreducible proposition system, in which there is only one minimal proposition (no superselection rules), every self-adjoint operator corresponds to an observable. Let  $P(H)$  be such a Hilbert realization.

We now state the Gleason theorem (Gleason, 1957) (see Piron, 1976, for the general case of a family of Hilbert spaces, for which there is a nontrivial set of minimal propositions):

**Theorem:** *Given a propositional system  $L = P(H)$ , where  $H$  is a Hilbert space (of dimension  $\geq 3$ ) over the reals, complex numbers or quaternions, there exists a unique function  $w(p, b)$  defined on the atoms  $p$  (corresponding to the one-dimensional subspaces of  $H$ ) and the propositions  $b$  of  $L$  which satisfies (as in Eq. (1.2) and Eq. (1.3))*

$$\begin{aligned} (1) \quad & 0 \leq w(p, b) \leq 1 \\ (2) \quad & p \subset b \Leftrightarrow w(p, b) = 1 \\ (3) \quad & b \perp c \Rightarrow w(p, b) + w(p, c) = w(p, b \cup c). \end{aligned} \tag{1A.1}$$

We begin the proof by noting that there is a vector  $f_p$  in  $H$ , associated with the atom  $p$ , satisfying

$$\langle f_p | f_p \rangle = \|f_p\|^2 = 1.$$

Each proposition  $b$  in  $P(H)$  can be represented by a projection operator  $Q$  into a linear closed subspace of  $H$ . Then

$$w(p, b) = \langle f_p | Q | f_p \rangle$$

satisfies the conditions of the theorem.

Our principal task is then to show uniqueness. If there were another function  $w(p, b)$  satisfying these conditions, it would have to have a different value on some pair  $p, b$ . For such functions, there would be another proposition  $q$  (an atom) for which, in this case,  $w(p, q)$  has a different value. However, if the function were unique, the value would necessarily be the same. Such a  $q$  can be constructed as follows. Note that

$$[(p \cup b') \cap b] \cup (p' \cap b) = b$$

and that, since  $p$  and  $p'$  are orthogonal,

$$[(p \cup b') \cap b] \perp p' \cap b.$$

However,  $w(p, p' \cap b) = 0$ , so

$$q = (p \cup b') \cap b \quad (1A.2)$$

for an atom. The other function would, by construction, have a different value for  $w(p, q)$ . We choose the two vectors  $f_p$  and  $f_q$  in such a way that  $\langle f_p | f_q \rangle$  is real. We may then consider just three vectors associated with the atoms  $p, q$ , i.e.  $f_p, f_q$  and a vector (real) orthogonal to these. The restriction of  $w(p, b)$  to the three-dimensional real Hilbert subspace generated by  $f_p, f_q$  and a third vector orthogonal to these still satisfies the conditions of the theorem. To complete the proof, it is then sufficient to prove the uniqueness of  $w$  in the case of the real three-dimensional Hilbert space ( $R^3$ ). This construction, therefore, has the minimum dimension necessary to carry out a proof of uniqueness.

To carry out the proof, let us assume  $p$  in  $w(p, b)$  to be fixed. The lattice of subspaces of  $R^3$  is then the points and lines of the projective plane realized as the intersection of  $R^3$  with the tangent plane at  $p$  to the unit sphere. In the same way as the complex plane is mapped onto the unit sphere including the point at infinity, we are considering the plane as a (projective) representation of the sphere of unit vectors in  $R^3$ . (It may be helpful for the reader to draw his own diagrams for the construction described here.)

We seek a unique function  $w(q)$ , where we drop reference to  $p$ , now fixed, defined at the points  $q$  of the plane which has the value 1 at  $p$  and 0 at the point(s) at infinity.

If  $q$  lies on some arbitrary line  $L$  in the plane, then  $w(q)$  takes on a maximal value at a point  $q_0$  where the line  $pq_0$  is perpendicular to the line  $L$ . This follows from the fact that if  $q$  is a point on  $L$ , and  $q'$  is its orthogonal complement on  $L$ ,  $q \cup q'$  on the line is just  $q_0$ . Hence, by (3) of Eq. (1A.1),

$$\begin{aligned} w(q) + w(q') &= w(q_0) \\ \text{or } w(q_0) &\geq w(q). \end{aligned}$$

We now note that  $w(q)$  decreases along the line  $L$ . To see this, consider a point at  $q$  and a line  $L_q$  perpendicular to  $pq$ . Move along this line to  $q_1$ ; we know by the foregoing argument that

$$w(q) \geq w(q_1).$$

Now erect a line at  $q_1$  perpendicular to  $pq_1$  and move to a point on this new line,  $r$ . Clearly,

$$w(q_1) \geq w(r).$$

Now put another line at this point  $r$ , and connect it back to  $L_q$  at the point  $q_2$ . Since

$$w(r) \geq w(q_2)$$

along  $L_q$ , it follows that

$$w(q) \geq w(q_1) \geq w(q_2), \quad (1A.3)$$

forming a decreasing sequence.

We prove now the first lemma of four leading to the uniqueness of the function  $w(p, q)$ . The method we follow is to prove each lemma making some crucial assumptions, and each succeeding lemma proves those assumptions. In the fourth lemma the proof is complete.

**Lemma 1:** *If the value of  $w(p, q)$  depends only on the angle  $\theta$  between the rays  $p$  and  $q$ , then it is unique and given by*

$$w(q) = \cos^2 \theta. \quad (1A.4)$$

To prove this lemma, we work as before in the plane tangent to  $R^3$  at the point  $p$  and erect another point  $q$  at a “distance”  $\lambda$  (corresponding to the square of the actual distance), say, below  $p$ . We then erect another point  $q'$  at an equal distance  $\lambda$  from  $p$ , labeling the midpoint of the line  $qq'$  by  $q_1$ . By the rules of ordinary geometry, the line  $pq_1$  is orthogonal to the line  $qq'$ ; it is the closest point on that line to  $p$ . It then follows from our previous arguments ( $q'$  is the orthogonal complement of  $q$  on this line) that

$$w(q') + w(q) = w(q_1).$$

But the angles  $q'q_1$  and  $q_1q$  are equal, and by the assumptions of our lemma, it then follows that

$$2w(q) = w(q_1).$$

There is a line  $L_q$ , perpendicular to  $pq$  at a point  $r$ , passing through  $q'$ , and a right triangle that can be constructed from  $r$  to the apex  $q_2$  to  $q$ , with the line  $rpq$  as hypotenuse. To satisfy Pythagoras’s theorem, we see that the distance  $pr$  is  $\frac{1}{\lambda}$ .  $pq_2$  is unity (this line is orthogonal to  $qp$ ). The distance  $qq_2$  is  $1 + \lambda$ , and the distance  $r q_2$  is  $1 + \frac{1}{\lambda}$ . Finally,  $q'r$  is  $\lambda - \frac{1}{\lambda}$ . Now we denote the total length of  $q'q$  as  $2y$  (this line is bisected by  $q_1$ ). Again, by Pythagoras, the length of  $qr$  is  $1 + \lambda + 1 + \frac{1}{\lambda}$ . Adding this to  $q'r$ , which is  $\lambda - \frac{1}{\lambda}$ , we find the simple result that  $4y = 2(1 + \lambda)$ . Finally, using the fact that  $pq$  has length (squared)  $\lambda$ , the length of  $pq_1$ , which we call  $z$ , is

$$z = \lambda - y = \lambda - \frac{1}{2}(1 + \lambda) = \frac{1}{2}(\lambda - 1).$$

We now rewrite the relation previously obtained,  $2w(q) = w(q_1)$ , as

$$2w(\lambda) = w\left(\frac{1}{2}(\lambda - 1)\right)$$

for  $\lambda > 1$ . Since by our construction,  $r \perp q$ ,

$$w(\lambda) + w\left(\frac{1}{\lambda}\right) = w(p) = 1,$$

we have that

$$1 - w(\lambda) = w\left(\frac{1}{\lambda}\right).$$

If we now define

$$x = (1 + \lambda)^{-1} = \cos^2 \theta,$$

the rest of the demonstration follows by simple algebra.

Since  $\lambda = \frac{1}{x} - 1$ , by defining

$$f(x) = w(\lambda) = w\left(\frac{1-x}{x}\right),$$

one easily finds that

$$2f(x) = f(2x) \tag{1A.5}$$

for  $0 \leq x \leq \frac{1}{2}$  (i.e.  $\lambda > 1$ ), and for a second relation,

$$1 - f(x) = f(1 - x). \tag{1A.6}$$

To see this, set  $y = \frac{2}{\lambda+1} = 2x$ ; then, using the definition,

$$f(y) = w\left(\frac{1-y}{y}\right) = w\left(\frac{1}{2}(\lambda - 1)\right) = 2w(\lambda),$$

it follows that  $f(y) = f(2x) = 2f(x)$ .

The second relation follows from the fact that

$$f(1 - x) = w\left(\frac{x}{1-x}\right) = w\left(\frac{1}{\lambda}\right),$$

so that  $1 - f(x) = f(1 - x)$ , for  $0 \leq x \leq 1$ .

The identification  $f(x) = x$  with  $x = \cos^2 \theta$  for some  $\theta$  satisfies both these relations and satisfies the statement of the lemma. To see that this solution is the only solution which increases, we may expand both sides of the equation  $2f(x) = f(2x)$  in Taylor series about  $x = 0$ . The condition  $f(0) = 0$  follows from the requirement that  $w \rightarrow 0$  at  $\infty$ ; it follows that all derivatives equal to or higher than second order must vanish, and the function must therefore be linear. Substituting  $f(x) = \alpha x$  into the second relation, Eq. (1A.6), we see that  $1 - \alpha x = \alpha(1 - x)$  so that  $\alpha$  must be unity. The solution is therefore unique.

We now prove one of the assumptions of Lemma 1.



**Lemma 2:** *If  $w(q)$  is continuous, then its value depends only on the angle between the rays  $p$  and  $q$ .*

The remaining two lemmas (lemmas 3 and 4) prove continuity.

To prove this lemma, let  $q$  and  $r$  be two points on the projective plane situated at the same distance from  $p$ . To prove that  $w(q) = w(r)$ , we start by proving that for any  $q_0 \in qp$  sufficiently close to  $q$ , the signs of  $w(q_0) - w(r)$  and  $\lambda - \lambda_0$ , where  $\lambda$  and  $\lambda_0$  are the distances  $pq$  and  $pq_0$  respectively, are the same. If  $\lambda > \lambda_0$ , we can join  $q_0$  to  $r$  by a sequence  $q_0, q_1, q_2, \dots$  of sequentially perpendicular steps, since at each step  $\lambda_1 \geq \lambda_0, \lambda_2 \geq \lambda_1, \dots$  up to  $r$ , which reaches  $\lambda$ , by construction (note that we started with  $\lambda_0 < \lambda$ ). Then

$$w(q_0) \geq w(q_1) \geq w(q_2) \geq \dots \geq w(r), \quad (1A.7)$$

since the lengths increase at every step. But we can take  $q_0$  arbitrarily close to  $q$ . The same set of inequalities can be established in the other direction, starting with a point  $r_0$  on  $pr$ , and hence  $w(q) = w(r)$ ; i.e. the value of  $w(q)$  depends only on the distance between  $p$  and  $q$  (the angle).

**Lemma 3:** *If  $w(q)$  is continuous at some point  $q_0$ , then it is continuous at every point.*

We first show that if  $w(q)$  is continuous at  $q_0$ , it is continuous at each point  $q_1$  orthogonal to  $q_0$ . Then  $q_0$  and  $q_1$  lie symmetrically on both sides of the point of a line from  $p$  perpendicular to  $q_0q_1$ . Denote an  $\varepsilon$  neighborhood of  $q_0$  by  $U$ , and take a point  $q'$  on the line  $q_0q_1$  in  $U$ ; further, consider the point  $q$  on the line  $q_0q_1$  orthogonal to  $q'$ . As we have done before, we use the relations

$$\begin{aligned} w(q) + w(q') &= w(q_0) + w(q_1) \\ w(r_0) + w(r_1) &= w(r') + w(q_0), \end{aligned}$$

where  $r_0, r_1$  and  $r'$  are defined in a similar way on a line passing at some angle through  $q$ , for which  $q$  and  $r'$  are orthogonal and  $r_0 \in U$  and  $r_1$  are orthogonal. It follows from these relations that

$$\begin{aligned} |w(r_1) - w(q_1)| &= |w(q_0) - w(r_0) + w(r') - w(q')| \\ &= |w(q_0) - w(r_0) + w(r') - w(q_0) + w(q_0) - w(q')| \\ &\leq |w(q_0) - w(r_0)| + |w(r') - w(q_0)| + |w(q_0) - w(q')| \\ &\leq 3\varepsilon, \end{aligned}$$

where we have used the bounding inequalities between the relation between the  $w(q)$ 's and the distances. Our construction, furthermore, requires  $r', q' \in U_{q_0}$ . The subset  $r_0 \ni r_1 \in U$  then forms an  $\varepsilon$  neighborhood of  $q_1$  and is therefore

continuous at  $q_1$ . We finally note that there always exists a point  $q^\perp$  perpendicular to two arbitrary points  $q', r'$ .

**Lemma 4:** *The function  $w(q)$  is continuous at some point  $q_0$ .*

On a line  $L$  through  $p$ ,  $w(q)$  is a decreasing function of  $\lambda$  (distance from  $p$ ). A decreasing bounded function is continuous almost everywhere. Hence  $w(q)$  is continuous on  $L$  at some point  $q_0$ . Finally, if  $w(q_2) - w(q_1) < \varepsilon$ , then  $|w(q) - w(q_0)| < \varepsilon$  at every point in the triangle formed by  $rr'q_1$  (all points in this triangle are farther away from  $p$  than the distance  $\lambda$  at  $q_2$ , in the  $\varepsilon$  neighborhood of  $q_0$ ).

This completes the lemmas for the proof of the Gleason theorem, in general.

## References

- Balescu, R. (1975). *Equilibrium and Non-equilibrium Statistical Mechanics* (New York, John Wiley), revised 1999 as *Matter out of Equilibrium* (London, Imperial College Press).
- Birkhoff, G. (1961). *Lattice Theory* (Providence, American Mathematical Society).
- Birkhoff, G. and von Neumann, J. (1936). *Ann. Math.* **37**, 823.
- Dirac, P. A. M. (1958). *Quantum Mechanics*, 4th edn. (London, Oxford University Press).
- Dvurecenskij, A. (1993). *Gleason's Theorem and Its Applications* (Dordrecht, Kluwer).
- Farquahar, I. E. (1964). *Ergodic Theory in Statistical Mechanics* (London, Interscience).
- Gleason, A. M. (1957). *J. Math. Mech.* **6**, 885.
- Huang, K. (1987). *Statistical Mechanics*, 2nd edn. (New York, John Wiley).
- Jauch, J. M. (1968). *Foundations of Quantum Mechanics* (Reading, Addison-Wesley).
- Landau, L. D. and Lifshitz, E. M. (1970). *Statistical Mechanics* (Reading, Addison-Wesley).
- Mackey, G. W. (1963). *Mathematical Foundations of Quantum Mechanics* (Reading, Benjamin).
- Piron, C. (1976). *Foundations of Quantum Physics* (Reading, Benjamin).
- Tolman, R. C. (1938). *The Principles of Statistical Mechanics* (London, Oxford).
- von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics* (Princeton, Princeton University Press).
- Wigner, E. (1936). *Phys. Rev.* **40**, 749.

## 2

### Elementary examples

#### 2.1 Introduction

Now we will turn to some elementary and familiar examples of quantum mechanics to remind us of matters which will be used in the subsequent discussions. The focus will be the harmonic oscillator and also the two-level atom and spin  $\frac{1}{2}$  systems (Dirac, 1958; Louisell, 1973; Cohen-Tannoudji *et al.*, 1977; Jordan, 1986; Liboff, 1998).

#### 2.2 Harmonic oscillator

The Hamiltonian operator is

$$\hat{H} = \frac{1}{2} (\hat{p}^2 + \omega^2 \hat{q}^2) = \hat{H}^\dagger. \quad (2.1)$$

The classical equations of motion are

$$\begin{aligned} \frac{dq}{dt} &= \frac{\partial H}{\partial p} = p \\ \frac{dp}{dt} &= -\frac{\partial H}{\partial q} = -\omega^2 q. \end{aligned} \quad (2.2)$$

In quantum mechanics,

$$[\hat{q}, \hat{p}] = i\hbar. \quad (2.3)$$

The “hat” denotes operator.

The time-dependent Heisenberg equations are of the same form as the classical counterpart:

$$\begin{aligned} \frac{d\hat{q}(t)}{dt} &= \hat{p}(t) \\ \frac{d\hat{p}(t)}{dt} &= -\omega^2 \hat{q}(t). \end{aligned} \quad (2.4)$$

This is generally true in one dimension, where we have

$$\begin{aligned}\frac{d\hat{q}(t)}{dt} &= \frac{1}{i} \left[ \hat{q}(t), \hat{H}(\hat{p}(t), \hat{q}(t)) \right] = \frac{\partial \hat{H}(t)}{\partial \hat{p}(t)} \\ \frac{d\hat{p}(t)}{dt} &= \frac{1}{i} \left[ \hat{p}(t), \hat{H}(\hat{p}(t), \hat{q}(t)) \right] = -\frac{\partial \hat{H}(t)}{\partial \hat{q}(t)},\end{aligned}$$

where  $\hat{H}(\hat{p}(t), \hat{q}(t))$  is the Heisenberg Hamiltonian operator. This, of course, is the classical correspondence rule

$$\{A, B\} \rightarrow \frac{1}{\hbar i} [A, B]$$

$$[\hat{q}(t), \hat{p}(t)] = i\hbar,$$

where the Heisenberg operators  $\hat{q}(t)$ ,  $\hat{p}(t)$  are related to the Schrödinger  $\hat{q}$ ,  $\hat{p}$  by

$$\begin{aligned}\hat{q}(t) &= U^\dagger(t, 0) \hat{q} U(t, 0) \\ \hat{p}(t) &= U^\dagger(t, 0) \hat{p} U(t, 0).\end{aligned}\tag{2.5}$$

Here  $U(t) = \exp(-i\hat{H}t)$ ,  $\hbar = 1$ . Utilizing this, we obtain the solutions to Eq. (2.4):

$$\begin{aligned}\hat{q}(t) &= \hat{q} \cos \omega t + \frac{\hat{p}}{\omega} \sin \omega t \\ \hat{p}(t) &= -\omega \hat{q} \sin \omega t + \hat{p} \cos \omega t.\end{aligned}\tag{2.6}$$

These operator equations have exactly the same form as the solutions to the classical equations. For this reason, this is one of the few cases in which an exact Heisenberg operator solution may be obtained. It is easily shown that the time-dependent commutation laws follow.

The Schrödinger equation is

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle.\tag{2.7}$$

In this “picture” the operators,  $\hat{H}$  etc., are time independent. From this the von Neumann equation for  $\hat{\rho}(t)$  is obtained (see the previous chapter):

$$i \frac{d\hat{\rho}}{dt} = [\hat{H}, \hat{\rho}]\tag{2.8}$$

Keep in mind that we are working in the Schrödinger picture. For the harmonic oscillator,

$$\begin{aligned}\psi(t) &= \exp(-iHt) |\psi(0)\rangle = U(t, 0) |\psi(0)\rangle \\ &= -i \left[ \cos \hat{H}t + i \sin \hat{H}(t) \right] |\psi(0)\rangle.\end{aligned}\tag{2.9}$$

To reduce this further, let us introduce the well-known creation ( $a^\dagger$ ) and annihilation ( $a$ ) operators. (Both are non-Hermitian.)

$$\hat{a} = \frac{1}{\sqrt{2\omega}} (\omega \hat{q} + i \hat{p}) \quad (2.10)$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2\omega}} (\omega \hat{q} - i \hat{p}) \quad (2.11)$$

From the commutation law, Eq. (2.3), we obtain

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (2.12)$$

Also important are

$$\begin{aligned} [\hat{a}, \hat{a}^\dagger \hat{a}] &= \hat{a} \\ [\hat{a}^\dagger, \hat{a}^\dagger \hat{a}] &= -\hat{a}^\dagger. \end{aligned} \quad (2.13)$$

In this representation,

$$\hat{H} = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right). \quad (2.14)$$

These relations are true in the Heisenberg as well as the Schrödinger picture.

Now, for the harmonic oscillator,

$$U(t, 0) = \exp(-i\omega \hat{a}^\dagger \hat{a} t) \exp\left(\frac{-i\omega t}{2}\right).$$

Let us introduce the number representation

$$\hat{N} |n\rangle = n |n\rangle, \quad (2.15)$$

equivalent to the energy representation

$$\hat{H} |E\rangle = E |E\rangle$$

$$\hat{N} = \hat{a}^\dagger \hat{a} = \hat{N}^\dagger.$$

From Eq. (2.13),

$$\begin{aligned} aN - Na &= a \\ a^\dagger N - Na^\dagger &= a^\dagger. \end{aligned} \quad (2.16)$$

With these raising and lowering operators, we may construct a complete set of states (Dirac, 1958). For normalized states we have

$$\begin{aligned}
 \hat{N} |n\rangle &= n |n\rangle \quad n \text{ integer and positive} \\
 \langle n | n' \rangle &= \delta_{nn'} \\
 \hat{a}^\dagger |n\rangle &= \sqrt{n+1} |n+1\rangle \\
 \hat{a} |n\rangle &= \sqrt{n} |n-1\rangle \\
 \hat{a} |0\rangle &= 0 \\
 |n\rangle &= \frac{\hat{a}^{\dagger n} |0\rangle}{\sqrt{n!}}
 \end{aligned} \tag{2.17}$$

and completeness

$$\sum_{n=0}^{\infty} |n\rangle \langle n| = I.$$

The energy is

$$E_n = \omega \left( n + \frac{1}{2} \right).$$

In the number states, the harmonic oscillator von Neumann equation is

$$\begin{aligned}
 i \dot{\rho}_{nn'} &= (E_n - E_{n'}) \rho_{nn'} \\
 &= \omega (n - n') \rho_{nn'}.
 \end{aligned}$$

The solution is simply

$$\rho_{nn'}(t) = \exp - (i\omega (n - n') t) \rho_{nn'}(0). \tag{2.18}$$

The diagonal and off-diagonal elements are uncoupled. Diagonal elements are constant, and the off-diagonal elements oscillate, and

$$\sum_n \rho_{nn}(t) = \sum_n \rho_{nn}(0) = 1. \tag{2.19}$$

In the so-called random phase approximation, we replace  $\rho_{nn'}(t)$  by its average over  $n - n'$ . Then the oscillations cancel, and  $\bar{\rho}_{nn'}(t) = \rho_{nn'}(0)$  is time independent. The comments made are also true for any exact diagonal representation, not just the harmonic oscillator being discussed here. We may write the coordinate representation  $u_n(q)$ . From

$$a |0\rangle = 0 = (q + ip) |0\rangle,$$

we have

$$\left( \omega q' + \frac{d}{dq'} \right) u_0(q') = 0, \tag{2.20}$$

whose normalized solution is the Gaussian

$$u_0(q) \equiv \langle q | 0 \rangle = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp\left(\frac{-\omega q^2}{2}\right). \quad (2.21)$$

The time-dependent solution is

$$u_0(q, t) = \exp\left(-i\frac{\omega}{2}t\right) u_0(q).$$

It is easily seen that the ground state is a minimum uncertainty state  $\Delta q \Delta p = \frac{1}{2}\hbar$ .

Let us now consider the coherent state representation. We introduce the non-Hermitian eigenvalue problem,

$$a |\alpha\rangle = \alpha |\alpha\rangle. \quad (2.22)$$

The eigenvalues are not real, nor are they orthogonal.

To solve this, we use the completeness of the number representation  $|\alpha\rangle = \sum_{n=0}^{\infty} c_n(\alpha) |n\rangle$ . Next, we form

$$a |\alpha\rangle = \sum_{n=1}^{\infty} c_n(\alpha) \sqrt{n} |n-1\rangle = \sum_{n=0}^{\infty} \alpha c_n(\alpha) |n\rangle \quad (2.23)$$

and shift indices  $n \rightarrow n+1$ . Take the scalar product with  $|m\rangle$ . We obtain the recursion relation

$$c_{n+1}(\alpha) \sqrt{n+1} = \alpha c_n(\alpha). \quad (2.24)$$

This gives

$$c_n(\alpha) = \frac{\alpha^n}{\sqrt{n!}} c_0.$$

Thus,

$$|\alpha\rangle = c_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle.$$

It is easy to show

$$|\langle n | \alpha \rangle|^2 = \frac{\alpha^{2n} \exp\left(-\frac{\alpha^2}{2}\right)}{n!},$$

a Poisson distribution. From this  $\langle n \rangle = \alpha^* \alpha$ , and

$$\frac{\langle (n - \langle n \rangle)^2 \rangle^{\frac{1}{2}}}{\langle n \rangle} = \frac{1}{|\alpha|} = \frac{1}{\langle n \rangle^{\frac{1}{2}}}.$$

We take  $\langle \alpha | \alpha \rangle = 1$  and obtain

$$\langle \alpha | \alpha \rangle = |c_0|^2 \exp |\alpha|^2,$$

so

$$|\alpha\rangle = \exp \frac{(-|\alpha|^2)}{2} \exp \alpha \hat{a}^\dagger \exp (-\alpha^* \hat{a}) |0\rangle, \quad (2.25)$$

taking  $\alpha$  to be complex. The completeness relation is

$$\int d^2\alpha |\alpha\rangle \langle \alpha| = 1 = \sum_0^\infty |n\rangle \langle n|, \quad (2.26)$$

where  $d^2\alpha = r dr d\theta$ , and the non-orthogonality is seen by

$$|\langle \beta | \alpha \rangle|^2 = \exp (-|\alpha - \beta|^2). \quad (2.27)$$

The expansion in terms of coherent states is not unique (Nussenzweig, 1973). They are overcomplete and non-orthogonal. In spite of this, one may expand an arbitrary vector in Hilbert space in terms of them. If we assume that the expansion is an entire function,  $f(\alpha\alpha^*)$ , of the complex  $\alpha$  plane, then the representation is unique.

We may show

$$\langle q \rangle = \sqrt{\frac{1}{2\omega}} (\alpha + \alpha^*) \quad (2.28)$$

$$\langle p \rangle = i\sqrt{\frac{\omega}{2}} (\alpha^* - \alpha)$$

$$\langle q^2 \rangle = \frac{1}{2\omega} (\alpha^{*2} + \alpha^2 + 2\alpha^* \alpha + 1) \quad (2.29)$$

$$\langle p^2 \rangle = \frac{-\omega}{2} (\alpha^{*2} + \alpha^2 - 2\alpha^* \alpha - 1).$$

Thus,  $\Delta p \Delta q = \frac{1}{2}$ , since  $(\Delta q)^2 = \frac{1}{2\omega}$  and  $(\Delta p)^2 = \frac{\omega}{2}$ . All the coherent states are minimum uncertainty. They are quasi-classical. We may obtain  $\langle q | \alpha \rangle$  to verify this. It is the generalized Gaussian

$$\langle q | \alpha \rangle = \left(\frac{\omega}{\pi}\right)^{\frac{1}{4}} \exp \left[ \frac{-\omega}{2} (q - \langle \hat{q} \rangle)^2 + i \langle \hat{p} \rangle q + iu \right], \quad (2.30)$$

where  $u$  is an arbitrary phase and as above,

$$\begin{aligned} \langle \Delta q \rangle^2 &= \frac{1}{2\omega} \\ \langle \Delta p \rangle^2 &= \frac{\omega}{2}. \end{aligned}$$



Now we introduce the first example met here of a phase space distribution function,  $P(\alpha\alpha^*, t)$ , of Glauber (1963) and Sudarshan (1963). Here the “phase space” is  $\alpha, \alpha^*$ . Now

$$\int d^2\alpha P(\alpha\alpha^*, t) = 1. \quad (2.31)$$

$P(\alpha\alpha^*)$  is a “diagonal” representation of the density operator in coherent states

$$\rho = \int d^2\alpha P(\alpha\alpha^*) |\alpha\rangle \langle\alpha|.$$

It has the important property

$$\text{tr} \hat{\rho} \hat{O} = \langle O(\hat{a}, \hat{a}^\dagger) \rangle = \int d^2\alpha O_{cl}(\alpha\alpha^*) P(\alpha\alpha^*). \quad (2.32)$$

Quantum averages are calculated quasi-classically. There is a correspondence rule, the normal ordering rule. In  $\hat{O}$  the  $\hat{a}$  is placed to the right of the  $\hat{a}^\dagger$ . For instance, by commutation,  $aa^\dagger \rightarrow a^\dagger a + 1$ . Phase space distribution functions, such as the Wigner function, will be discussed in greater detail in subsequent chapters. We must remark  $P(\alpha\alpha^*, t) \not\equiv 0$ . It is real and normalizable. Let

$$P(\alpha\alpha^*, t) = \text{tr} \rho(t) \delta(\alpha^* - a^\dagger) \delta(\alpha - a). \quad (2.33)$$

This is a somewhat sophisticated statement because of the operator  $\delta$  functions. Utilizing this definition and the von Neumann equation, we may write for the harmonic oscillator

$$i \frac{\partial P}{\partial t} = \text{Tr} [\rho(t) [\delta(\alpha^\dagger - a^*) \delta(\alpha - a), \omega a^\dagger a]].$$

We will evaluate this in the appendix to this chapter. We obtain a Fokker–Planck equation for  $P(\alpha\alpha^*, t)$  (Gardiner, 1991).

$$\frac{\partial P(\alpha\alpha^*, t)}{\partial t} = i\omega \left[ \alpha \frac{\partial P}{\partial \alpha} - \alpha^* \frac{\partial P}{\partial \alpha^*} \right]. \quad (2.34)$$

It is a first-order partial differential equation in  $t, \alpha, \alpha^*$ . The general solution may be obtained from the characteristic equations

$$dt = \frac{d\alpha}{-i\omega\alpha} = \frac{d\alpha^*}{i\omega\alpha^*}, \quad (2.35)$$

which are the “Hamilton equations” of the  $\alpha, \alpha^*$  “phase space.” The solution is

$$\begin{aligned} \alpha(t) &= \alpha_0 \exp(-i\omega t) \\ \alpha^*(t) &= \alpha_0^* \exp(i\omega t). \end{aligned} \quad (2.36)$$

The general solution is an arbitrary function  $f(\alpha(t), \alpha^*(t))$ . If the initial value is Gaussian in  $\alpha$ , i.e.

$$P(\alpha, \alpha^*, 0) = N \exp(-|\alpha - \alpha_0|^2),$$

then

$$P(\alpha, \alpha^*, t) = N \exp(-|\alpha(t) - \alpha_0|^2).$$

For

$$P(\alpha\alpha^*, t) = \delta^2(\alpha(t) - \alpha_0),$$

the coherent state propagates in time as  $\exp i\omega t$ . This was first seen by Schrödinger (1926).

Let us consider an extension of the harmonic oscillator by including a damping term. A particularly simple example is the phase damped oscillator with the interaction

$$V = \Gamma a^\dagger a + \Gamma^\dagger a^\dagger a \quad (2.37)$$

(Walls and Milburn, 1985; Gardiner, 1991). The von Neumann equation may be written

$$\dot{\rho} = -i\omega [a^\dagger a, \rho] + \frac{1}{2}K (\bar{N} + 1) 2a^\dagger a \rho a^\dagger a - (a^\dagger a)^2 \rho - \rho (a^\dagger a)^2. \quad (2.38)$$

This is the Lindblad form and is discussed in detail in Chapters 4, 5 and 6. Here  $\bar{N} = \frac{1}{\exp(\frac{\omega}{kT}) - 1}$ , and  $K$  is a damping constant. In the number representation,

$$\langle n | \dot{\rho} | m \rangle = \left\{ -i\omega (n - m) - \frac{1}{2}K (2\bar{N} + 1) (n - m)^2 \right\} \langle n | \rho | m \rangle.$$

The diagonal and off-diagonal elements  $\langle n | \rho | m \rangle$  are still uncoupled. The solution is immediate:

$$\langle n | \rho(t) | m \rangle = \exp(-i\omega (n - m)t) \exp - \left[ (2\bar{N} + 1)K (n - m)^2 \frac{t}{2} \right] \langle n | \rho(0) | m \rangle.$$

The off-diagonal elements decay as  $(n - m)^2 K (2\bar{N} + 1)$  to the constant diagonal initial state  $\langle n | \rho(0) | m \rangle$ . More will be said of this in the discussion of decoherence in Chapter 12.

To obtain the equation for  $P(\alpha)$ , we use the operator correspondence discussed in the appendix:

$$\begin{aligned} a\rho &\rightarrow \alpha P(\alpha\alpha^*) \\ a^\dagger\rho &\rightarrow \left(\alpha^* - \frac{\partial}{\partial\alpha}\right) P(\alpha\alpha^*) \\ \rho a &\rightarrow \left(\alpha - \frac{\partial}{\partial\alpha^*}\right) P(\alpha\alpha^*) \\ \rho a^\dagger &\rightarrow \alpha^* P(\alpha\alpha^*) \end{aligned} \quad (2.39)$$

to obtain the Fokker–Planck equation,

$$\frac{\partial P}{\partial t} = \left\{ \frac{1}{2}K \left( \frac{\partial}{\partial\alpha}\alpha + \frac{\partial}{\partial\alpha^*}\alpha^* \right) - i\omega \left( \frac{\partial}{\partial\alpha}\alpha - \frac{\partial}{\partial\alpha^*}\alpha^* \right) + K\bar{N} \frac{\partial^2}{\partial\alpha\partial\alpha^*} \right\} P. \quad (2.40)$$

By introducing  $\alpha = x + iy$  (Scully and Zubairy, 1997), we find the average:

$$\langle\alpha(t)\rangle = \alpha(0) \exp \left[ - \left( \frac{K}{2} \right) - i\omega \right] t. \quad (2.41)$$

In the coherent state, we obtain a classical damped oscillator solution.

$P(\alpha\alpha^*, t)$  need not be positive. If it is, then the state of the system is classical,  $P(\alpha\alpha^*)$  being a true probability distribution.  $P(\alpha\alpha^*)$  may exist for nonclassical or truly quantum states. However, if  $\alpha = x + iy$ , we obtain a Fokker–Planck equation in  $x, y$  with positive diffusion coefficient, so  $P(\alpha\alpha^*, t) > 0$ .

### 2.3 Spin one-half and two-level atoms

The spin of the electron is

$$\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma} \quad \text{Let } \hbar = 1 \quad (2.42)$$

(Cohen-Tannoudji *et al.*, 1977).  $\boldsymbol{\sigma}$  obeys  $\mathbf{m}_S = -\frac{e}{2\mu}\boldsymbol{\sigma}$ , and  $\mathbf{m}_S$  is the spin magnetic moment.  $\sigma_j$  has the properties

$$\begin{aligned} [\sigma_i, \sigma_j]_- &= 2i\sigma_k \\ i, j &= 1, 2, 3. \end{aligned} \quad (2.43)$$

These are angular momentum commutation laws for half integer  $\mathbf{l}$ . Now  $\sigma_i^2 = 1$ , so

$$\sigma_i\sigma_j = i\sigma_k. \quad (2.44)$$