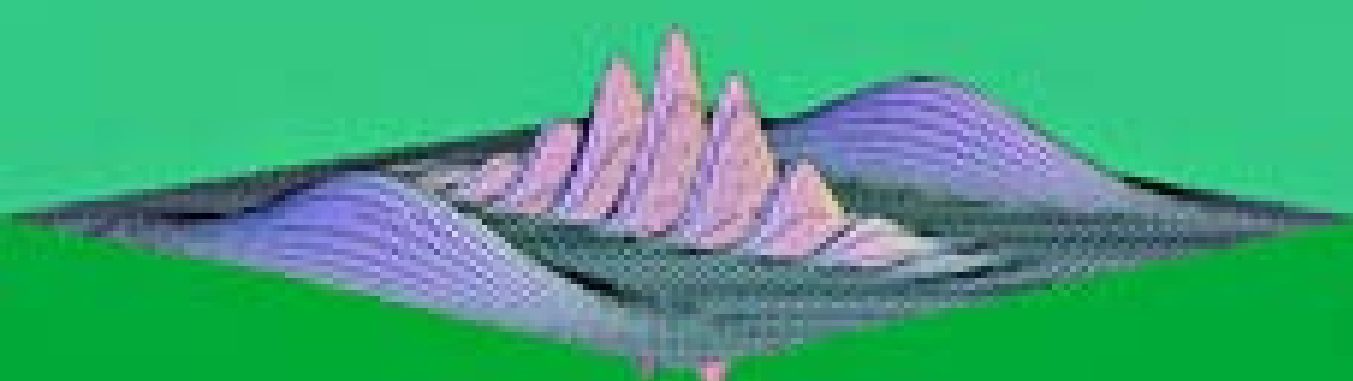
Christopher C. Gerry and Peter L. Knight

Introductory Quantum Optics



CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521820356

This page intentionally left blank

Introductory Quantum Optics

This book provides an elementary introduction to the subject of quantum optics, the study of the quantum-mechanical nature of light and its interaction with matter.

The presentation is almost entirely concerned with the quantized electromagnetic field. Topics covered include single-mode field quantization in a cavity, quantization of multimode fields, quantum phase, coherent states, quasi-probability distribution in phase space, atom–field interactions, the Jaynes–Cummings model, quantum coherence theory, beam splitters and interferometers, nonclassical field states with squeezing etc., tests of local realism with entangled photons from down-conversion, experimental realizations of cavity quantum electrodynamics, trapped ions, decoherence, and some applications to quantum information processing, particularly quantum cryptography. The book contains many homework problems and a comprehensive bibliography.

This text is designed for upper-level undergraduates taking courses in quantum optics who have already taken a course in quantum mechanics, and for first- and second-year graduate students.

A solutions manual is available to instructors via solutions@cambridge.org.

Christopher Gerry is Professor of Physics at Lehman College, City University of New York. He was one of the first to exploit the use of group theoretical methods in quantum optics and is also a frequent contributor to *Physical Review A*. In 1992 he co-authored, with A. Inomata and H. Kuratsuji, *Path Integrals and Coherent States for Su (2) and SU (1, 1)*.

Peter Knight is a leading figure in quantum optics, and in addition to being President of the Optical Society of America in 2004, he is a Fellow of the Royal Society. In 1983 he co-authored *Concepts of Quantum Optics* with L. Allen. He is currently Head of the Physics Department of Imperial College and Chief Scientific Advisor at the UK National Physical Laboratory.

Introductory Quantum Optics

Christopher Gerry
Lehman College, City University of New York

Peter Knight
Imperial College London and UK National Physical Laboratory

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

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

www.cambridge.org
Information on this title: www.cambridge.org/9780521820356

© C. C. Gerry and P. L. Knight 2005

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 2004

ISBN-13 978-0-511-22949-7 eBook (EBL)
ISBN-10 0-511-22949-6 eBook (EBL)

ISBN-13 978-0-521-82035-6 hardback
ISBN-10 0-521-82035-9 hardback

ISBN-13 978-0-521-52735-4 paperback
ISBN-10 0-521-52735-X paperback

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.

C. C. G. dedicates this book to his son, Eric.
P. L. K. dedicates this book to his wife Chris.

Contents

Acknowledgements

This book developed out of courses that we have given over the years at Imperial College London, and the Graduate Center of the City University of New York, and we are grateful to the many students who have sat through our lectures and acted as guinea pigs for the material we have presented here.

We would like to thank our many colleagues who, over many years have given us advice, ideas and encouragement. We particularly thank Dr. Simon Capelin at Cambridge University Press who has had much more confidence than us that this would ever be completed. Over the years we have benefited from many discussions with our colleagues, especially Les Allen, Gabriel Barton, Janos Bergou, Keith Burnett, Vladimir Buzek, Richard Campos, Bryan Dalton, Joseph Eberly, Rainer Grobe, Edwin Hach III, Robert Hilborn, Mark Hillery, Ed Hinds, Rodney Loudon, Peter Milonni, Bill Munro, Geoffrey New, Edwin Power, George Series, Wolfgang Schleich, Bruce Shore, Carlos Stroud Jr, Stuart Swain, Dan Walls and Krzysztof Wodkiewicz. We especially thank Adil Benmoussa for creating all the figures for this book using his expertise with Mathematica, Corel Draw, and Origin Graphics, for working through the homework problems, and for catching many errors in various drafts of the manuscript. We also thank Mrs. Ellen Calkins for typing the initial draft of several of the chapters.

Our former students and postdocs, who have taught us much, and have gone on to become leaders themselves in this exciting subject: especially Stephen Barnett, Almut Beige, Artur Ekert, Barry Garraway, Christoph Keitel, Myungshik Kim, Gerard Milburn, Martin Plenio, Barry Sanders, Stefan Scheel, and Vlatko Vedral: they will recognize much that is here!

As this book is intended as an introduction to quantum optics, we have not attempted to be comprehensive in our citations. We apologize to authors whose work is not cited.

C. C. G. wishes to thank the members of the Lehman College Department of Physics and Astronomy, and many other members of the Lehman College community, for their encouragement during the writing of this book.

P. L. K. would like especially to acknowledge the support throughout of Chris Knight, who has patiently provided encouragement, chauffeuring and vast amounts of tea during the writing of this book.

Our work in quantum optics over the past four decades has been funded by many sources: for P. L. K. in particular the UK SRC, SERC, EPSRC, the Royal Society, The European Union, the Nuffield Foundation, and the U. S. Army are thanked for their support; for C. C. G. the National Science Foundation, The Research Corporation, Professional Staff Congress of the City University of New York (PSC-CUNY).

Chapter 1
Introduction

1.1 Scope and aims of this book

Quantum optics is one of the liveliest fields in physics at present. While it has been a dominant research field for at least two decades, with much graduate activity, in the past few years it has started to impact the undergraduate curriculum. This book developed from courses we have taught to final year undergraduates and beginning graduate students at Imperial College London and City University of New York. There are plenty of good research monographs in this field, but we felt that there was a genuine need for a straightforward account for senior undergraduates and beginning postgraduates, which stresses basic concepts. This is a field which attracts the brightest students at present, in part because of the extraordinary progress in the field (e.g. the implementation of teleportation, quantum cryptography, Schrödinger cat states, Bell violations of local realism and the like). We hope that this book provides an accessible introduction to this exciting subject.

Our aim was to write an elementary book on the essentials of quantum optics directed to an audience of upper-level undergraduates, assumed to have suffered through a course in quantum mechanics, and for first- or second-year graduate students interested in eventually pursuing research in this area. The material we introduce is not simple, and will be a challenge for undergraduates and beginning graduate students, but we have tried to use the most straightforward approaches. Nevertheless, there are parts of the text that the reader will find more challenging than others. The problems at the end of each chapter similarly have a range of difficulty. The presentation is almost entirely concerned with the quantized electromagnetic field and its effects on atoms, and how nonclassical light behaves. One aim of this book is to connect quantum optics with the newly developing subject of quantum information processing.

Topics covered are: single-mode field quantization in a cavity, quantization of multimode fields, the issue of the quantum phase, coherent states, quasi-probability distributions in phase space, atom–field interactions, the Jaynes–Cummings model, quantum coherence theory, beam splitters and interferometers, nonclassical field states with squeezing, etc., test of local realism with entangled

photons from down-conversion, experimental realizations of cavity quantum electrodynamics, trapped ions, etc., issues regarding decoherence, and some applications to quantum information processing, particularly quantum cryptography. The book includes many homework problems for each chapter and bibliographies for further reading. Many of the problems involve computational work, some more extensively than others.

1.2 History

In this chapter we briefly survey the historical development of our ideas of optics and photons. A detailed account can be found in the "Historical Introduction" for example in the 6th edition of Born and Wolf. A most readable account of the development of quantum ideas can be found in a recent book by Whitaker [1]. A recent article by A. Muthukrishnan, M. O. Scully and M. S. Zubairy [2] ably surveys the historical development of our ideas on light and photons in a most readable manner.

The ancient world already was wrestling with the nature of light as rays. By the seventeenth century the two rival concepts of waves and corpuscles were well established. Maxwell, in the second half of the nineteenth century, laid the foundations of modern field theory, with a detailed account of light as electromagnetic waves and at that point classical physics seemed triumphant, with "minor" worries about the nature of black-body radiation and of the photoelectric effect. These of course were the seeds of the quantum revolution. Planck, an inherently conservative theorist, was led rather reluctantly, it seems, to propose that thermal radiation was emitted and absorbed in discrete quanta in order to explain the spectra of thermal bodies. It was Einstein who generalized this idea so that these new quanta represented the light itself rather than the processes of absorption and emission, and was able to describe how matter and radiation could come into equilibrium (introducing on the way the idea of stimulated emission), and how the photoelectric effect could be explained. By 1913, Bohr applied the basic idea of quantization to atomic dynamics and was able to predict the positions of atomic spectral lines.

Gilbert Lewis, a chemist, coined the word photon well after the light quanta idea itself was introduced. In 1926 Lewis said

> It would seem appropriate to speak of one of these hypothetical entities as a particle of light, a corpuscle of light, a light quantum, or light quant, if we are to assume that it spends only a minute fraction of its existence as a carrier of radiant energy, while the rest of the time it remains as an important structural element within the atom . . . I therefore take the liberty of proposing for this hypothetical new atom, which is not light but plays an important part in every process of radiation, the name photon [3].

Clearly Lewis's idea and ours are rather distantly connected!

De Broglie in a remarkable leap of imagination generalized what we knew about light quanta, exhibiting wave and particle properties to matter itself. Heisenberg, Schrödinger and Dirac laid the foundations of quantum mechanics in an amazingly short period from 1925 to 1926. They gave us the whole machinery we still use: representations, quantum-state evolution, unitary transformations, perturbation theory and more. The intrinsic probabilistic nature of quantum mechanics was uncovered by Max Born, who proposed the idea of probability amplitudes which allowed a fully quantum treatment of interference.

Fermi and Dirac, pioneers of quantum mechanics, were also among the first to address the question of how quantized light interacts with atomic sources and propagates. Fermi's *Reviews of Modern Physics* article in the 1930s, based on lectures he gave in Ann Arbor, summarize what was known at that time within the context of nonrelativistic quantum electrodynamics in the Coulomb gauge. His treatment of interference (especially Lipmann fringes) still repays reading today. It is useful to quote Willis Lamb in this context:

> Begin by deciding how much of the universe needs to be brought into the discussion. Decide what normal modes are needed for an adequate treatment. Decide how to model the light sources and work out how they drive the system [4].

This statement sums up the approach we will take throughout this book.

Weisskopf and Wigner applied the newly developed ideas of non-relativistic quantum mechanics to the dynamics of spontaneous emission and resonance fluorescence, predicting the exponential law for excited-state decay. This work already exhibited the self-energy problems, which were to plague quantum electrodynamics for the next 20 years until the development of the renormalization programme by Schwinger, Feynman, Tomonaga, and Dyson. The observation of the anomalous magnetic moment of the electron by Kusch, and of radiative level shifts of atoms by Lamb and Retherford, were the highlights of this era. The interested reader will find the history of this period very ably described by Schweber in his magisterial account of QED [5]. This period of research demonstrated the importance of considering the vacuum as a field which had observable consequences. In a remarkable development in the late 1940s, triggered by the observation that colloids were more stable than expected from considerations of van der Waals interactions, Casimir showed that long-range intermolecular forces were intrinsically quantum electrodynamic. He linked them to the idea of zero-point motion of the field and showed that metal plates in vacuum attract as a consequence of such zero-point motion.

Einstein had continued his study of the basic nature of quantum mechanics and in 1935 in a remarkable paper with Podolsky and Rosen was able to show how peculiar quantum correlations were. The ideas in this paper were to explode into one of the most active parts of modern physics with the development by Bohm and Bell of concrete predictions of the nature of these correlations; this laid the

foundations of what was to become the new subject of quantum information processing.

Optical coherence had been investigated for many years using amplitude interference: a first-order correlation. Hanbury Brown and Twiss in the 1950s worked on intensity correlations as a tool in stellar interferometry, and showed how thermal photon detection events were "bunched." This led to the development of the theory of photon statistics and photon counting and to the beginnings of quantum optics as a separate subject. At the same time as ideas of photon statistics were being developed, researchers had begun to investigate coherence in light–matter interactions. Radio-frequency spectroscopy had already been initiated with atomic beams with the work of Rabi, Ramsey and others. Sensitive optical pumping probes of light interaction with atoms were developed in the 1950s and 1960s by Kastler, Brossel, Series, Dodd and others.

By the early 1950s, Townes and his group, and Basov and Prokhorov, had developed molecular microwave sources of radiation: the new masers, based on precise initial state preparation, population inversion and stimulated emission. Ed Jaynes in the 1950s played a major role in studies of whether quantization played a role in maser operation (and this set the stage for much later work on fully quantized atom–field coupling in what became known as the Jaynes–Cummings model). Extending the maser idea to the optical regime and the development of lasers of course revolutionized modern physics and technology.

Glauber, Wolf, Sudarshan, Mandel, Klauder and many others developed a quantum theory of coherence based on coherent states and photodetection. Coherent states allowed us to describe the behaviour of light in phase space, using the quasi-probabilities developed much earlier by Wigner and others.

For several years after the development of the laser there were no tuneable sources: researchers interested in the details of atom–light or molecule–light interactions had to rely on molecular chance resonances. Nevertheless, this led to the beginning of the study of coherent interactions and coherent transients such as photon echoes, self-induced transparency, optical nutation and so on (well described in the standard monograph by Allen and Eberly). Tuneable lasers became available in the early 1970s, and the dye laser in particular transformed precision studies in quantum optics and laser spectroscopy. Resonant interactions, coherent transients and the like became much more straightforward to study and led to the beginnings of quantum optics proper as we now understand it: for the first time we were able to study the dynamics of single atoms interacting with light in a non-perturbative manner. Stroud and his group initiated studies of resonance fluorescence with the observation of the splitting of resonance fluorescence spectral lines into component parts by the coherent driving predicted earlier by Mollow. Mandel, Kimble and others demonstrated how the resonance fluorescence light was antibunched, a feature studied by a number of theorists including Walls, Carmichael, Cohen-Tannoudji, Mandel and Kimble. The observation of

antibunching, and the associated (but inequivalent) sub-Poissonian photon statistics laid the foundation of the study of "non-classical light". During the 1970s, several experiments explored the nature of photons: their indivisibility and the build up of interference at the single photon level. Laser cooling rapidly developed in the 1980s and 1990s and allowed the preparation of states of matter under precise control. Indeed, this has become a major subject in its own right and we have taken the decision here to exclude laser cooling from this text.

Following the development of high-intensity pulses of light from lasers, a whole set of nonlinear optical phenomena were investigated, starting with the pioneering work in Ann Arbor by Franken and co-workers. Harmonic generation, parametric down-conversion and other phenomena were demonstrated. For the most part, none of this early work on nonlinear optics required field quantization and quantum optics proper for its description. But there were early signs that some could well do so: quantum nonlinear optics was really initiated by the study by Burnham and Weinberg (see Chapter 9) of unusual nonclassical correlations in down-conversion. In the hands of Mandel and many others, these correlations in down-conversion became the fundamental tool used to uncover fundamental insights into quantum optics.

Until the 1980s, essentially all light fields investigated had phase-independent noise; this changed with the production of squeezed light sources with phase-sensitive noise. These squeezed light sources enabled us to investigate Heisenberg uncertainty relations for light fields. Again, parametric down-conversion proved to be the most effective tool to generate such unusual light fields.

Quantum opticians realized quite early that were atoms to be confined in resonators, then atomic radiative transition dynamics could be dramatically changed. Purcell, in a remarkable paper in 1946 within the context of magnetic resonance, had already predicted that spontaneous emission rates, previously thought of as pretty immutable were in fact modified by enclosing the source atom within a cavity whose mode structure and densities are significantly different from those of free space. Putting atoms within resonators or close to mirrors became possible at the end of the 1960s. By the 1980s the theorists' dream of studying single atoms interacting with single modes of the electromagnetic field became possible. At this point the transition dynamics becomes wholly reversible, as the atom coherently exchanges excitation with the field, until coherence is eventually lost through a dissipative "decoherence" process. This dream is called the Jaynes–Cummings model after its proposers and forms a basic building block of quantum optics (and is discussed in detail in this book).

New fundamental concepts in information processing, leading to quantum cryptography and quantum computation, have been developed in recent years by Feynman, Benioff, Deutsch, Jozsa, Bennett, Ekert and others. Instead of using classical bits that can represent either the values 0 or 1, the basic unit of a

quantum computer is a quantum mechanical two-level system (qubit) that can exist in coherent superpositions of the logical values 0 and 1. A set of n qubits can then be in a superposition of up to 2^n different states, each representing a binary number. Were we able to control and manipulate say 1500 qubits, we could access more states than there are particles in the visible universe. Computations are implemented by unitary transformations, which act on all states of a superposition simultaneously. Quantum gates form the basic units from which these unitary transformations are built up. In related developments, absolutely secure encryption can be guaranteed by using quantum sources of light.

The use of the quantum mechanical superpositions and entanglement results in a high degree of parallelism, which can increase the speed of computation exponentially. A number of problems which cannot feasibly be tackled on a classical computer can be solved efficiently on a quantum computer. In 1994 a quantum algorithm was discovered by Peter Shor that allows the solution of a practically important problem, namely factorization, with such an exponential increase of speed. Subsequently, possible experimental realizations of a quantum computer have been proposed, for example in linear ion traps and nuclear magnetic resonance schemes. Presently we are at a stage where quantum gates have been demonstrated in these two implementations. Quantum computation is closely related to quantum cryptography and quantum communication. Basic experiments demonstrating the in-principle possibility of these ideas have been carried out in various laboratories.

The linear ion trap is one of the most promising systems for quantum computation and is one we study in this book in detail. The quantum state preparation (laser cooling and optical pumping) in this system is a well-established technique, as is the state measurement by electron shelving and fluorescence. Singly charged ions of an atom such as calcium or beryllium are trapped and laser cooled to micro-Kelvin temperatures, where they form a string lying along the axis of a linear radio-frequency (r.f.) Paul trap. The internal state of any one ion can be exchanged with the quantum state of motion of the whole string. This can be achieved by illuminating the ion with a pulse of laser radiation at a frequency tuned below the ion's internal resonance by the vibrational frequency of one of the normal modes of oscillation of the string. This couples single phonons into and out of the vibrational mode. The motional state can then be coupled to the internal state of another ion by directing the laser onto the second ion and applying a similar laser pulse. In this way general transformations of the quantum state of all the ions can be generated. The ion trap has several features to recommend it. It can achieve processing on quantum bits without the need for any new technological breakthroughs, such as micro-fabrication techniques or new cooling methods. The state of any ion can be measured and re-prepared many times without problem, which is an important feature for implementing quantum error correction protocols.

Trapped atoms or ions can be strongly coupled to an electromagnetic field mode in a cavity, which permits the powerful combination of quantum processing and long-distance quantum communication. This suggests ways in which we may construct quantum memories. These systems can in principle realize a quantum processor larger than any which could be thoroughly simulated by classical computing but the decoherence generated by dephasing and spontaneous emission is a formidable obstacle.

Entangled states are the key ingredient for certain forms of quantum cryptography and for quantum teleportation. Entanglement is also responsible for the power of quantum computing, which, under ideal conditions, can accomplish certain tasks exponentially faster than any classical computer. A deeper understanding of the role of quantum entanglement in quantum information theory will allow us to improve existing applications and to develop new methods of quantum information manipulation. These are all described in later chapters.

What then is the future of quantum optics? It underpins a great deal of laser science and novel atomic physics. It may even be the vehicle by which we can realize a whole new technology whereby quantum mechanics permits the processing and transmission of information in wholly novel ways. But of course, whatever we may predict now to emerge will be confounded by the unexpected: the field remains an adventure repeatedly throwing up the unexpected.

1.3 The contents of this book

The layout of this book is as follows. In Chapter 2, we show how the electromagnetic field can be quantized in terms of harmonic oscillators representing modes of the electromagnetic field, with states describing how many excitations (photons) are present in each normal mode. In Chapter 3 we introduce the coherent states, superposition states carrying phase information. In Chapter 4 we describe how light and matter interact. Chapter 5 quantifies our notions of coherence in terms of optical field correlation functions. Chapter 6 introduces simple optical elements such as beam splitters and interferometers, which manipulate the states of light. Chapter 7 describes those nonclassical states whose basic properties are dictated by their fundamental quantum nature. Spontaneous emission and decay in an open environment are discussed in Chapter 8. Chapter 9 describes how quantum optical sources of radiation can be used to provide tests of fundamental quantum mechanics, including tests of nonlocality and Bell inequalities. Chapter 10 discusses how atoms confined in cavities and trapped laser-cooled ions can be used to study basic interaction phenomena. Chapter 11 applies what we have learnt to the newly emerging problems of quantum information processing. Appendices set out some mathematical ideas needed within the main body of the text. Throughout we have tried to illustrate the ideas we have been developing through homework problems.

References

[1] A. Whitaker, *Einstein, Bohr and the Quantum Dilemma* (Cambridge: Cambridge University Press, 1996).
[2] A. Muthukrishnan, M. O. Scully and M. S. Zubairy, *Optics and Photonics News Trends*, **3**, No. 1 (October 2003).
[3] G. N. Lewis, *Nature,* **118** (1926), 874.
[4] W. E. Lamb, Jr., *Appl. Phys. B,* **66** (1995), 77.
[5] S. S. Schweber, *QED and the Men Who Made It: Dyson, Feynman, Schwinger and Tomonaga* (Princeton University Press, Princeton, 1994).

Suggestions for further reading

Many books on quantum optics exist, most taking the story much further than we do, in more specialized monographs.

L. Allen and J. H. Eberly, *Optical Resonance and Two Level Atoms* (New York: Wiley, 1975 and Mineola: Dover, 1987).
H. Bachor, *A Guide to Experiments in Quantum Optics* (Berlin & Weinheim: Wiley-VCH, 1998).
S. M. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics* (Oxford: Oxford University Press, 1997).
C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons and Atoms* (New York: Wiley-Interscience, 1989).
C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Atom–Photon Interactions* (New York: Wiley-Interscience, 1992).
V. V. Dodonov and V. I. Man'ko (editors), *Theory of Nonclassical States of Light* (London: Taylor and Francis, 2003).
P. Ghosh, *Testing Quantum Mechanics on New Ground* (Cambridge: Cambridge University Press, 1999).
H. Haken, *Light, Volume I*: *Waves, Photons, and Atoms* (Amsterdam: North Holland, 1981).
J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (New York: W. A. Benjamin, 1968).
U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge: Cambridge University Press, 1997).
W. H. Louisell, *Quantum Statistical Properties of Radiation* (New York: Wiley, 1973).
R. Loudon, *The Quantum Theory of Light*, 3rd edition (Oxford: Oxford University Press, 2000).
L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press, 1995).
P. Meystre and M. Sargent III, *Elements of Quantum Optics*, 2nd edition (Berlin: Springer-Verlag, 1991).
G. J. Milburn and D. F. Walls, *Quantum Optics* (Berlin: Springer-Verlag, 1994).
H. M. Nussenzveig, *Introduction to Quantum Optics* (London: Gordon and Breach, 1973).
M. Orszag, *Quantum Optics*: *Including Noise, Trapped Ions, Quantum Trajectories, and Decoherence* (Berlin: Springer, 2000).

J. Peřina, *Quantum Statistics of Linear and Nonlinear Optical Phenomena*, 2nd edition (Dordrecht: Kluwer, 1991).
V. Peřinová, A. Lukš, and J. Peřina, *Phase in Optics* (Singapore: World Scientific, 1998).
R. R. Puri, *Mathematical Methods of Quantum Optics* (Berlin: Springer, 2001).
M. Sargent, III, M. O. Scully and W. E. Lamb, Jr., *Laser Physics* (Reading: Addison-Wesley, 1974).
M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge: Cambridge University Press, 1997).
W. P. Schleich, *Quantum Optics in Phase Space* (Berlin: Wiley-VCH, 2001).
B. W. Shore, *The Theory of Coherent Atomic Excitation* (New York: Wiley-Interscience, 1990).
W. Vogel and D.-G. Welsch, *Lectures in Quantum Optics* (Berlin: Akademie Verlag, 1994).
M. Weissbluth, *Photon–Atom Interactions* (New York: Academic Press, 1989).
Y. Yamamoto and A. İmamoğlu, *Mesoscopic Quantum Optics* (New York: Wiley-Interscience, 1999).

A useful reprint collection of papers on coherent states, including the early work by Glauber, Klauder, and others, is the following.

J. R. Klauder and B.-S. Skagerstam (editors), *Coherent States* (Singapore: World Scientific, 1985).

The history of quantum optics and of fundamental tests of quantum theory can be found in a number of places. We have found the following invaluable.

R. Baeierlin, *Newton to Einstein* (Cambridge: Cambridge University Press, 1992).
M. Born and E. Wolf, *Principles of Optics* (Cambridge: Cambridge University Press, 1998).
A. Whitaker, *Einstein, Bohr and the Quantum Dilemma* (Cambridge: Cambridge University Press, 1996).

Chapter 2
Field quantization

In this chapter we present a discussion of the quantization of the electromagnetic field and discuss some of its properties with particular regard to the interpretation of the photon as an elementary excitation of a normal mode of the field. We start with the case of a single-mode field confined by conducting walls in a one-dimensional cavity and later generalize to multimode fields in free space. The photon number states are introduced and we discuss the fluctuations of the field observables with respect to these states. Finally, we discuss the problem of the quantum description of the phase of the quantized electromagnetic field.

2.1 Quantization of a single-mode field

We begin with the rather simple but very important case of a radiation field confined to a one-dimensional cavity along the z-axis with perfectly conducting walls at $z = 0$ and $z = L$ as shown in Fig. 2.1.

The electric field must vanish on the boundaries and will take the form of a standing wave. We assume there are no sources of radiation, i.e. no currents or charges nor any dielectric media in the cavity. The field is assumed to be polarized along the x-direction, $\mathbf{E}(\mathbf{r}, t) = \mathbf{e}_x E_x(z, t)$, where $\mathbf{e}_x$ is a unit polarization vector. Maxwell's equations without sources are, in SI units,

$$\nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} \tag{2.1}$$

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{2.2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.3}$$

$$\nabla \cdot \mathbf{E} = 0. \tag{2.4}$$

A single-mode field satisfying Maxwell's equations and the boundary conditions is given by

$$E_x(z, t) = \left(\frac{2\omega^2}{V\varepsilon_0}\right)^{1/2} q(t) \sin(kz) \tag{2.5}$$

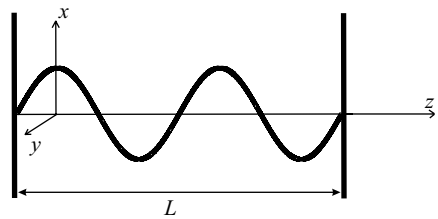


Fig. 2.1. Cavity with perfectly conducting walls located at $z = 0$ and $z = L$. The electric field is polarized along the x-direction.

where ω is the frequency of the mode and k is the wave number related to the frequency according to $k = \omega/c$. The boundary condition at $z = L$ yields the allowed frequencies $\omega_m = c(m\pi/L)$, $m = 1, 2, \ldots$ We assume that ω in Eq. (2.5) is one of these frequencies and ignore the rest for now. V in Eq. (2.5) is the effective volume of the cavity and $q(t)$ is a time-dependent factor having the dimension of length. As we shall see, $q(t)$ will act as a canonical position. The magnetic field in the cavity, from Eq. (2.5) and Eq. (2.2) is $\mathbf{B}(\mathbf{r}, t) = \mathbf{e}_y B_y(z, t)$ where

$$B_y(z, t) = \left(\frac{\mu_0 \varepsilon_0}{k}\right)\left(\frac{2\omega^2}{V\varepsilon_0}\right)^{1/2} \dot{q}(t)\cos(kz). \tag{2.6}$$

Here, $\dot{q}(t)$ will play the role of a canonical momentum for a "particle" of unit mass, i.e. $p(t) = \dot{q}(t)$.

The classical field energy, or Hamiltonian H, of the single-mode field is given by

$$\begin{aligned} H &= \frac{1}{2}\int dV\left[\varepsilon_0 \mathbf{E}^2(\mathbf{r}, t) + \frac{1}{\mu_0}\mathbf{B}^2(\mathbf{r}, t)\right] \\ &= \frac{1}{2}\int dV\left[\varepsilon_0 E_x^2(z, t) + \frac{1}{\mu_0}B_y^2(z, t)\right]. \end{aligned} \tag{2.7}$$

From Eqs. (2.5) and (2.6) it is straightforward to show (and is left as an exercise) that

$$H = \frac{1}{2}(p^2 + \omega^2 q^2), \tag{2.8}$$

from which it is apparent that a single-mode field is formally equivalent to a harmonic oscillator of unit mass, where the electric and magnetic fields, apart from some scale factors, play the roles of canonical position and momentum.

Every elementary textbook on quantum mechanics discusses the quantization of the one-dimensional harmonic oscillator. Here we take the approach that having identified the canonical variables q and p for the classical system, we simply use the correspondence rule to replace them by their operator equivalents $\hat{q}$ and $\hat{p}$ where operators will be distinguished from c-numbers by the caret. These operators must satisfy the canonical commutation relation

$$[\hat{q}, \hat{p}] = i\hbar\hat{I}. \tag{2.9}$$

Henceforth we follow custom and drop the identity operator $\hat{I}$ and write $[\hat{q}, \hat{p}] = i\hbar$. Then the electric and magnetic fields of the single mode become the operators

$$\hat{E}_x(z,t) = \left(\frac{2\omega^2}{V\varepsilon_0}\right)^{1/2} \hat{q}(t)\sin(kz), \tag{2.10}$$

and

$$\hat{B}_y(z,t) = \left(\frac{\mu_0\varepsilon_0}{k}\right)\left(\frac{2\omega^2}{V\varepsilon_0}\right)^{1/2} \hat{p}(t)\,\cos(kz), \tag{2.11}$$

respectively. The Hamiltonian becomes

$$\hat{H} = \frac{1}{2}(\hat{p}^2 + \omega^2\hat{q}^2). \tag{2.12}$$

The operators $\hat{q}$ and $\hat{p}$ are Hermitian and therefore correspond to observable quantities. However, it is convenient, and traditional, to introduce the non-Hermitian (and therefore non-observable) annihilation ($\hat{a}$) and creation ($\hat{a}^\dagger$) operators through the combinations

$$\hat{a} = (2\hbar\omega)^{-1/2}(\omega\hat{q} + i\hat{p}) \tag{2.13}$$

$$\hat{a}^\dagger = (2\hbar\omega)^{-1/2}(\omega\hat{q} - i\hat{p}). \tag{2.14}$$

The electric and magnetic field operators then become, respectively,

$$\hat{E}_x(z,t) = \mathcal{E}_0(\hat{a} + \hat{a}^\dagger)\sin(kz), \tag{2.15}$$

$$\hat{B}_y(z,t) = \mathcal{B}_0\frac{1}{i}(\hat{a} - \hat{a}^\dagger)\cos(kz), \tag{2.16}$$

where $\mathcal{E}_0 = (\hbar\omega/\varepsilon_0 V)^{1/2}$ and $\mathcal{B}_0 = (\mu_0/k)(\varepsilon_0\hbar\omega^3/V)^{1/2}$ represent respectively the electric and magnetic fields "per photon". The quotation marks indicate that this is not exactly correct since, as we shall show, the average of these fields for a definite number of photons is zero. Nevertheless, they are useful measures of the fluctuations of the quantized field. Operators $\hat{a}$ and $\hat{a}^\dagger$ satisfy the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{2.17}$$

and, as a result, the Hamiltonian operator takes the form

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \tag{2.18}$$

So far, we have said nothing of the time dependence of the operators $\hat{a}$ and $\hat{a}^\dagger$. For an arbitrary operator $\hat{O}$ having no explicit time dependence, Heisenberg's equation reads

$$\frac{d\hat{O}}{dt} = \frac{i}{\hbar}[\hat{H}, \hat{O}]. \tag{2.19}$$

For the annihilation operator $\hat{a}$ this becomes

$$\begin{aligned}\frac{d\hat{a}}{dt} &= \frac{i}{\hbar}[\hat{H}, \hat{a}] \\ &= \frac{i}{\hbar}\left[\hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \hat{a}\right] \\ &= i\omega(\hat{a}^\dagger\hat{a}\hat{a} - \hat{a}\hat{a}^\dagger\hat{a}) \\ &= i\omega[\hat{a}, \hat{a}^\dagger]\hat{a} = -i\omega\,\hat{a},\end{aligned} \tag{2.20}$$

which has the solution

$$\hat{a}(t) = \hat{a}\,(0)\,e^{-i\omega t}. \tag{2.21}$$

By the same method, or simply by taking the Hermitian conjugate of Eq. (2.21), we have

$$\hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}. \tag{2.22}$$

An alternate way of obtaining these solutions is to write the formal solution to Eq. (2.19) in the form

$$\hat{O}(t) = e^{i\hat{H}t/\hbar}\hat{O}(0)e^{-i\hat{H}t/\hbar} \tag{2.23}$$

and then to use the Baker–Hausdorf lemma [1] to obtain

$$\begin{aligned}\hat{O}(t) = \hat{O}(0) &+ \frac{it}{\hbar}[\hat{H}, \hat{O}\,(0)] \\ &+ \frac{1}{2!}\left(\frac{it}{\hbar}\right)^2[\hat{H}, [\hat{H}, \hat{O}\,(0)]] + \cdots \\ &+ \frac{1}{n!}\left(\frac{it}{\hbar}\right)^n[\hat{H}, [\hat{H}, [\hat{H}, \ldots [\hat{H}, \hat{O}(0)]]]] + \cdots\end{aligned} \tag{2.24}$$

For the operator $\hat{a}$ this results in

$$\begin{aligned}\hat{a}(t) = &\ \hat{a}(0)\left[1 - i\omega\,t - \frac{\omega^2t^2}{2!} + i\frac{\omega^3t^3}{3!} + \cdots\right] \\ &= \hat{a}(0)e^{-i\omega\,t}.\end{aligned} \tag{2.25}$$

The use of this method of solution may seem analogous to the use of a sledge-hammer to crack a nut but will turn out to be quite useful later when we take up cases involving nonlinear interactions.

The operator product $\hat{a}^\dagger\hat{a}$ has a special significance and is called the number operator, which we denote as $\hat{n}$. We let $|n\rangle$ denote an energy eigenstate of the single mode field with the energy eigenvalue E_n such that

$$\hat{H}|n\rangle = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|n\rangle = E_n|n\rangle. \tag{2.26}$$

If we multiply Eq. (2.26) by $\hat{a}^\dagger$ then we can generate a new eigenvalue equation

$$\hbar\omega\left(\hat{a}^\dagger\hat{a}\hat{a} + \frac{1}{2}\hat{a}^\dagger\right)|n\rangle = E_n\hat{a}^\dagger|n\rangle. \tag{2.27}$$

Using the commutation relations of Eq. (2.17) we can rewrite this as

$$\hbar\omega\left[(\hat{a}^\dagger\hat{a}-\hat{a}^\dagger)+\frac{1}{2}\hat{a}\right]|n\rangle = E_n\hat{a}^\dagger|n\rangle, \tag{2.28}$$

or

$$\hbar\omega\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)(\hat{a}^\dagger|n\rangle) = (E_n+\hbar\omega)(\hat{a}^\dagger|n\rangle), \tag{2.29}$$

which is the eigenvalue problem for the eigenstate $(\hat{a}^\dagger|n\rangle)$ with the energy eigenvalue $E_n+\hbar\omega$. It should be clear now why $\hat{a}^\dagger$ is called the creation operator: it creates a "quantum" of energy $\hbar\omega$. One could also say, rather loosely, that a "photon" of energy $\hbar\omega$ is created by $\hat{a}^\dagger$. Similarly, if we multiply Eq. (2.26) by the operator $\hat{a}$ and use the commutation relation we obtain

$$\hat{H}\,(\hat{a}\,|n\rangle) = (E_n-\hbar\omega)\,(\hat{a}\,|n\rangle) \tag{2.30}$$

where it is evident that the operator $\hat{a}$ destroys or annihilates one quantum of energy or one photon, the eigenstate $(\hat{a}|n\rangle)$ possessing the energy eigenvalue $E_n-\hbar\omega$. Evidently, repeating the procedure on Eq. (2.30) will result in the lowering of the energy eigenvalue by integer multiples of $\hbar\omega$. But the energy of the harmonic oscillator must always be positive so there must be a lowest-energy eigenvalue, $E_0>0$, with the corresponding eigenstate $|0\rangle$ such that

$$\hat{H}(\hat{a}\,|0\rangle) = (E_0-\hbar\omega)\,(\hat{a}\,|0\rangle) = 0 \tag{2.31}$$

because

$$\hat{a}\,|0\rangle = 0. \tag{2.32}$$

Thus, the eigenvalue problem for the ground state is

$$\hat{H}|0\rangle = \hbar\omega\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)|0\rangle = \frac{1}{2}\hbar\omega|0\rangle \tag{2.33}$$

so that the lowest-energy eigenvalue is the so-called zero-point energy $\hbar\omega/2$. Since $E_{n+1}=E_n+\hbar\omega$, the energy eigenvalues are

$$E_n = \hbar\omega\left(n+\frac{1}{2}\right), \qquad n=0,1,2,\ldots \tag{2.34}$$

(These energy levels are pictured, against the harmonic oscillator potential, in Fig. 2.2.)

Thus for the number operator $\hat{n}=\hat{a}^\dagger\hat{a}$ we have

$$\hat{n}\,|n\rangle = n\,|n\rangle\,. \tag{2.35}$$

These number states must be normalized according to $\langle n|n\rangle=1$. For the state $\hat{a}|n\rangle$ we have

$$\hat{a}\,|n\rangle = c_n|n-1\rangle\,, \tag{2.36}$$

where c_n is a constant to be determined. Then the inner product of $\hat{a}|n\rangle$

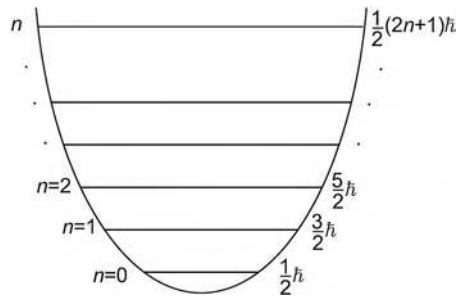


Fig. 2.2. The energy levels of a harmonic oscillator of frequency ω.

with itself is

$$\begin{aligned}(\langle n|\,\hat{a}^\dagger)(\hat{a}\,|n\rangle) &= \langle n|\,\hat{a}^\dagger\hat{a}\,|n\rangle = n \\ &= \langle n-1|\,c_n^* c_n\,|n-1\rangle = \left|c_n^2\right|,\end{aligned} \tag{2.37}$$

and thus $|c_n^2| = n$ so we can take $c_n = \sqrt{n}$. Thus

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{2.38}$$

Similarly we can show that

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.39}$$

From this last result it is straightforward to show that the number states $|n\rangle$ may be generated from the ground state $|0\rangle$ by the repeated action of the creation operator $\hat{a}$:

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}\,|0\rangle\,. \tag{2.40}$$

Because $\hat{H}$ and $\hat{n}$ are Hermitian operators, states of different number are orthogonal, i.e. $\langle n'|n\rangle = \delta_{nn'}$ and furthermore, the number states form a complete set, i.e.

$$\sum_{n=0}^{\infty} |n\rangle\,\langle n| = 1. \tag{2.41}$$

The only nonvanishing matrix elements of the annihilation and creation operators are

$$\langle n-1|\,\hat{a}\,|n\rangle = \sqrt{n}\,\langle n-1|n-1\rangle = \sqrt{n} \tag{2.42}$$

$$\langle n+1|\,\hat{a}\,|n\rangle = \sqrt{n+1}\,\langle n+1|n+1\rangle = \sqrt{n+1}\,. \tag{2.43}$$

2.2 Quantum fluctuations of a single-mode field

The number state $|n\rangle$ is a state of well-defined energy but it is not a state of well-defined electric field since

$$\langle n|\hat{E}_x(z,t)\,|n\rangle = \mathcal{E}_0 \sin(kz)\,[\langle n|\,\hat{a}\,|n\rangle + \langle n|\,\hat{a}^\dagger\,|n\rangle] = 0, \tag{2.44}$$

i.e. the mean field is zero. But the mean of the square of this field, which contributes to the energy density, is not zero:

$$\begin{aligned}\langle n|\,\hat{E}_x^2(z,t)\,|n\rangle &= \mathcal{E}_0^2\sin^2(kz)\,\langle n|\,\hat{a}^{\dagger 2}+\hat{a}^2+\hat{a}^\dagger\hat{a}+\hat{a}\hat{a}^\dagger\,|n\rangle \\ &= \mathcal{E}_0^2\sin^2(kz)\,\langle n|\,\hat{a}^{\dagger 2}+\hat{a}^2+2\hat{a}^\dagger\hat{a}+1\,|n\rangle \\ &= 2\mathcal{E}_0^2\sin^2(kz)\left(n+\frac{1}{2}\right).\end{aligned} \tag{2.45}$$

The fluctuations in the electric field may be characterized by the variance

$$\langle(\Delta\hat{E}_x(z,t))^2\rangle = \left\langle\hat{E}_x^2(z,t)\right\rangle - \langle\hat{E}_x(z,t)\rangle^2 \tag{2.46}$$

or by the standard deviation $\Delta E_x = \langle(\Delta E_x(z,t))^2\rangle^{1/2}$, which is sometimes referred to as the uncertainty of the field. For the number state $|n\rangle$ we have

$$\Delta E_x = \sqrt{2\mathcal{E}_0}\,\sin(kz)\left(n+\frac{1}{2}\right)^{1/2}. \tag{2.47}$$

Note that even when $n=0$, the field has fluctuations, the so-called vacuum fluctuations. Now the number states $|n\rangle$ are taken to represent a state of the field containing n photons. Yet as we have seen, the average field is zero. This is all in accordance with the uncertainty principle because the number operator $\hat{n}$ does not commute with the electric field:

$$[\hat{n},\,\hat{E}_x] = \mathcal{E}_0\sin(kz)\,(\hat{a}^\dagger-\hat{a}). \tag{2.48}$$

Thus $\hat{n}$ and $\hat{E}_x$ are complementary quantities for which their respective uncertainties obey the inequality*

$$\Delta n\,\Delta E_x \geq \frac{1}{2}\mathcal{E}_0|\sin(kz)||\langle\hat{a}^\dagger-\hat{a}\rangle|. \tag{2.49}$$

For a number state $|n\rangle$, the right-hand side vanishes but $\Delta n = 0$ as well. If the field were accurately known, then the number of photons would be uncertain. There is a connection here to the notion of the phase of the electric field. In classical physics, the amplitude and phase of a field can be simultaneously well defined. This is not so in quantum mechanics. In fact, the history of the concept of a quantum phase operator is long and contentious and we shall deal with this issue at length later. For now, we simply take a heuristic point of view for which the phase is in some sense complementary to number much in the way that time is complementary to energy. By analogy to the time–energy uncertainty relation there should be a number–phase uncertainty relation of the form

$$\Delta n\,\Delta\phi \geq 1. \tag{2.50}$$

From this, one could argue that for a well-defined (accurately known) phase the photon number is uncertain, whereas for a well-defined photon number, the phase is uncertain and, in fact, the phase is randomly distributed over the range

* Recall that for operators $\hat{A}$ and $\hat{B}$ satisfying $[\hat{A},\,\hat{B}] = \hat{C}$, $\Delta A\,\Delta B \geq \frac{1}{2}|\langle\hat{C}\rangle|$.

$0 < \phi < 2\pi$. We shall examine the issue of the quantum phase in more detail in Section 2.7.

2.3 Quadrature operators for a single-mode field

When we explicitly include the time dependence of the electric field operator we have

$$\hat{E}_x = \mathcal{E}_0(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t})\sin(kz) \tag{2.51}$$

where $\hat{a}(0) \equiv \hat{a}$ and $\hat{a}^\dagger(0) \equiv \hat{a}^\dagger$. We now introduce the so-called quadrature operators

$$\hat{X}_1 = \frac{1}{2}(\hat{a} + \hat{a}^\dagger) \tag{2.52}$$

$$\hat{X}_2 = \frac{1}{2i}(\hat{a} - \hat{a}^\dagger) \tag{2.53}$$

in terms of which the field operator may be recast as

$$\hat{E}_x(t) = 2\mathcal{E}_0 \sin(kz)[\hat{X}_1 \cos(\omega t) + \hat{X}_2 \sin(\omega t)]. \tag{2.54}$$

It is evident that $\hat{X}_1$ and $\hat{X}_2$ are associated with field amplitudes oscillating out of phase with each other by 90° (and hence are in quadrature). Note that $\hat{X}_1$ and $\hat{X}_2$ are essentially the position and momentum operators obtainable from Eqs. (2.13) and (2.14) but scaled to be dimensionless. They satisfy the commutation relation

$$[\hat{X}_1, \hat{X}_2] = \frac{i}{2} \tag{2.55}$$

from which it follows that

$$\langle(\Delta\hat{X}_1)^2\rangle\langle(\Delta\hat{X}_2)^2\rangle \geq \frac{1}{16}. \tag{2.56}$$

For the number states, $\langle n|\hat{X}_1|n\rangle = 0 = \langle n|\hat{X}_2|n\rangle$ but

$$\begin{aligned}\langle n|\hat{X}_1^2|n\rangle &= \frac{1}{4}\langle n|\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger|n\rangle \\ &= \frac{1}{4}\langle n|\hat{a}^2 + \hat{a}^\dagger 2 + 2\hat{a}^\dagger\hat{a} + 1|n\rangle \\ &= \frac{1}{4}(2n{+}1)\end{aligned} \tag{2.57}$$

and similarly

$$\langle n|\hat{X}_2^2|n\rangle = \frac{1}{4}(2n{+}1). \tag{2.58}$$

Thus for a number state, the uncertainties in both quadratures are the same and furthermore the vacuum state ($n = 0$) minimizes the uncertainty product since

$$\left\langle(\Delta\hat{X}_1)^2\right\rangle_{\text{vac}} = \frac{1}{4} = \left\langle(\Delta\hat{X}_2)^2\right\rangle_{\text{vac}}. \tag{2.59}$$

Before moving on to multimode fields, we want to stress that the quanta of the single-mode cavity field are the excitations of energy in discrete amounts of $\hbar\omega$. These quanta, universally referred to as photons, are not localized particles (in field theory, there is no position operator for photons) but rather are spread out over the entire mode volume. This is in sharp contrast to the view of photons as "corpuscles" of light as in the old quantum theory.

2.4 Multimode fields

The results for the single-mode field confined to a cavity can be generalized to multimode radiation fields. We shall consider these fields to be in free space where it is assumed that there are no sources of radiation and no charges so that Eqs. (2.1–2.4) still hold. The electric and magnetic radiation fields may be given in terms of the vector potential $\mathbf{A}(\mathbf{r}, t)$, which satisfies the wave equation

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0 \tag{2.60}$$

and the Coulomb gauge condition

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) = 0 \tag{2.61}$$

where

$$\mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} \tag{2.62}$$

and

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t). \tag{2.63}$$

The reason for this choice of gauge will be explained in Chapter 4.

We now imagine that free space can be modeled as a cubic cavity of side length L with perfectly reflecting walls. The idea here is that L should be very large compared with the dimensions of anything inside the cube with which the radiation could interact (e.g. atoms). We also assume that L is much larger than the wavelengths of the field. All physical results obtained from such a model should be independent of the size of the cavity as, after all calculations are done, we take $L \to \infty$.

The purpose of the cubical cavity is to allow us to impose periodic boundary conditions on the faces of the cube. For example, in the x-direction we shall require that plane waves satisfy the condition

$$e^{ik_x x} = e^{ik_x(x+L)} \tag{2.64}$$

from which it follows that

$$k_x = \left(\frac{2\pi}{L}\right) m_x, \qquad m_x = 0, \pm 1, \pm 2, \ldots \tag{2.65}$$

Similarly for the y- and z-directions we have

$$k_y = \left(\frac{2\pi}{L}\right) m_y, \qquad m_y = 0, \pm 1, \pm 2, \ldots \tag{2.66}$$

$$k_z = \left(\frac{2\pi}{L}\right) m_z, \qquad m_z = 0, \pm 1, \pm 2, \ldots \tag{2.67}$$

The wave vector is then

$$\mathbf{k} = \frac{2\pi}{L}(m_x, m_y, m_z) \tag{2.68}$$

and its magnitude is related to the frequency ω_k according to $k = \omega_k/c$. A set of integers (m_x, m_y, m_z) specifies a normal mode of the field (apart from polarization), the number of modes being infinite but denumerable. This is mathematically simpler than dealing with the continuum of modes in free space. The total number of modes in the intervals Δm_x, Δm_y, Δm_z is

$$\Delta m = \Delta m_x \Delta m_y \Delta m_z = 2\left(\frac{L}{2\pi}\right)^3 \Delta k_x \Delta k_y \Delta k_z, \tag{2.69}$$

where the factor of 2 takes into account the two independent polarizations. In a quasi-continuous limit, wherein we assume that wavelengths are small compared to L, we shall have waves densely packed in k-space and may therefore approximate Δm by the differential

$$dm = 2\left(\frac{V}{8\pi^3}\right) dk_x dk_y dk_z \tag{2.70}$$

where we have set $V = L^3$. In k-space spherical polar coordinates

$$\mathbf{k} = k\,(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{2.71}$$

and we have

$$dm = 2\left(\frac{V}{8\pi^3}\right) k^2 dk d\Omega \tag{2.72}$$

where $d\Omega = \sin\theta d\theta d\phi$ is the element of solid angle around the direction of $\mathbf{k}$. By using the relation $k = \omega_k/c$ we can transform Eq. (2.72) into

$$dm = 2\left(\frac{V}{8\pi^3}\right) \frac{\omega^2 k}{c^3} d\omega_k d\Omega. \tag{2.73}$$

Integrating Eq. (2.72) over the solid angle gives us:

$$\left.\begin{array}{l}\text{the numbers of modes}\\ \text{in all directions}\\ \text{in the range } k \text{ to}\\ k + dk\end{array}\right\} = V\frac{k^2}{\pi^2} dk = V\rho_k dk \tag{2.74}$$

where $\rho_k dk$ is the mode density (number of modes per unit volume) and obviously $\rho_k = k^2/\pi^2$. Integrating Eq. (2.73) in the same fashion yields:

$$\left.\begin{array}{l}\text{the numbers of modes}\\ \text{in all directions}\\ \text{in the range } \omega_k \text{ to}\\ \omega_k + d\omega_k\end{array}\right\} = V\frac{\omega_k^2}{\pi^2 c^3}d\omega_k \equiv V\rho(\omega_k)\,d\omega_k \tag{2.75}$$

where $\rho(\omega_k)d\omega_k$ is also the mode density with $\rho(\omega_k) = \omega_k^2/(\pi^2 c^3)$.

The vector potential can be expressed as a superposition of plane waves in the form

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k},s} \mathbf{e}_{\mathbf{k}s}\left[A_{\mathbf{k}s}(t)e^{i\mathbf{k}\cdot\mathbf{r}} + A^*_{\mathbf{k}s}(t)\,e^{-i\mathbf{k}\cdot\mathbf{r}}\right] \tag{2.76}$$

where $A_{\mathbf{k}s}$ is the complex amplitude of the field and where $\mathbf{e}_{\mathbf{k}s}$ is a real polarization vector. The sum over $\mathbf{k}$ simply means the sum over the set of integers (m_x, m_y, m_z) and the sum over s is the sum over the two independent polarizations. These polarizations must be orthogonal

$$\mathbf{e}_{\mathbf{k}s}\cdot\mathbf{e}_{\mathbf{k}s'} = \delta_{ss'} \tag{2.77}$$

and from the gauge condition of Eq. (2.66) must satisfy

$$\mathbf{k}\cdot\mathbf{e}_{\mathbf{k}s} = 0, \tag{2.78}$$

known as the *transversality* condition. The Coulomb gauge is sometimes known as the transverse gauge wherein the polarization is orthogonal to the propagation direction. The polarization vectors $\mathbf{e}_{\mathbf{k}1}$ and $\mathbf{e}_{\mathbf{k}2}$ form a right-handed system such that

$$\mathbf{e}_{\mathbf{k}1}\times\mathbf{e}_{\mathbf{k}2} = \frac{\mathbf{k}}{|\mathbf{k}|} = \boldsymbol{\kappa}. \tag{2.79}$$

In free space, the sum in Eq. (2.76) is replaced by the integral:

$$\sum_k \rightarrow \frac{V}{\pi^2}\int k^2 dk. \tag{2.80}$$

Now from Eqs. (2.60) and (2.61) we obtain for the complex amplitudes $A_{\mathbf{k},s}(t)$ the harmonic oscillator equation

$$\frac{d^2 A_{\mathbf{k}s}}{dt^2} + \omega_k^2 A_{\mathbf{k}s} = 0 \tag{2.81}$$

where $\omega_k = ck$. The solution is

$$A_{\mathbf{k}s}(t) = A_{\mathbf{k}s}e^{-i\omega_k t} \tag{2.82}$$

where we have set $A_{\mathbf{k}s}(0) \equiv A_{\mathbf{k}s}$. From Eqs. (2.62) and (2.63), the electric and magnetic fields respectively are

$$\mathbf{E}(\mathbf{r}, t) = i\sum_{\mathbf{k},s}\omega_k\mathbf{e}_{\mathbf{k}s}\left[A_{\mathbf{k}s}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} - A^*_{\mathbf{k}s}e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right], \tag{2.83}$$

$$\mathbf{B}(\mathbf{r}, t) = \frac{i}{c}\sum_{\mathbf{k},s}\omega_k(\boldsymbol{\kappa}\times\hat{\mathbf{e}}_{\mathbf{k}s})\left[A_{\mathbf{k}s}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} - A^*_{\mathbf{k}s}e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right]. \tag{2.84}$$

The energy of the field is given by

$$H = \frac{1}{2}\int_V \left(\varepsilon_0 \mathbf{E}\cdot\mathbf{E} + \frac{1}{\mu_0}\mathbf{B}\cdot\mathbf{B}\right) dV. \tag{2.85}$$

The periodic boundary condition results in

$$\int_0^L e^{\pm ik_x x} dx = \begin{cases} L & k_x = 0 \\ 0 & k_x \neq 0 \end{cases} \tag{2.86}$$

with similar results for the y- and z-directions. These may collectively be written as

$$\int_V e^{\pm i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} dV = \delta_{\mathbf{k}\mathbf{k}'} V. \tag{2.87}$$

From this we find that the contribution to H from the electric field is

$$\frac{1}{2}\int_V \mathcal{E}_0 \mathbf{E}\cdot\mathbf{E} dV = \varepsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A^*_{\mathbf{k}s}(t) - R \tag{2.88}$$

where

$$R = \frac{1}{2}\varepsilon_0 V \sum_{\mathbf{k}ss'} \omega_k^2 \mathbf{e}_{\mathbf{k}s}\cdot\hat{\mathbf{e}}_{-\mathbf{k}s'}[A_{\mathbf{k}s}(t)A_{-\mathbf{k}s'}(t) + A^*_{\mathbf{k}s}(t)A^*_{-\mathbf{k}s'}(t)]. \tag{2.89}$$

To obtain the magnetic contribution we need the vector identity

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C}) \tag{2.90}$$

from which we obtain

$$(\mathbf{k}\times\mathbf{e}_{\mathbf{k}s})\cdot(\mathbf{k}\times\mathbf{e}_{\mathbf{k}s'}) = \delta_{ss'} \tag{2.91}$$

$$(\mathbf{k}\times\mathbf{e}_{\mathbf{k}s})\cdot(-\mathbf{k}\times\mathbf{e}_{-\mathbf{k}s'}) = -\mathbf{e}_{\mathbf{k}s}\cdot\mathbf{e}_{-\mathbf{k}s'}. \tag{2.92}$$

Using these results we have

$$\frac{1}{2}\int \frac{1}{\mu_0}\mathbf{B}\cdot\mathbf{B} dV = \varepsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A^*_{\mathbf{k}s}(t) + R. \tag{2.93}$$

Thus adding Eqs. (2.88) and (2.93) we obtain the field energy

$$H = 2\varepsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A^*_{\mathbf{k}s}(t) \tag{2.94}$$

$$= 2\varepsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s} A^*_{\mathbf{k}s} \tag{2.95}$$

where we have used Eq. (2.82).

The energy of Eq. (2.95) has a very simple form in terms of the amplitudes $A_{\mathbf{k}s}$. In order to quantize the field, the canonical variables $p_{\mathbf{k}s}$ and $q_{\mathbf{k}s}$ must be introduced. We set

$$A_{\mathbf{k}s} = \frac{1}{2\omega_k(\varepsilon_0 V)^{1/2}}\left[\omega_k q_{\mathbf{k}s} + i p_{\mathbf{k}s}\right], \tag{2.96}$$

$$A^*_{\mathbf{k}s} = \frac{1}{2\omega_k(\varepsilon_0 V)^{1/2}}\left[\omega_k q_{\mathbf{k}s} - i p_{\mathbf{k}s}\right], \tag{2.97}$$

such that upon substitution into Eq. (2.95) we obtain

$$H = \frac{1}{2}\sum_{\mathbf{k}s}\left(p_{\mathbf{k}s}^2 + \omega_k^2 q_{\mathbf{k}s}^2\right), \tag{2.98}$$

each term of which is the energy of a simple harmonic oscillator of unit mass. The quantization of the field proceeds by demanding that the canonical variables become operators satisfying the commutation relations

$$[\hat{q}_{\mathbf{k}s}, \hat{q}_{\mathbf{k}'s'}] = 0 = [\hat{p}_{\mathbf{k}s}, \hat{p}_{\mathbf{k}'s'}] \tag{2.99}$$

$$[\hat{q}_{\mathbf{k}s}, \hat{p}_{\mathbf{k}'s'}] = i\,\hbar\delta_{\mathbf{k}\mathbf{k}'}\delta_{ss'}. \tag{2.100}$$

As for the single-mode field, annihilation and creation operators may be defined as

$$\hat{a}_{\mathbf{k}s} = \frac{1}{(2\,\hbar\omega_k)^{1/2}}\left[\omega_k\hat{q}_{\mathbf{k}s} + i\hat{p}_{\mathbf{k}s}\right], \tag{2.101}$$

$$\hat{a}_{\mathbf{k}s}^\dagger = \frac{1}{(2\,\hbar\omega_k)^{1/2}}\left[\omega_k\hat{q}_{\mathbf{k}s} - i\hat{p}_{\mathbf{k}s}\right], \tag{2.102}$$

which satisfy

$$[\hat{a}_{\mathbf{k}s}, \hat{a}_{\mathbf{k}'s'}] = 0 = \left[\hat{a}_{\mathbf{k}s}^\dagger, \hat{a}_{\mathbf{k}'s'}^\dagger\right] \tag{2.103}$$

$$\left[\hat{a}_{\mathbf{k}s}, \hat{a}_{\mathbf{k}'s'}^\dagger\right] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{ss'}. \tag{2.104}$$

The energy of the field becomes the Hamiltonian operator

$$\hat{H} = \sum_{\mathbf{k}s}\hbar\omega_k\left(\hat{a}_{\mathbf{k}s}^\dagger\hat{a}_{\mathbf{k}s} + \frac{1}{2}\right) \tag{2.105}$$

$$= \sum_{\mathbf{k}s}\hbar\omega_k\left(\hat{n}_{\mathbf{k}s} + \frac{1}{2}\right), \tag{2.106}$$

where

$$\hat{n}_{\mathbf{k}s} = \hat{a}_{\mathbf{k}s}^\dagger\hat{a}_{\mathbf{k}s} \tag{2.107}$$

is the number operator for the mode $\mathbf{k}s$. Each of these modes, being independent of all the others, has an associated set of number eigenstates $|n_{\mathbf{k}s}\rangle$. For the jth mode, let $\hat{a}_{\mathbf{k}_j s_j} \equiv \hat{a}_j$, $\hat{a}_{\mathbf{k}_j s_j}^\dagger \equiv \hat{a}_j$ and $\hat{n}_{\mathbf{k}_j s_j} \equiv \hat{n}_j$. The Hamiltonian for the field is then

$$\hat{H} = \sum_j \hbar\omega_j\left(\hat{n}_j + \frac{1}{2}\right) \tag{2.108}$$

and a multimode photon number state is just a product of the number states of all the modes which we write as

$$\begin{aligned} |n_1\rangle\,|n_2\rangle\,|n_3\rangle\ldots &\equiv |n_1, n_2, n_3\ldots\rangle \\ &= |\{n_j\}\rangle. \end{aligned} \tag{2.109}$$

This is an eigenstate of $\hat{H}$ such that

$$\hat{H}|\{n_j\}\rangle = E|\{n_j\}\rangle \tag{2.110}$$

where the eigenvalue E is

$$E = \sum_j \hbar\omega_j \left(n_j + \frac{1}{2}\right). \tag{2.111}$$

Of course, these number states are orthogonal according to

$$\langle n_1, n_2, \ldots | \, n'_1, n'_2, \ldots\rangle = \delta_{n_1 n'_1}\delta_{n_2 n'_2}\ldots \tag{2.112}$$

The action of the annihilation operator of the jth mode on the multimode number state is

$$\hat{a}_j|n_1, n_2, \ldots n_j, \ldots\rangle = \sqrt{n_j}\,|n_1, n_2, \ldots n_j - 1, \ldots\rangle. \tag{2.113}$$

Similarly, for the creation operator

$$\hat{a}_j^\dagger|n_1, n_2, \ldots n_j, \ldots\rangle = \sqrt{n_j + 1}\,|n_1, n_2, \ldots n_j + 1, \ldots\rangle. \tag{2.114}$$

The multimode vacuum state is denoted

$$|\{0\}\rangle = |0_1, 0_2, \ldots 0_j, \ldots\rangle \tag{2.115}$$

for which

$$\hat{a}_j|\{0\}\rangle = 0 \tag{2.116}$$

for all j. All the number states can be generated from the vacuum according to

$$|\{n_j\}\rangle = \prod_j \frac{\left(\hat{a}_j^\dagger\right)^{n_j}}{\sqrt{n_j!}}\,|\{0\}\rangle\,. \tag{2.117}$$

Upon quantization of the field, the amplitudes $A_{\mathbf{k}s}$ become operators which, from Eqs. (2.98) and (2.101), have the form

$$\hat{A}_{\mathbf{k}s} = \left(\frac{\hbar}{2\omega_k\varepsilon_0 V}\right)^{\frac{1}{2}} \hat{a}_{\mathbf{k}s} \tag{2.118}$$

and thus the quantized vector potential has the form

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}s}\left(\frac{\hbar}{2\omega_k\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s}\left[\hat{a}_{\mathbf{k}s}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right]. \tag{2.119}$$

The electric field operator is then

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i\sum_{\mathbf{k}s}\left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s}\left[\hat{a}_{\mathbf{k}s}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} - \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right] \tag{2.120}$$

while the magnetic field operator is

$$\hat{\mathbf{B}}(\mathbf{r}, t) = \frac{i}{c}\sum_{\mathbf{k}s}(\boldsymbol{\kappa}\times\mathbf{e}_{\mathbf{k}s})\left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s}\left[\hat{a}_{\mathbf{k}s}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} - \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right], \tag{2.121}$$

where $\boldsymbol{\kappa} = \mathbf{k}/|\mathbf{k}|$. The annihilation and creation operators appearing in Eqs. (2.119–2.121) are to be understood as Heisenberg picture operators evaluated

at time $t = 0$. As in the single-mode case, the time-dependent annihilation operator for a free field is given by

$$\hat{a}_{\mathbf{k}s}(t) = \hat{a}_{\mathbf{k}s}(0)e^{-i\omega_k t}. \tag{2.122}$$

Thus the electric field, for instance, can be written as

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i \sum_{\mathbf{k}s} \left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s}\left(\hat{a}_{\mathbf{k}s}(t)e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}^{\dagger}_{\mathbf{k}s}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\right). \tag{2.123}$$

Sometimes this field is written as

$$\hat{\mathbf{E}}(\mathbf{r}, t) = \hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) + \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \tag{2.124}$$

where

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) = i \sum_{\mathbf{k}s} \left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s}\hat{a}_{\mathbf{k}s}(t)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.125}$$

and where

$$\hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) = \left[\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)\right]^{\dagger}. \tag{2.126}$$

$\hat{\mathbf{E}}^{(+)}$ is called the positive frequency part of the field as it contains all terms that oscillate as $e^{-i\omega t}$ for $\omega > 0$, while $\hat{\mathbf{E}}^{(-)}$ is called the negative frequency part. The former is essentially a collective annihilation operator while the latter is a collective creation operator. Similar expressions can be written for the magnetic field and for the vector potential.

In most quantum optical situations, the coupling of the field to matter is through the electric field interacting with a dipole moment or through some nonlinear type of interaction involving powers of the electric field. Thus we shall be mostly interested in the electric field throughout the rest of the book. Furthermore, note that the magnetic field is "weaker" than the electric field by a factor of $1/c$. The field couples to the spin magnetic moment of the electrons and this interaction is negligible for essentially all the aspects of quantum optics that we are concerned with.

For a single-mode plane wave field the electric field is

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_x[\hat{a}e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} - \hat{a}^{\dagger}e^{-i\mathbf{k}\cdot\mathbf{r}+i\omega t}]. \tag{2.127}$$

In much of quantum optics, the spatial variation of the field over the dimensions of the atomic system may be negligible. For optical radiation, λ is on the order of several thousand ångströms so that

$$\frac{\lambda}{2\pi} = \frac{1}{|\mathbf{k}|} \gg |\mathbf{r}_{\text{atom}}|, \tag{2.128}$$

where $|\mathbf{r}_{\text{atom}}|$ is a length characteristic of the size of an atom. Under this condition

$$e^{\pm i\mathbf{k}\cdot\mathbf{r}} \approx 1 \pm i\mathbf{k}\cdot\mathbf{r} \tag{2.129}$$

and we can replace the exponential by unity to obtain

$$\begin{aligned}\hat{\mathbf{E}}(\mathbf{r},t) &\approx \hat{\mathbf{E}}(t)\\ &= i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{\frac{1}{2}} \mathbf{e}_x[\hat{a}e^{-i\omega t} - \hat{a}^\dagger e^{i\omega t}].\end{aligned} \tag{2.130}$$

This approximation, which will be discussed again in Chapter 4, is called the "dipole" approximation.

2.5 Thermal fields

As is well known, quantum theory originated with Planck's discovery of the radiation law that now bears his name. We refer, of course, to the law describing the radiation emitted by an ideal object known as a black body – a perfect emitter and absorber of radiation. A black body can be modeled as a cavity (or actually a small hole in the cavity) containing radiation at thermal equilibrium with its walls. The radiation is thus coupled to a heat bath and so is not, unlike in the preceding sections of this chapter, a truly free field. But assuming the coupling is weak, we can, according to the theory of statistical mechanics, treat the field as if it were an isolated system that can be described as a microcanical ensemble.

We consider then, for the moment, a single-mode field in thermal equilibrium with the walls of a cavity at temperature T. According to statistical mechanics, the probability P_n that the mode is thermally excited in the nth level is

$$P_n = \frac{\exp(-E_n/k_\mathrm{B}T)}{\sum_n \exp(-E_n/k_\mathrm{B}T)} \tag{2.131}$$

where the E_n are given in Eq. (2.131) and where k_B is the Boltzmann constant ($k_\mathrm{B} = 1.38 \times 10^{-23}$J/K). We introduce here the density operator (whose general properties are described in Appendix A) for the thermal field:

$$\hat{\rho}_\mathrm{Th} = \frac{\exp(-\hat{H}/k_\mathrm{B}T)}{\mathrm{Tr}[\exp(-\hat{H}/k_\mathrm{B}T)]} \tag{2.132}$$

where $\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$ and where

$$\begin{aligned}\mathrm{Tr}[\exp(-\hat{H}/k_\mathrm{B}T)] &= \sum_{n=0}^{\infty} \langle n|\exp(-\hat{H}/k_\mathrm{B}T)|n\rangle\\ &= \sum_{n=0}^{\infty} \exp(-E_n/k_\mathrm{B}T) \equiv Z\end{aligned} \tag{2.133}$$

is the partition function. With $E_n = \hbar\omega(n + \frac{1}{2})$

$$Z = \exp(-\hbar\omega/2k_\mathrm{B}T)\sum_{n=0}^{\infty}\exp(-\hbar\omega n/k_\mathrm{B}T). \tag{2.134}$$

Since $\exp(-\hbar\omega/k_{\mathrm{B}}T) < 1$, the sum is a geometric series and thus

$$\sum_{n=0}^{\infty} \exp(-\hbar\omega n/k_{\mathrm{B}}T) = \frac{1}{1-\exp(-\hbar\omega/k_{\mathrm{B}}T)} \tag{2.135}$$

so that

$$Z = \frac{\exp(-\hbar\omega/2k_{\mathrm{B}}T)}{1-\exp(-\hbar\omega/k_{\mathrm{B}}T)}. \tag{2.136}$$

Evidently

$$P_n = \langle n\,|\hat{\rho}_{\mathrm{Th}}|\,n\rangle = \frac{1}{Z}\exp(-E_n/k_{\mathrm{B}}T). \tag{2.137}$$

Also note the density operator itself can be written as

$$\begin{aligned}\hat{\rho}_{\mathrm{Th}} &= \sum_{n'=0}^{\infty}\sum_{n=0}^{\infty} |n'\rangle\langle n'|\hat{\rho}_{\mathrm{Th}}\,|n\rangle\,\langle n| \\ &= \frac{1}{Z}\sum_{n=0}^{\infty}\exp(-E_n/k_{\mathrm{B}}T)\,|n\rangle\,\langle n| \\ &= \sum_{n=0}^{\infty} P_n\,|n\rangle\,\langle n|\,.\end{aligned} \tag{2.138}$$

The average photon number of the thermal field is calculated as

$$\begin{aligned}\bar{n} &= \langle\hat{n}\rangle = \mathrm{Tr}\,(\hat{n}\hat{\rho}_{\mathrm{Th}}) = \sum_{n=0}^{\infty}\langle n\,|\hat{n}\hat{\rho}_{\mathrm{Th}}|\,n\rangle \\ &= \sum_{n=0}^{\infty} n P_n = \exp(-\hbar\omega/2k_{\mathrm{B}}T)\frac{1}{Z}\sum_{n=0}^{\infty} n\,\exp(-\hbar\omega n/k_{\mathrm{B}}T).\end{aligned} \tag{2.139}$$

Noting that with $x = \hbar\omega/k_{\mathrm{B}}T$, we have

$$\begin{aligned}\sum_{n=0}^{\infty} n e^{-nx} &= -\frac{d}{dx}\sum_{n=0}^{\infty} e^{-nx} \\ &= -\frac{d}{dx}\left(\frac{1}{1-e^{-x}}\right) \\ &= \frac{e^{-x}}{(1-e^{-x})^2}.\end{aligned} \tag{2.140}$$

Thus we have

$$\begin{aligned}\bar{n} &= \frac{\exp(-\hbar\omega/k_{\mathrm{B}}T)}{1-\exp(-\hbar\omega/k_{\mathrm{B}}T)} \\ &= \frac{1}{\exp(\hbar\omega/k_{\mathrm{B}}T)-1}.\end{aligned} \tag{2.141}$$

Evidently

$$\bar{n} \approx \begin{cases} \dfrac{k_{\mathrm{B}}T}{\hbar\omega} & (k_{\mathrm{B}}T \gg \hbar\omega) \\ \dfrac{\hbar\omega}{k_{\mathrm{B}}T} & (k_{\mathrm{B}}T \ll \hbar\omega) \end{cases}. \tag{2.142}$$

At room temperatures, the average number of photons at optical frequencies is very small (on the order of 10^{-40}) . At the surface temperature of the sun (6000 K) and at the frequency of yellow light (6×10^{14} Hz, $\lambda = 500$ nm) the average photon number is about 10^{-2}. On the other hand, the average photon number rapidly increases with increasing wavelength. Again at room temperature, $\bar{n} \simeq 1$ for λ in the range λ =10–100 μm. In the microwave part of the spectrum, $\bar{n} \gg 1$.

From Eq. (2.141) it follows that

$$\exp(-\hbar\omega/k_{\mathrm{B}}T) = \frac{\bar{n}}{1+\bar{n}} \tag{2.143}$$

and from Eqs. (2.137) and (2.138) it follows that $\hat{\rho}_{\mathrm{Th}}$ can be written in terms of $\bar{n}$ as

$$\hat{\rho}_{\mathrm{Th}} = \frac{1}{1+\bar{n}} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{1+\bar{n}} \right)^n |n\rangle \langle n| . \tag{2.144}$$

The probability of finding n photons in the field is given in terms of $\bar{n}$ as

$$P_n = \frac{\bar{n}^n}{(1+\bar{n})^{n+1}}. \tag{2.145}$$

In Fig. 2.3 we plot P_n versus n for two different values of $\bar{n}$. It is clear in both cases that the most probable photon number is the vacuum, P_n decreasing monotonically with n. There is obviously nothing special about P_n for n near or at $\bar{n}$ (which need not be integer).

The fluctuations in the average photon number are given as

$$\langle (\Delta n)^2 \rangle = \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2 . \tag{2.146}$$

It can be shown, in a manner similar to the derivation of, $\bar{n}$ that

$$\begin{aligned} \langle \hat{n}^2 \rangle &= \mathrm{Tr}\left(\bar{n}^2 \hat{\rho}_{\mathrm{Th}} \right) \\ &= \bar{n} + 2\bar{n}^2 \end{aligned} \tag{2.147}$$

so that

$$\langle (\Delta n)^2 \rangle = \bar{n} + \bar{n}^2 \tag{2.148}$$

from which it is apparent that the *fluctuations* of $\hat{n}$ are larger than the *average* $\bar{n}$. The root-mean-square (r.m.s.) deviation is

$$\Delta n = (\bar{n} + \bar{n}^2)^{1/2} \tag{2.149}$$

which for $\bar{n} \gg 1$ is approximately

$$\Delta n \approx \bar{n} + \frac{1}{2}. \tag{2.150}$$

The relative uncertainty is given by the ratio $\Delta n/\bar{n}$, which is approximately 1 for $\bar{n} \gg 1$ and is approximately $1/\sqrt{\bar{n}}$ for $\bar{n} \ll 1$. Obviously, $\Delta n/\bar{n} \to \infty$ as $\bar{n} \to 0$.

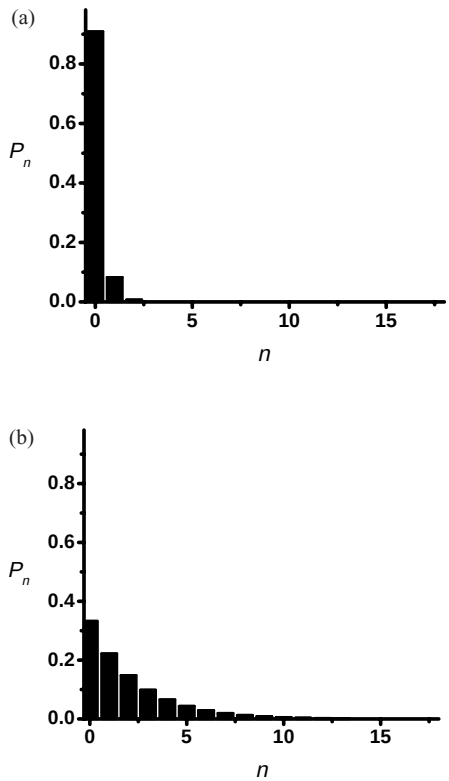


Fig. 2.3. Thermal photon number distributions for (a) $\bar{n} = 0.1$, and (b) $\bar{n} = 2$.

The average energy of the photons in the cavity is $\hbar\omega\,\bar{n}$. Planck's radiation law is obtained by multiplying the average energy of the photons by the density of modes per unit interval in ω in a unit volume, $\rho(\omega) = \omega^2/\pi^2c^3$ (where the two independent polarization directions have been taken into account), to obtain the average energy density per unit interval in ω as

$$\begin{aligned}\bar{U}(\omega) &= \hbar\omega\,\bar{n}\,\rho(\omega) \\ &= \frac{\hbar\omega^3}{\pi^2c^3}\,\frac{1}{\exp(\hbar\omega/k_{\mathrm{B}}T)-1}.\end{aligned} \tag{2.151}$$

For $k_{\mathrm{B}}T \gg \hbar\omega$, this takes the simpler form

$$\bar{U}(\omega) \approx \frac{\omega^3 k_{\mathrm{B}}T}{\pi^2c^3} \qquad (k_{\mathrm{B}}T \gg \hbar\omega) \tag{2.152}$$

which is known as Rayleigh's law. This is sometimes called the "classical limit" obtained from Planck's law of Eq. (2.150) by setting $\hbar \to 0$. (We note, however, that setting $\hbar$ to zero is not a well-defined limit: $\hbar$ contains dimensional information.) On the other hand, for low temperatures where $k_{\mathrm{B}}T \ll \hbar\omega$ we obtain

$$\bar{U}(\omega) \approx \frac{\hbar\omega^3}{\pi^2c^3}\exp\left(-\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) \qquad (k_{\mathrm{B}}T \ll \hbar\omega)\,. \tag{2.153}$$

This is Wien's law. By differentiation, it follows that $\bar{U}(\omega)$ has a maximum at

$$\omega_{\max} = \frac{2.8 k_{\mathrm{B}} T}{\hbar} = \frac{2\pi c}{\lambda_{\max}} \tag{2.154}$$

which is Wien's displacement law.

The average energy per unit volume is obtained by integrating over all frequencies:

$$\begin{aligned} \bar{U} &= \int_0^\infty \bar{U}(\omega) d\omega \\ &= \frac{\hbar^2}{\pi^2 c^3} \int_0^\infty \frac{\omega^3}{\exp(\hbar\omega/k_{\mathrm{B}} T) - 1}\, d\omega \\ &= \frac{\pi^2 k_{\mathrm{B}}^4 T^4}{15 c^3 \hbar^3}. \end{aligned} \tag{2.155}$$

This is the Stefan–Boltzmann law.

2.6 Vacuum fluctuations and the zero-point energy

We have seen that the quantized radiation field fluctuates. For a single-mode field the fluctuations in the electric field strength are given by Eq. (2.47). With the field mode in the vacuum state $|0\rangle$, the r.m.s. fluctuation of the field strength is

$$\Delta E_x = \mathcal{E}_0 \sin(kz). \tag{2.156}$$

These vacuum fluctuations and the zero-point energy have a common origin in the noncommutability of the operators $\hat{a}$ and $\hat{a}^\dagger$. The zero-point energy and the vacuum fluctuations actually present severe problems in quantum field theory. The most glaring of these comes about as follows: the universe contains an infinite number of radiation modes, each with a finite zero-point energy, $\hbar\omega/2$. The total zero-point energy (ZPE) of the universe then is

$$E_{\mathrm{ZPE}} = \frac{\hbar}{2} \sum_\omega \omega \to \infty \tag{2.157}$$

unless somehow the high-frequency modes are excluded. It is frequently said that energy differences only are important but this cannot quite be the whole story because according to general relativity, it is the total energy that counts, not just energy differences [2]. Other "infinites" appearing in the theory of quantum electrodynamics have been "swept under the rug" through the renormalization procedure, but this particular one still sticks out like a sore thumb. In fact, the vacuum energy and fluctuations actually give rise to observable effects. For example, spontaneous emission, which generates most of the visible light around us as thermal radiation, is a direct result of the vacuum fluctuations as we will show in Chapter 4. The ZPE gives rise to at least two other effects, one being the Lamb shift and the other being the Casimir effect.

The Lamb shift is a discrepancy between experiment and the Dirac relativistic theory of the hydrogen atom. The theory predicts that the $2^2S_{1/2}$ and $2^2P_{1/2}$ levels should be degenerate. Early optical work suggested that these states were not degenerate but separated by about 0.033 cm. By using an elegant combination of atomic beam and microwave techniques, Lamb and Retherford [3] showed that the $2^2S_{1/2}$ state has a higher energy than the $2^2P_{1/2}$ state by the equivalent of about 1000 MHz. In 1947, Bethe [4] explained this splitting as being caused by the interaction of the bound electron with the ZPE. Here we present a simple intuitive interpretation originally given by Welton in 1948 [5]. For this calculation only, we follow the lead of Welton who uses cgs units.

Each mode contains ZPE $h\nu/2$, where $\nu = \omega/2\pi$. The number of modes in a cavity of volume V with frequency between ν and $\nu + d\nu$ is $(8\pi/c^3)\nu^2 d\nu$. Thus the ZPE field energy is

$$\begin{aligned}\left(\frac{8\pi}{c^3}\nu^2 d\nu V\right)\frac{1}{2}h\nu &= \frac{1}{8\pi}\int_V \left(E_\nu^2 + B_\nu^2\right) dV \\ &= \frac{1}{8\pi}E_\nu^2 V\end{aligned} \tag{2.158}$$

where E_ν is the amplitude of the electric field component of frequency ν. Thus

$$E_\nu^2 = \frac{32\pi^2}{c^3}h\nu^3 d\nu. \tag{2.159}$$

The electron bound in the hydrogen atom interacts with the fluctuating zero-point electric field and with the Coulomb potential of the proton $-e^2/r$. If r represents the electron's "standard orbit" and Δr represents the fluctuations from this orbit, then the change in the potential energy is $\Delta V = V(r + \Delta r) - V(r)$ which by Taylor's theorem gives

$$\begin{aligned}\Delta V = {} & \Delta x\frac{\partial V}{\partial x} + \Delta y\frac{\partial V}{\partial y} + \Delta z\frac{\partial V}{\partial z} \\ & + \frac{1}{2}(\Delta x)^2\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}(\Delta y)^2\frac{\partial^2 V}{\partial y^2} + \frac{1}{2}(\Delta z)^2\frac{\partial^2 V}{\partial z^2} + \cdots\end{aligned} \tag{2.160}$$

Since the fluctuations are isotropic $\langle\Delta x\rangle = \langle\Delta y\rangle = \langle\Delta z\rangle = 0$ and $\langle(\Delta x)^2\rangle = \langle(\Delta y)^2\rangle = \langle(\Delta z)^2\rangle = \langle(\Delta r)^2\rangle/3$. Then

$$\langle\Delta V\rangle = \frac{1}{6}\langle(\Delta r)^2\rangle\nabla^2 V. \tag{2.161}$$

For the atomic state $|nlm_l\rangle$ the energy shift to first order is

$$\begin{aligned}\Delta E &= \langle nlm_l|\,\langle\Delta V\rangle\,|nlm_l\rangle \\ &= \frac{1}{6}\langle(\Delta r)^2\rangle\,\langle nlm_l|\,\nabla^2 V\,|nlm_l\rangle\,.\end{aligned} \tag{2.162}$$

With $V = -e^2/r$ and $\nabla^2(1/r) = -4\pi\delta(r)$ we obtain

$$\langle nlm_l|\,\nabla^2 V\,|nlm_l\rangle = 4\pi e^2\left|\psi_{nml_l}(r=0)\right|^2. \tag{2.163}$$

All the atomic wave functions of nonrelativistic quantum theory vanish at the origin except for the s-states with $l = 0$, where

$$|\psi_{n00}(r = 0)^2| = \frac{1}{\pi n^3 a_0^3} \tag{2.164}$$

where a_0 is the Bohr radius. For p-states the wave function vanishes at the origin and therefore so does the energy shift. To obtain $\langle(\Delta r)^2\rangle$ we assume that the important field frequencies greatly exceed the atomic resonance frequencies, the lower frequencies being shielded by the atomic binding and unable to influence the motion of the electrons. The displacement Δr_ν induced with frequency between ν and $\nu + d\nu$ is determined by

$$\frac{d^2 \Delta r_\nu}{dt^2} = \frac{eE_\nu}{m} \exp(2\pi\nu i t). \tag{2.165}$$

The solution is

$$\Delta r_\nu = -\frac{e^2}{m} \frac{E_\nu}{4\pi^2\nu^2} \exp(2\pi\nu i t). \tag{2.166}$$

The mean square displacement induced by these modes is

$$\langle(\Delta r_\nu)^2\rangle = -\frac{e^2}{m^2} \frac{E_\nu^2}{32\pi^4\nu^4} = \frac{e^2 h}{\pi^2 m^2 c^3} \frac{d\nu}{\nu}. \tag{2.167}$$

The s-state energy shift obtained by summing over all frequencies is

$$\Delta E = \frac{2}{3} \left(\frac{e^2}{\hbar c}\right)^2 \left(\frac{\hbar}{mc}\right) \frac{hc}{\pi^2 n^3 a_0^3} \int \frac{d\nu}{\nu} \tag{2.168}$$

where $e^2/\hbar c$ is the fine structure constant, and $\hbar/mc$ is the Compton wavelength of the electron. The integral is divergent but may be cut off at both high and low frequencies. At low frequencies, the atom does not respond to the fluctuating field, the frequency of the electron's orbit $\nu_0 = e^2/\hbar a_0^3 n^3$ being the natural cut-off. At high frequencies, relativistic effects show up in the electron's motion. But the preceding analysis is nonrelativistic so that

$$\frac{v}{c} = \left(\frac{p/m}{c}\right) = \frac{pc}{mc^2} = \frac{\hbar k}{mc} < 1 \tag{2.169}$$

which restricts k to less than $(mc/\hbar)$ and angular frequencies to less than $mc^2/\hbar$ in the integral of Eq. (2.168). Thus for the $2^2S_{1/2}$ state of hydrogen, with $a_0 = \hbar^2/me^2$, the energy shift is

$$\Delta E = \frac{1}{6\pi} \left(\frac{e^2}{\hbar c}\right)^3 \frac{me^4}{\hbar^2} \log\left(\frac{mc^2}{h\nu_0}\right) \tag{2.170}$$

which gives $\Delta E/h \sim 1000$ MHz. The $2^2P_{1/2}$ state is unaffected to this order.

The Casimir effect [6], in the simplest version, is the occurrence of a force between two parallel perfectly conducting plates owing to a change in the ZPE

resulting from the boundary conditions on the plates. We will follow the discussion of Milonni and Shih [7] to show how this force arises.

Consider a parallelepiped with perfectly conducting walls of lengths $L_x = L_y = L$ and $L_z = d$. The boundary conditions on the walls restrict the allowed frequencies to those given by

$$\omega_{lmn} = \pi c \left(\frac{l^2}{L^2} + \frac{m^2}{L^2} + \frac{n^2}{d^2} \right)^{\frac{1}{2}} \tag{2.171}$$

where l, m, n take on non-negative integer values. If there is no box, all frequencies are allowed. The ZPE in the box is

$$E_0(d) = \sum\nolimits'_{l,m,n} (2) \frac{1}{2} \hbar \omega_{lmn} \tag{2.172}$$

where the factor of two accounts for two independent polarizations and where the prime on the summation sign means that the two is to be removed if one of the integers l, m, n is zero, as there is only one independent polarization for that case. We shall be interested only in the case where $L \gg d$ so that we can replace the sums of l and m by integrals to write

$$E_0(d) = \frac{\hbar c L^2}{\pi} \sum_{n=0}^{\infty} \int_0^{\infty} dx \int_0^{\infty} dy \left(x^2 + y^2 + \frac{\pi^2 n^2}{d^2} \right)^{\frac{1}{2}}. \tag{2.173}$$

On the other hand, if d is arbitrarily large, the sum over n can be replaced by an integral so that

$$E_0(\infty) = \frac{\hbar c L^2}{\pi^2} \frac{d}{\pi} \int_0^{\infty} dx \int_0^{\infty} dy \int_0^{\infty} dz (x^2 + y^2 + z^2)^{\frac{1}{2}}. \tag{2.174}$$

When the plates are separated by the distance d, the potential energy of the system is just $U(d) = E_0(d) - E_0(\infty)$, which is the energy required to bring the plates from infinity to a distance d. Thus

$$\begin{aligned} U(d) = \frac{L^2 \hbar c}{\pi} \Bigg[& \sum_{n=0}^{\infty} \int_0^{\infty} dx \int_0^{\infty} dy \left(x^2 + y^2 + \frac{\pi^2 n^2}{d^2} \right)^{\frac{1}{2}} \\ & - \frac{d}{\pi} \int_0^{\infty} dx \int_0^{\infty} dy \int_0^{\infty} dz (x^2 + y^2 + z^2)^{\frac{1}{2}} \Bigg]. \end{aligned} \tag{2.175}$$

Transforming to polar coordinates in the x–y plane we have

$$\begin{aligned} U(d) = \frac{L^2 \hbar c}{\pi^2} \frac{\pi}{2} \Bigg[& \sum_{n=0}^{\infty} \int_0^{\infty} dr\, r \left(r^2 + \frac{n^2 \pi^2}{d^2} \right)^{\frac{1}{2}} \\ & - \frac{d}{\pi} \int_0^{\infty} dz \int_0^{\infty} dr\, r (r^2 + z^2)^{\frac{1}{2}} \Bigg]. \end{aligned} \tag{2.176}$$

Making the change of variable $w = r^2$

$$U(d) = \frac{L^2\hbar c}{4\pi^2}\frac{\pi^3}{d^3}\left[\sum_{n=0}^{\infty}\int_0^{\infty} dw\ (w+n^2)^{\frac{1}{2}} - \int_0^{\infty} dz \int_0^{\infty} dw (w+z^2)^{\frac{1}{2}}\right]. \tag{2.177}$$

Both ZPEs in Eq. (2.177) are infinite but the difference is finite. We can write this as

$$U(d) = \frac{\pi^2\hbar c}{4d^3}L^2\left[\frac{1}{2}F(0) + \sum_{n=1}^{\infty}F(n) - \int_0^{\infty} dz F(n)\right] \tag{2.178}$$

where

$$F(u) \equiv \int_0^{\infty} dw (w+u^2)^{\frac{1}{2}} \tag{2.179}$$

with $u = n$ or z. The difference can be estimated by the Euler–Maclaurin formula:

$$\sum_0^{\infty} F(n) - \int_0^{\infty} dz F(z) = -\frac{1}{2}F(0) - \frac{1}{12}F'(0) + \frac{1}{720}F'''(0)\ldots \tag{2.180}$$

since $F'(z) = 2z^2$, $F'(0) = 0$, $F'''(0) = -4$, and all higher-order derivations vanish. Thus

$$U(d) = \frac{\pi^2\hbar c}{4d^3}L^3\left(-\frac{4}{720}\right) = -\frac{\pi^2\hbar c}{720d^3}L^2. \tag{2.181}$$

This implies that the force per unit area between the plates is given by

$$F(d) = U'(d) = -\frac{\pi^2\hbar c}{240d^4} \tag{2.182}$$

which is the Casimir force. The existence of this force was confirmed by experiments carried out by Sparnaay [8] in 1957.

2.7 The quantum phase

Consider now a light wave as pictured in the classical electromagnetic theory. The electric field of a single mode can be written as

$$\begin{aligned}\mathbf{E}(\mathbf{r},t) &= \mathbf{e}_x E_0 \cos(\mathbf{k}\cdot\mathbf{r} - \omega t + \phi)\\ &= \mathbf{e}_x \frac{1}{2}E_0\{\exp[i\mathbf{k}\cdot\mathbf{r} - \omega t + \phi] + \exp[-i\mathbf{k}\cdot\mathbf{r} - \omega t + \phi]\}\end{aligned} \tag{2.183}$$

where E_0 is the amplitude of the field and ϕ is its phase. Compare this with Eq. (2.127). The equations would be quite similar if the operator $\hat{a}$ could be

factored into polar form. The earliest attempt at such a decomposition appears to be due to Dirac [9] who factored the annihilation and creation operators according to

$$\hat{a} = e^{i\hat{\phi}}\sqrt{\hat{n}} \tag{2.184}$$

$$\hat{a}^\dagger = \sqrt{\hat{n}}\, e^{-i\hat{\phi}} \tag{2.185}$$

where $\hat{\phi}$ was to be interpreted as a Hermitian operator for phase. From the fundamental commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$ it follows that

$$e^{i\hat{\phi}}\hat{n}e^{-i\hat{\phi}} - \hat{n} = 1 \tag{2.186}$$

or

$$e^{i\hat{\phi}}\hat{n} - \hat{n}e^{i\hat{\phi}} = e^{i\hat{\phi}}. \tag{2.187}$$

By expanding the exponentials one can see that Eq. (2.186) is satisfied as long as

$$[\hat{n}, \hat{\phi}] = i. \tag{2.188}$$

It thus appears that number and phase are complementary observables and therefore the fluctuations in these quantities should satisfy the uncertainty relation $\Delta n\, \Delta\phi \geq \frac{1}{2}$.

Unfortunately, things are not so simple. To see that something is quite wrong with the above, consider the matrix element of the commutator for the arbitrary number states $|n\rangle$ and $|n'\rangle$:

$$\langle n'|[\hat{n}, \hat{\phi}]|n\rangle = i\delta_{nn'}. \tag{2.189}$$

Expanding the left side results in

$$(n' - n)\langle n'|\hat{\phi}|n\rangle = i\delta_{nn'} \tag{2.190}$$

which contains an obvious contradiction in the case when $n' = n$ (giving $0 = i$). The Dirac approach fails because of the underlying assumption that a Hermitian phase operator $\hat{\phi}$ actually exists. There are in fact two reasons for the failure of the Dirac approach. If $\hat{\phi}$ exists as a Hermitian operator, then $\exp(i\hat{\phi})$ should be a unitary operator. From Eqs. (2.184) and (2.185) we should have

$$e^{i\phi} = \hat{a}(\hat{n})^{-\frac{1}{2}} \tag{2.191}$$

$$e^{-i\hat{\phi}} = (\hat{n})^{-\frac{1}{2}}\hat{a}^\dagger = (e^{i\hat{\phi}})^\dagger. \tag{2.192}$$

Now

$$(e^{i\hat{\phi}})^\dagger(e^{i\hat{\phi}}) = 1 \tag{2.193}$$

but

$$(e^{i\hat{\phi}})(e^{i\hat{\phi}})^\dagger = \hat{a}\frac{1}{\hat{n}}\hat{a}^\dagger \neq 1. \tag{2.194}$$

So, in fact, $\exp(i\hat{\phi})$ is not a unitary operator from which it follows that $\hat{\phi}$ is not Hermitian. The root of the problem is that the operator $\hat{n}$ has a spectrum bounded from below; it does not include the negative integers. One way to fix this is simply

to include negative integers into the spectrum. The negative number states are of course nonphysical but, as Barnett and Pegg have shown [10], it is possible to construct unitary operators of the form

$$e^{i\hat{\phi}} \equiv \sum_{n=-\infty}^{\infty} |n\rangle\langle n+1| \tag{2.195}$$

$$(e^{i\hat{\phi}})^{\dagger} = e^{-i\hat{\phi}} = \sum_{n=-\infty}^{\infty} |n+1\rangle\,\langle n| \tag{2.196}$$

where it is easy to see that

$$e^{i\hat{\phi}}(e^{i\hat{\phi}})^{\dagger} = (e^{i\hat{\phi}})^{\dagger}e^{i\hat{\phi}} = 1. \tag{2.197}$$

It must be understood that the introduction of the negative number states is only formal and that they are decoupled from the positive (physical) number states. No new predictions arise from including these negative number states. But by using Eq. (2.194) it is easy to show that Eq. (2.186) still holds from which Eq. (2.187) follows and we are still back to the contradiction encountered in Eq. (2.189).

This brings us to the second problem, which is connected with the fact that $\hat{\phi}$ is supposed to be an angle operator. The situation is very similar to the more familiar problem of angular momentum of a particle moving in the x–y plane with angular momentum about the z-axis. If ϕ is the azimuthal angle defined as

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{2.198}$$

(modulo 2π), then the orbital angular momentum is

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = i\,\hbar\frac{\partial}{\partial\phi} \tag{2.199}$$

from which it follows that

$$[\phi, \hat{L}_z] = i. \tag{2.200}$$

But $\hat{L}_z$ is Hermitian only in a space of periodic functions, i.e. wave functions of the form

$$\psi_m(\phi) = \frac{1}{\sqrt{2\pi}}e^{im\phi}, \qquad m = 0, \pm1, \pm2, \ldots \tag{2.201}$$

where

$$\hat{L}_z\psi_m(\phi) = \hbar m\psi_m(\phi). \tag{2.202}$$

(Note that the spectrum of $\hat{L}_z$ contains negative integers.) But ϕ itself is not a periodic function, having the range $-\infty < \phi < \infty$. In fact the fluctuations $\Delta\phi$ can be greater than 2π and for a number state, Eq. (2.199) would seem to imply that $\Delta\phi \to \infty$, a nonsensical behavior. A possible solution to this problem is to introduce a periodic coordinate $\Phi(\phi)$ [10] behaving in a discontinuous fashion according to Fig. 2.4.

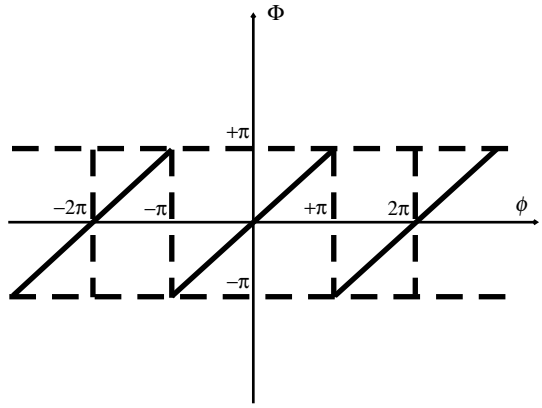


Fig. 2.4. The periodic function Φ, arbitrarily chosen to be discontinuous.

The problem with this is that the commutation relation is no longer of the canonical form but rather is

$$[\Phi, \hat{L}_z] = i \left\{ 1 - 2\pi \sum_{n=-\infty}^{\infty} \delta \left[\phi - (2n+1)\pi \right] \right\}. \tag{2.203}$$

The delta functions arise from the discontinuities of Φ and they will occur no matter where the discontinuities are placed. One of the delta functions would always be present in any particular 2π interval so no sensible uncertainty relation can be formulated.

Over the years, there have been many attempts to create a formalism for the description of the quantum phase that in one way or another overcomes the obstacles discussed above. These schemes have been reviewed extensively in the recent literature. For pedagogical reasons, we first introduce the approach due to Susskind and Glogover [11] and its improvements by Carruthers and Nieto [12] wherein a kind of one-sided "unitary" phase operator (actually an operator analogous to an exponential phase factor) is introduced. Eigenstates of this operator, the phase eigenstates, are introduced and then used to construct a phase distribution for an arbitrary state of the field [13]. Various averages over the phase may be calculated from this distribution. Results obtained from this procedure are identical to those obtained by the scheme of Pegg and Barnett [14] who factor the annihilation operator à la Dirac but in a truncated Hilbert space. After the calculation of expectation values in this truncated space the dimension of the space is allowed to go to infinity. By using the phase eigenstates we shall avoid such calculational complications.

The Susskind–Glogower (SG) operators are defined by the relations

$$\hat{E} \equiv (\hat{n}+1)^{-\frac{1}{2}} \hat{a} = (\hat{a}\hat{a}^\dagger)^{-\frac{1}{2}} \hat{a} \tag{2.204}$$

$$\hat{E}^\dagger = \hat{a}^\dagger (\hat{n}+1)^{-\frac{1}{2}} = \hat{a}^\dagger (\hat{a}\hat{a}^\dagger)^{-\frac{1}{2}} \tag{2.205}$$

where $\hat{E}$ and $\hat{E}^\dagger$ are to be the analogs of the phase factors $\exp(\pm i\phi)$. These $\hat{E}$ operators are sometimes called "exponential" operators and, from context, should not be confused with the field operators. When applied to the number states $|n\rangle$ they yield

$$\begin{aligned} \hat{E}\,|n\rangle &= |n-1\rangle \qquad \text{for } n \neq 0 \\ &= 0 \qquad\qquad\ \ \text{for } n = 0, \end{aligned} \tag{2.206}$$

$$\hat{E}^\dagger|n\rangle = |n+1\rangle. \tag{2.207}$$

From this, it is easy to see that useful and equivalent expressions for these exponential operators are

$$\hat{E} = \sum_{n=0}^{\infty} |n\rangle\langle n+1| \qquad \hat{E}^\dagger = \sum_{n=0}^{\infty} |n+1\rangle\langle n|. \tag{2.208}$$

It is easy to show that

$$\hat{E}\hat{E}^\dagger = \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty} |n\rangle\langle n+1|n'+1\rangle\langle n'| = \sum_{n=0}^{\infty} |n\rangle\langle n| = 1 \tag{2.209}$$

but that

$$\hat{E}^\dagger\hat{E} = \sum_{n=0}^{\infty}\sum_{n=0}^{\infty} |n+1\rangle\langle n|n'\rangle\langle n'+1| = \sum_{n=0}^{\infty} |n+1\rangle\langle n+1| = 1 - |0\rangle\langle 0|. \tag{2.210}$$

The presence of the projection operator $|0\rangle\langle 0|$ spoils the unitarity of $\hat{E}$. But for a state with average photon number $\bar{n} \geq 1$, the contribution from the vacuum state will be small and $\hat{E}$ will be approximately unitary.

Of course, $\hat{E}$ and $\hat{E}^\dagger$ themselves are not observables but the operators

$$\hat{C} \equiv \frac{1}{2}(\hat{E} + \hat{E}^\dagger), \qquad \hat{S} \equiv \frac{1}{2i}(\hat{E} - \hat{E}^\dagger) \tag{2.211}$$

are the obvious analogs of $\cos\phi$ and $\sin\phi$. These operators are Hermitian and satisfy the commutation relation

$$[\hat{C}, \hat{S}] = \frac{1}{2} i\,|0\rangle\langle 0| \tag{2.212}$$

and thus commute for all states but the vacuum. Furthermore, the quantum form of the familiar trigonometric identity becomes

$$\hat{C}^2 + \hat{S}^2 = 1 - \frac{1}{2}|0\rangle\langle 0| \tag{2.213}$$

where once again we see the spoiling effect of the vacuum. It can also be shown that

$$[\hat{C}, \hat{n}] = i\hat{S} \quad \text{and} \quad [\hat{S}, \hat{n}] = -i\hat{C} \tag{2.214}$$

the first of which is similar to Eq. (2.48). The uncertainty relations obeyed in these cases respectively are

$$(\Delta n)(\Delta C) \geq \frac{1}{2}|\langle \hat{S} \rangle| \tag{2.215}$$

and

$$(\Delta n)(\Delta S) \geq \frac{1}{2}|\langle \hat{C} \rangle| \tag{2.216}$$

where the Δ means the root-mean-square deviation.

In the case of number states $|n\rangle$, $\Delta n = 0$,

$$\langle n|\, \hat{C}\, |n\rangle = \langle n|\, \hat{S}\, |n\rangle = 0 \tag{2.217}$$

and

$$\langle n|\, \hat{C}^2\, |n\rangle = \langle n|\, \hat{S}^2\, |n\rangle = \begin{cases} \frac{1}{2}, & n \geq 1 \\ \frac{1}{4}, & n = 0 \end{cases}. \tag{2.218}$$

Thus for $n \geq 1$ the uncertainties in $\hat{C}$ and $\hat{S}$ are

$$\Delta C = \Delta S = \frac{1}{\sqrt{2}} \tag{2.219}$$

which would seem to correspond to a phase angle equally likely to have any value in the range 0 to 2π. (Note that this will not be true for the vacuum state!) In any case, the right-hand sides of Eqs. (2.214) and (2.215) are zero for a number state as required in order that the uncertainty relations be satisfied.

The eigenstates $|\phi\rangle$ of the exponential operator satisfying the eigenvalue equation

$$\hat{E}\, |\phi\rangle = e^{i\phi}\, |\phi\rangle \tag{2.220}$$

are given by

$$|\phi\rangle = \sum_{n=0}^{\infty} e^{in\phi}\, |n\rangle\,. \tag{2.221}$$

These states are not normalizable nor are they orthogonal as the scalar product of $|\phi\rangle$ and $|\phi'\rangle$ is not the delta function $\delta(\phi - \phi')$. By virtue of the fact that

$$\int_0^{2\pi} e^{i(n-n')\phi} d\phi = 2\pi\, \delta_{nn'} \tag{2.222}$$

it is easy to show that the phase eigenstates resolve to unity according to

$$\frac{1}{2\pi} \int_0^{2\pi} d\phi\, |\phi\rangle\, \langle\phi| = 1. \tag{2.223}$$

The expectation values of the $\hat{C}$ and $\hat{S}$ operators obviously take the form of $\cos\phi$ and $\sin\phi$ respectively.

Now an arbitrary state $|\psi\rangle$ of the field will be given as a superposition of all the number states, i.e.

$$|\psi\rangle = \sum_{n=0}^{\infty} C_n |n\rangle \tag{2.224}$$

where the coefficients C_n must satisfy

$$\sum_{n=0}^{\infty} |C_n|^2 = 1 \tag{2.225}$$

in order that $|\psi\rangle$ be normalized. We may associate a phase distribution $\mathcal{P}(\phi)$ with the state $|\psi\rangle$ according to the prescription

$$\begin{aligned} \mathcal{P}(\phi) &\equiv \frac{1}{2\pi} |\langle \phi \mid \psi \rangle|^2 \\ &= \frac{1}{2\pi} \left| \sum_{n=0}^{\infty} e^{-in\phi} C_n \right|^2 . \end{aligned} \tag{2.226}$$

Clearly, $\mathcal{P}(\phi)$ is always positive and can be shown to be normalized as follows: writing

$$\mathcal{P}(\phi) \equiv \frac{1}{2\pi} |\langle \phi \mid \psi \rangle|^2 = \frac{1}{2\pi} \langle \phi \mid \psi \rangle \langle \psi \mid \phi \rangle \tag{2.227}$$

and then integrating over ϕ using the resolution of unity as given by Eq. (2.223) we have

$$\int_0^{2\pi} \mathcal{P}(\phi)\, d\phi = \langle \psi \mid \psi \rangle = 1. \tag{2.228}$$

More generally, for a state described by a density operator $\hat{\rho}$, we would have

$$\mathcal{P}(\phi) = \frac{1}{2\pi} \langle \phi | \hat{\rho} | \phi \rangle. \tag{2.229}$$

The distribution can be used to calculate the average of any function of ϕ, $f(\phi)$ according to

$$\langle f(\phi) \rangle = \int_0^{2\pi} f(\phi) \mathcal{P}(\phi)\, d\phi. \tag{2.230}$$

This is essentially the same as would be obtained from the Pegg–Barnett formalism of Reference [14].

For the present we consider only a number state $|n\rangle$ for which all the coefficients in Eq. (2.224) vanish but for $C_n = 1$. This leads to

$$\mathcal{P}(\phi) = \frac{1}{2\pi} \tag{2.231}$$

which is uniform, as expected. Furthermore, the average of ϕ is π and the average over ϕ^2 is $4\pi^2/3$ and thus the fluctuations in ϕ are

$$\Delta\phi = \sqrt{\langle \phi^2 \rangle - \langle \phi \rangle^2} = \frac{\pi}{\sqrt{3}} \tag{2.232}$$

as we would expect for a uniform probability distribution over the range 0 to 2π. Note that this applies to all number states, including the vacuum. We shall

apply this formalism to other field states, in particular the coherent states, in later chapters.

One difficulty with the attempt to find a sensible quantum description of phase is that it is not obvious how to connect the various formalisms for phase with realistic experiments on the measurement of phase. The approach taken here is supported by the work of Shapiro and Shepard [15] who have shown, by using quantum estimation theory [16], that phase states $|\phi\rangle$ generate the probability operator measure for maximum likelihood phase estimation. On the other hand, Mandel and co-workers [17] have taken an operational approach to phase operators based on classical phase measurements. They show that phase measurements are difficult even in classical optics and that these difficulties carry over into quantum optics and, further, that different measurement procedures lead to different phase operators. It does not appear, at least from the operational point of view, that there exists one correct phase operator. Henceforth we shall use the eigenstates of the SG operator under the assumption that the corresponding phase distribution, at the very least, qualitively, if not quantitatively, describes a quantum phase variable.

Problems

1. For the single-mode field given by Eq. (2.5), use Maxwell's equations to obtain the corresponding magnetic field given by Eq. (2.6).
2. For the single-mode field of the previous problem, obtain the Hamiltonian and show that it has the form of a simple harmonic oscillator.
3. Give a proof of the Baker–Hausdorf lemma: for any two operators $\hat{A}$ and $\hat{B}$,

$$e^{i\lambda\hat{A}}\hat{B}\,e^{-i\lambda\hat{A}} = \hat{B} + i\lambda[\hat{A},\hat{B}] + \frac{(i\lambda)^2}{2!}[\hat{A},[\hat{A},\hat{B}]] + \cdots$$

4. In the special case where $[\hat{A},\hat{B}] \neq 0$, but where $[\hat{A},[\hat{A},\hat{B}]] = 0 = [\hat{B},[\hat{A},\hat{B}]]$, show that

$$e^{\hat{A}+\hat{B}} = \exp\left(-\frac{1}{2}[\hat{A},\hat{B}]\right)e^{\hat{A}}e^{\hat{B}} = \exp\left(\frac{1}{2}[\hat{A},\hat{B}]\right)e^{\hat{B}}e^{\hat{A}}.$$

   This is known as the Baker–Hausdorf–Campbell theorem.
5. Suppose the state of a single-mode cavity field is given at time $t = 0$ by

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|n\rangle + e^{i\varphi}\,|n+1\rangle)$$

   where φ is some phase. Find the state $|\psi(t)\rangle$ at times $t > 0$. For this time-evolved state, verify the uncertainty relation of Eq. (2.49).
6. Consider the superposition state $|\psi_{01}\rangle = \alpha|0\rangle + \beta|1\rangle$ where α and β are complex and satisfy $|\alpha|^2 + |\beta|^2 = 1$. Calculate the variances of the quadrature operators $\hat{X}_1$ and $\hat{X}_2$. Are there any values of the parameters α and β for which either of the quadrature

variances become *less* than for a vacuum state? If so, check to see if the uncertainty principle is violated. Repeat with the state $|\psi_{02}\rangle = \alpha|0\rangle + \beta|2\rangle$.

7. Many processes involve the absorption of single photons from a quantum field state, the process of absorption being represented by the action of the annihilation operator $\hat{a}$. For an arbitrary field state $|\psi\rangle$, the absorption of a single photon yields the state $|\psi'\rangle \sim \hat{a}|\psi\rangle$. Normalize this state. Compare the average photon numbers $\bar{n}$ of the state $|\psi\rangle$ and $\bar{n}'$ of $|\psi'\rangle$. Do you find that $\bar{n}' = \bar{n} - 1$?
8. Consider the superposition of the vacuum and 10 photon number state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |10\rangle\right).$$

Calculate the average photon number for this state. Next assume that a single photon is absorbed and recalculate the average photon number. Does your result seem sensible in comparison with your answer to the previous question?

9. Consider the multimode expressions for the electric and magnetic fields of Eqs. (2.83) and (2.84) respectively. (a) Show that they satisfy the free-space Maxwell equations. (b) Use these fields and follow through the derivation of the field energy given by Eq. (2.95).
10. It is sometimes useful to characterize the photon number probability distribution P_n by its factorial moments. The rth factorial moment is defined as

$$\begin{aligned}&\langle \hat{n}\,(\hat{n}-1)\,(\hat{n}-2)\ldots(\hat{n}-r+1)\rangle \\ &\quad = \sum_n n\,(n-1)\,(n-2)\ldots(n-r+1)\,P_n.\end{aligned}$$

Show that for a thermal field, the right-hand side has the value $r!\bar{n}^r$.

11. Work out the commutator of the cosine and sine operators, $\hat{C}$ and $\hat{S}$ respectively. Calculate the matrix elements of the commutator and show that only the diagonal ones are nonzero.
12. Consider the mixed state (see Appendix A) described by the density operator

$$\hat{\rho} = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right)$$

and the pure superposition state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle).$$

Calculate the corresponding phase distributions $\mathcal{P}(\phi)$ and compare them.

13. Show that for the thermal state, $\mathcal{P}(\phi) = 1/2\pi$.

References

[1] See, for example, J. J. Sakurai, *Modern Quantum Mechanics*, revised edition (Reading: Addison-Wesley, 1994), p. 96.
[2] See the discussion in I. J. K. Aitchison, *Contemp. Phys.*, **26** (1985), 333.
[3] W. E. Lamb, Jr., and R. C. Retherford, *Phys. Rev.*, **72** (1947), 241.
[4] H. A. Bethe, *Phys. Rev.*, **72** (1947), 339.

[5] T. A. Welton, *Phys. Rev*., **74** (1948), 1157.
[6] H. B. G. Casimir, *Proc. K. Ned. Akad. Wet*., **51** (1948), 793; *Physica*, **19** (1953), 846.
[7] P. W. Milonni and M.-L. Shih, *Contemp. Phys*., **33** (1992), 313.
[8] M. J. Sparnaay, *Nature*, **180** (1957), 334.
[9] P. A. M. Dirac, *Proc. R. Soc. Lond.*, A **114** (1927), 243.
[10] S. M. Barnett and D. T. Pegg, *J. Phys. A*, **19** (1986), 3849.
[11] L. Susskind and J. Glogower, *Physics*, **1** (1964), 49.
[12] P. Carruthers and M. M. Nieto, *Rev. Mod. Phys*., **40** (1968), 411.
[13] See G. S. Agarwal, S. Chaturvedi, K. Tara and V. Srinivasan, *Phys. Rev. A*, **45** (1992), 4904.
[14] D. T. Pegg and S. M. Barnett, *Europhys. Lett*., **6** (1988), 483; *Phys. Rev. A*, **39** (1989), 1665; **43** (1991), 2579.
[15] J. H. Shapiro and S. R. Shepard, *Phys. Rev. A*, **43** (1991), 3795.
[16] C. W. Helstrom, *Quantum Detection and Estimation Theory* (New York: Academic Press, 1976).
[17] J. W. Noh, S. Fougères and L. Mandel, *Phys. Rev. A*, **45** (1992), 424; **46** (1992), 2840.

Bibliography

Quantum mechanics

Virtually every quantum mechanics textbook covers the harmonic oscillator through the operator method, sometimes known as the algebraic, or factorization, method. Some good undergraduate textbooks are as follows.

R. L. Liboff, *Introductory Quantum Mechanics,* 4th edition (Reading: Addison-Wesley, 2002).
D. J. Griffiths, *Introduction to Quantum Mechanics* (Englewood Cliffs: Prentice Hall, 1995).
J. S. Townsend, *A Modern Approach to Quantum Mechanics* (New York: McGraw-Hill, 1992).

The following is an excellent graduate level text.

J. J. Sakurai, *Modern Quantum Mechanics*, revised edition (Reading: Addison-Wesley, 1994).

Field theory

E. A. Power, *Introductory Quantum Electrodynamics* (New York: American Elsevier, 1965).
E. G. Harris, *A Pedestrian Approach to Quantum Field Theory* (New York: Wiley, 1972).
L. H. Ryder, *Quantum Field Theory*, 2nd edition (Cambridge: Cambridge University Press, 1996).

For an extensive set of references on the Casimir effect see the following.

S. K. Lamoreaux, “Resource Letter CF-1: Casimir Force”, *Am. J. Phys*., **67** (1999), 850.

Chapter 3
Coherent states

At the end of the preceding chapter, we showed that the photon number states $|n\rangle$ have a uniform phase distribution over the range 0 to 2π. Essentially, then, there is no well-defined phase for these states and, as we have already shown, the expectation value of the field operator for a number state vanishes. It is frequently suggested (see, for example, Sakurai [1]) that the classical limit of the quantized field is the limit in which the number of photons becomes very large such that the number operator becomes a continuous variable. However, this cannot be the whole story since the mean field $\langle n|\hat{E}_x|n\rangle = 0$ no matter how large the value of n. We know that at a fixed point in space a classical field oscillates sinusoidally in time. Clearly this does not happen for the expectation value of the field operator for a number state. In this chapter we present a set of states, the coherent states [2], which do give rise to a sensible classical limit; and, in fact, these states are the "most classical" quantum states of a harmonic oscillator, as we shall see.

3.1 Eigenstates of the annihilation operator and minimum uncertainty states

In order to have a non-zero expectation value of the electric field operator or, equivalently, of the annihilation and creation operators, we are required to have a superposition of number states differing only by ± 1. For example, this could contain only the states $|n\rangle$ and $|n \pm 1\rangle$:

$$|\psi\rangle = C_n|n\rangle + C_{n+1}|n \pm 1\rangle \tag{3.1}$$

(where $|C_n|^2 + |C_{n+1}|^2 = 1$). Generally, a superposition of all the number states will have the property that the expectation value of say, $\hat{a}$, will not vanish. Clearly, by inspecting Eqs. (2.15) and (2.16), the replacement of $\hat{a}$ and $\hat{a}^\dagger$ by continuous variables produces a classical field. A unique way to make this replacement is to seek eigenstates of the annihilation operator. These states are denoted as $|\alpha\rangle$ and satisfy the relation

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \tag{3.2}$$

where α is a complex number, otherwise arbitrary. (Note that complex eigenvalues are allowed here as $\hat{a}$ is non-Hermitian.) The states $|\alpha\rangle$ are "right" eigenstates of $\hat{a}$ and $\langle\alpha|$ are "left" eigenstates of $\hat{a}^\dagger$ with eigenvalue α^*:

$$\langle\alpha|\,\hat{a}^\dagger = \alpha^*\,\langle\alpha|\,. \tag{3.3}$$

Since the number states $|n\rangle$ form a complete set we may expand $|\alpha\rangle$ according to

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_n\,|n\rangle\,. \tag{3.4}$$

Acting with $\hat{a}$ on each term of the expansion, Eq. (3.2) becomes

$$\hat{a}\,|\alpha\rangle = \sum_{n=1}^{\infty} C_n\sqrt{n}\,|n-1\rangle = \alpha\sum_{n=0}^{\infty} C_n\,|n\rangle\,. \tag{3.5}$$

Equating coefficients of $|n\rangle$ on both sides leads to

$$C_n\sqrt{n} = \alpha C_{n-1} \tag{3.6}$$

or

$$\begin{aligned} C_n &= \frac{2}{\sqrt{n}}C_{n-1} = \frac{\alpha^2}{\sqrt{n(n-1)}}C_{n-2} = \ldots \\ &= \frac{\alpha^n}{\sqrt{n!}}C_0 \end{aligned} \tag{3.7}$$

and thus

$$|\alpha\rangle = C_0\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,. \tag{3.8}$$

From the normalization requirement we determine C_0:

$$\begin{aligned} \langle\alpha\mid\alpha\rangle = 1 &= |C_0|^2\sum_n\sum_{n'}\frac{\alpha^{*n}\alpha^{n'}}{\sqrt{n!n'!}}\,\langle n\mid n'\rangle \\ &= |C_0|^2\sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!} = |C_0|^2\,e^{|\alpha|^2} \end{aligned} \tag{3.9}$$

which implies that $C_0 = \exp(-\frac{1}{2}|\alpha|^2)$. Thus our normalized coherent states are

$$|\alpha\rangle = \exp\left(-\tfrac{1}{2}\,|\alpha|^2\right)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,. \tag{3.10}$$

Let us now consider the expectation value of the electric field operator

$$\hat{E}_x(\mathbf{r},t) = i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{\frac{1}{2}}\left[\hat{a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - \hat{a}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right]. \tag{3.11}$$

We obtain

$$\langle\alpha|\,\hat{E}_x(\mathbf{r},t)\,|\alpha\rangle = i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{\frac{1}{2}}\left[\alpha e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - \alpha^* e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right]. \tag{3.12}$$

Writing α in polar form, $\alpha = |\alpha| e^{i\theta}$, we have

$$\langle \alpha | \hat{E}_x(\mathbf{r}, t) | \alpha \rangle = 2 |\alpha| \left(\frac{\hbar\omega}{2\varepsilon_0 V} \right)^{\frac{1}{2}} \sin(\omega t - \mathbf{k} \cdot \mathbf{r} - \theta) \tag{3.13}$$

which looks like a classical field. Furthermore, we can show that

$$\langle \alpha | \hat{E}_x^2(\mathbf{r}, t) | \alpha \rangle = \frac{\hbar\omega}{2\varepsilon_0 V} [1 + 4 |\alpha|^2 \sin^2 (\omega t - \mathbf{k} \cdot \mathbf{r} - \theta)] . \tag{3.14}$$

Thus the fluctuations in $\hat{E}_x(\mathbf{r}, t)$,

$$\Delta E_x \equiv \langle (\Delta \hat{E}_x)^2 \rangle^{\frac{1}{2}} = \left(\frac{\hbar\omega}{2\varepsilon_0 V} \right)^{\frac{1}{2}} , \tag{3.15}$$

are identical to those for a vacuum state. The coherent state is nearly a classical-like state because it not only yields the correct form for the field expectation values but contains only the noise of the vacuum. Indeed, using the quadrature operators of Eqs. (2.52) and (2.53), it is an easy exercise to show that for the coherent states

$$\langle (\Delta \hat{X}_1)^2 \rangle_\alpha = \frac{1}{4} = \langle (\Delta \hat{X}_2)^2 \rangle_\alpha \tag{3.16}$$

which again shows that these states have the fluctuations of the vacuum. Thus with respect to the field quadrature operators, the coherent states both minimize (actually equalize) the uncertainty product and exhibit equal uncertainties, those of the vacuum, in each quadrature. As a matter of fact, this property can be used as an alternate definition for the coherent states. Consider three Hermitian operators $\hat{A}$, $\hat{B}$, and $\hat{C}$, satisfying

$$[\hat{A}, \hat{B}] = i\hat{C} \tag{3.17}$$

which implies the uncertainty relation

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \frac{1}{4} \langle (\hat{C})^2 \rangle . \tag{3.18}$$

States that equalize this relation are those that satisfy the eigenvalue equation [3]

$$\left[\hat{A} + \frac{i\langle \hat{C} \rangle}{2\langle (\Delta \hat{B})^2 \rangle} \hat{B} \right] |\psi\rangle = \left[\langle \hat{A} \rangle + \frac{i\langle \hat{C} \rangle}{2\langle (\Delta \hat{B})^2 \rangle} \langle \hat{B} \rangle \right] |\psi\rangle . \tag{3.19}$$

The states $|\psi\rangle$ satisfying Eq. (3.19) sometimes go by the name "intelligent" states [4]. For the case where

$$\langle (\Delta \hat{A})^2 \rangle = \langle (\Delta \hat{B})^2 \rangle = \frac{1}{4} \langle \hat{C} \rangle , \tag{3.20}$$

the eigenvalue equation becomes

$$[\hat{A} + i\hat{B}] |\psi\rangle = [\langle \hat{A} \rangle + i \langle \hat{B} \rangle] \, |\psi\rangle . \tag{3.21}$$

For $\hat{A} = \hat{X}_1$ and $\hat{B} = \hat{X}_2$ this is equivalent to Eq. (3.2) with $\alpha = \langle \hat{X}_1 \rangle + i \langle \hat{X}_2 \rangle$.

What is the physical meaning of the complex parameter α? From Eq. (3.13) it is apparent that $|\alpha|$ is related to the amplitude of the field. The expectation value

of the photon number operator $\hat{n} = \hat{a}^\dagger \hat{a}$ is

$$\bar{n} = \langle \alpha | \, \hat{n} \, | \alpha \rangle = |\alpha|^2 \tag{3.22}$$

and thus $|\alpha|^2$ is just the average photon number of the field. To calculate the fluctuations of the photon number we need to calculate

$$\begin{aligned} \langle \alpha | \, \hat{n}^2 \, | \alpha \rangle &= \langle \alpha | \, \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a} \, | \alpha \rangle \\ &= \langle \alpha | \, \left(\hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} + \hat{a}^\dagger \hat{a} \right) | \alpha \rangle \\ &= |\alpha|^4 + |\alpha|^2 = \bar{n}^2 + \bar{n} \end{aligned} \tag{3.23}$$

and thus

$$\Delta n = \sqrt{\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2} = \bar{n}^{1/2} \tag{3.24}$$

which is characteristic of a Poisson process. In fact, for a measurement of the number of photons in the field, the probability of detecting n photons is

$$\begin{aligned} P_n &= |\langle n \mid \alpha \rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \\ &= e^{-\bar{n}} \frac{\bar{n}^n}{n!} \end{aligned} \tag{3.25}$$

which is a Poisson distribution with a mean of $\bar{n}$. Note that the fractional uncertainty in the photon number is

$$\frac{\Delta n}{\bar{n}} = \frac{1}{\sqrt{\bar{n}}} \tag{3.26}$$

which decreases with increasing $\bar{n}$. In Fig. 3.1 we plot a couple of examples of the photon number probability distribution for different $\bar{n}$.

Let us now look at the phase distribution of the coherent states. For a coherent state $|\alpha\rangle$, with $\alpha = |\alpha| \, e^{i\theta}$, the corresponding phase distribution is

$$\begin{aligned} \mathcal{P}(\varphi) &= \frac{1}{2\pi} \, |\langle \varphi \mid \alpha \rangle|^2 \\ &= \frac{1}{2\pi} \, e^{-|\alpha|^2} \left| \sum_{n=0}^{\infty} e^{in(\varphi - \theta)} \frac{|\alpha|^n}{\sqrt{n}\,!} \right|^2 . \end{aligned} \tag{3.27}$$

For large $|\alpha|^2$, the Poisson distribution may be approximated as a Gaussian [5]:

$$e^{-|\alpha|^2/2} \frac{|\alpha|^{2n}}{n!} \, e^{-|\alpha|^2} \approx (2\pi \, |\alpha|^2)^{-\frac{1}{2}} \exp\left[-\frac{(n - |\alpha|^2)^2}{2 \, |\alpha|^2} \right] \tag{3.28}$$

so that the sum in Eq. (3.27) may be evaluated to obtain an approximate form for $\mathcal{P}(\varphi)$ as

$$\mathcal{P}(\varphi) \approx \left(\frac{2 \, |\alpha|^2}{\pi} \right)^{\frac{1}{2}} \exp[-2 \, |\alpha|^2 \, (\varphi - \theta)^2] \, . \tag{3.29}$$

This is a Gaussian peaked at $\varphi = \theta$. Furthermore, the peak becomes narrower with increasing $\bar{n} = |\alpha|^2$, as illustrated in Fig. 3.2.

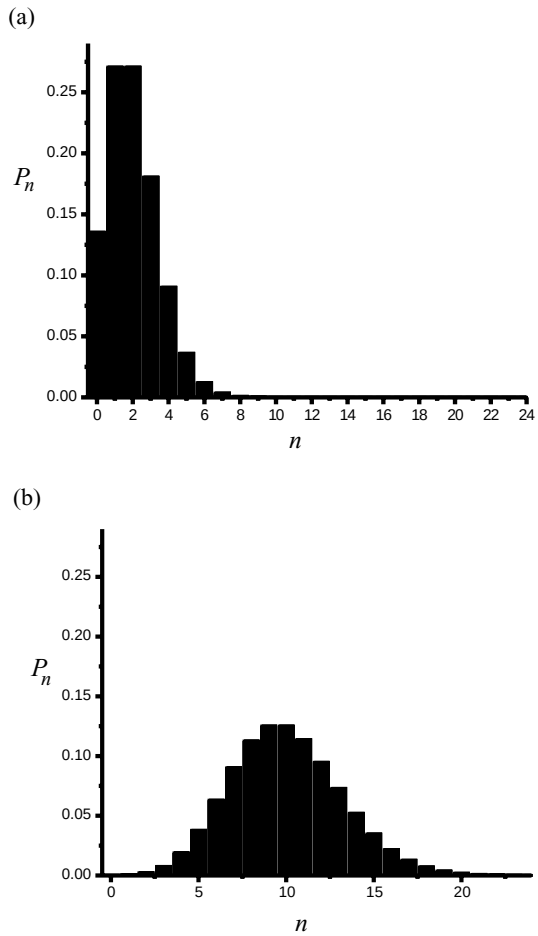


Fig. 3.1. Coherent state photon number probability distributions for (a) $\bar{n} = 2$ and (b) $\bar{n} = 10$.

The coherent states $|\alpha\rangle$ are quantum states very close to classical states because (i) the expectation value of the electric field has the form of the classical expression, (ii) the fluctuations in the electric field variables are the same as for a vacuum, (iii) the fluctuations in the fractional uncertainty for the photon number decrease with the increasing average photon number, and (iv) the states become well localized in phase with increasing average photon number. However, in spite of their near-classical properties, they are still quantum states. In Fig. 3.3 we illustrate this with a sketch where the expectation of the field operator and its fluctuations are plotted against time. The field clearly has the classical form of a sine wave but with the quantum fluctuations superimposed, indicating that it does have quantum features. In fact, *all* states of light must have some quantum features as the quantum theory of light is more fundamental than the classical theory. However, the quantum features of light are generally difficult to observe. We shall deal with some of these features in Chapter 7.

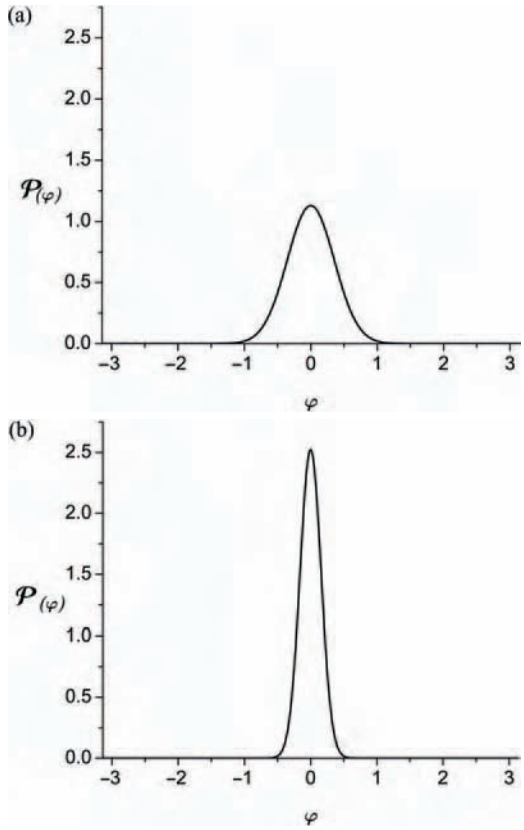


Fig. 3.2. Phase distributions for coherent states with $\theta = 0$ for (a) $\bar{n} = 2$ and (b) $\bar{n} = 10$.

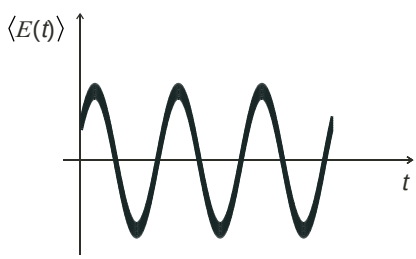


Fig. 3.3. Coherent state expectation value of the electric field as a function of time for a fixed position showing the quantum fluctuations. The fluctuations of the field are the same at all times such that the field is as close to a classical field as is possible for any quantum state.

3.2 Displaced vacuum states

We have discussed two ways in which the coherent states may be defined: as right eigenstates of the annihilation operator and as states that minimize the uncertainty relation for the two orthogonal field quadratures with equal uncertainties (identical to those of a vacuum state) in each quadrature. There is, in fact, a third

definition that leads to equivalent states. This involves the displacement of the vacuum. As we will show, this is closely related to a mechanism for generating the coherent states from classical currents.

The displacement operator $\hat{D}(\alpha)$ is defined as [2]

$$\hat{D}(\alpha) = \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a}) \tag{3.30}$$

and the coherent states are given as

$$|\alpha\rangle = \hat{D}(\alpha)\,|0\rangle\,. \tag{3.31}$$

To see this, consider the identity (the disentangling theorem)

$$\begin{aligned} e^{\hat{A}+\hat{B}} &= e^{\hat{A}} e^{\hat{B}} e^{-\frac{1}{2}[\hat{A},\hat{B}]} \\ &= e^{\hat{B}} e^{\hat{A}} e^{\frac{1}{2}[\hat{A},\hat{B}]} \end{aligned} \tag{3.32}$$

valid if $[\hat{A}, \hat{B}] \neq 0$ but where also

$$[\hat{A}, [\hat{A}, \hat{B}]] = [\hat{B}, [\hat{A}, \hat{B}]] = 0. \tag{3.33}$$

With $\hat{A} = \alpha\,\hat{a}^\dagger$ and $\hat{B} = -\alpha^* \hat{a}$, $[\hat{A}, \hat{B}] = |\alpha|^2$ and Eq. (3.33) holds. Thus

$$\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}} = e^{-\frac{1}{2}|\alpha|^2} e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}}\,. \tag{3.34}$$

Expanding $\exp(-\alpha^* \hat{a})$ makes it clear that

$$e^{-\alpha^* \hat{a}}\,|0\rangle = \sum_{l=0}^{\infty} \frac{(-\alpha^* \hat{a})^l}{l!}\,|0\rangle = |0\rangle \tag{3.35}$$

since $\hat{a}^l |0\rangle = 0$ except for $l = 0$. But

$$\begin{aligned} e^{\alpha \hat{a}^\dagger}\,|0\rangle &= \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} (\hat{a}^\dagger)^n\,|0\rangle \\ &= \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle \end{aligned} \tag{3.36}$$

where we have used the fact that $(\hat{a}^\dagger)^n |0\rangle = \sqrt{n!}|n\rangle$. Thus we have

$$\begin{aligned} |\alpha\rangle &= \hat{D}(\alpha)\,|0\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\,, \end{aligned} \tag{3.37}$$

in agreement with our previous definitions.

The displacement operator $\hat{D}$ is, of course, a unitary operator. It can be shown that

$$\begin{aligned} \hat{D}^\dagger(\alpha) &= \hat{D}(-\alpha) \\ &= e^{-\frac{1}{2}|\alpha|^2} e^{-\alpha\,\hat{a}^\dagger} e^{\alpha^* \hat{a}}. \end{aligned} \tag{3.38}$$

An alternate representation of $\hat{D}(\alpha)$ is

$$\hat{D}(\alpha) = e^{\frac{1}{2}|\alpha|^2} e^{-\alpha^* \hat{a}} e^{\alpha\,\hat{a}^\dagger}. \tag{3.39}$$

Evidently

$$\hat{D}(\alpha)\hat{D}^{\dagger}(\alpha) = \hat{D}^{\dagger}(\alpha)\hat{D}(\alpha) = 1 \tag{3.40}$$

as required for unitarity.

The displacement operator obeys the semigroup relation: the product of two displacement operators, say of $\hat{D}(\alpha)$ and $\hat{D}(\beta)$, is, up to an overall phase factor, the displacement operator $\hat{D}(\alpha+\beta)$. To see this let $\hat{A} = \alpha\hat{a}^{\dagger} - \alpha^{*}\hat{a}$, $\hat{B} = \beta\hat{a}^{\dagger} - \beta^{*}\hat{a}$ where

$$[\hat{A}, \hat{B}] = \alpha\beta^{*} - \alpha^{*}\beta = 2i\,\mathrm{Im}(\alpha\beta^{*}). \tag{3.41}$$

Then, by using Eq. (3.32),

$$\begin{aligned}\hat{D}(\alpha)\,\hat{D}(\beta) &= e^{\hat{A}}e^{\hat{B}} \\ &= \exp\left[i\,\mathrm{Im}\,(\alpha\beta^{*})\right] \\ &\quad \times \exp[(\alpha+\beta)\,\hat{a}^{\dagger} - (\alpha^{*}+\beta^{*})\,\hat{a}^{\dagger}] \\ &= \exp[i\,\mathrm{Im}(\alpha\beta^{*})]\hat{D}(\alpha+\beta)\ .\end{aligned} \tag{3.42}$$

Thus, applied to the vacuum

$$\begin{aligned}\hat{D}(\alpha)\,\hat{D}(\beta) &= \hat{D}(\alpha)\,|\beta\rangle \\ &= \exp[i\,\mathrm{Im}\,(\alpha\beta^{*})]\ |\alpha+\beta\rangle\ .\end{aligned} \tag{3.43}$$

The phase factor $\exp[i\,\mathrm{Im}(\alpha\beta^{*})]$ is an overall phase factor and so is physically irrelevant.

3.3 Wave packets and time evolution

From Eqs. (2.13) and (2.14) we obtain the "position" operator

$$\hat{q} = \sqrt{\frac{\hbar}{2\omega}}(\hat{a} + \hat{a}^{\dagger}) = \sqrt{\frac{2\,\hbar}{\omega}}\,\hat{X}_1, \tag{3.44}$$

where $\hat{X}_1$ is the quadrature operator of Eq. (2.52) The eigenstates of the operator $\hat{q}$ we denote as $|q\rangle$ where $\hat{q}\,|q\rangle = q\,|q\rangle$. The corresponding wave functions for the number states are [6]

$$\psi_n(q) = \langle q \mid n\rangle = (2^n n!)^{-1/2}\left(\frac{\omega}{\pi\,\hbar}\right)^{1/4}\exp(-\xi^2/2)H_n(\xi) \tag{3.45}$$

where $\xi = q\sqrt{\omega/\hbar}$ and where the $H_n(\xi)$ are Hermite polynomials. The corresponding wave function for the coherent state is then

$$\psi_{\alpha}(q) \equiv \langle q \mid \alpha\rangle = \left(\frac{\omega}{\pi\,\hbar}\right)^{1/4} e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{(\alpha/\sqrt{2})^n}{n!}H_n(\xi). \tag{3.46}$$

From the generating function for Hermite polynomials, we can obtain the closed-form expression for the coherent-state wave function

$$\psi_{\alpha}(q) = \left(\frac{\omega}{\pi\,\hbar}\right)^{1/4} e^{-|\alpha|^2/2}e^{\xi^2/2}e^{-(\xi-\alpha/\sqrt{2})^2}, \tag{3.47}$$

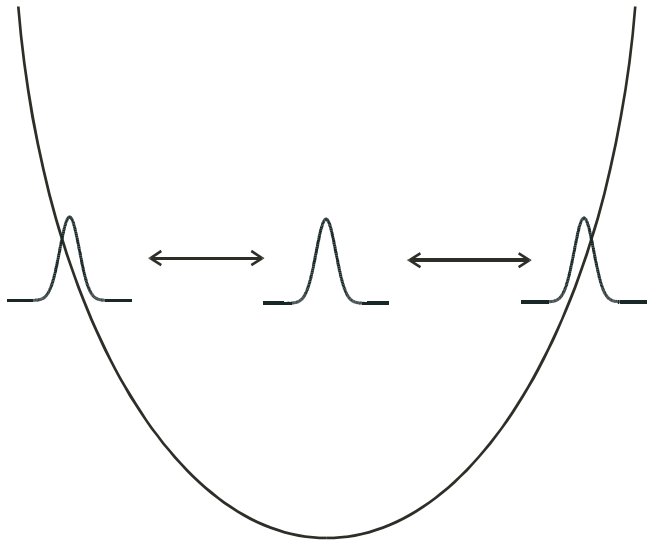


Fig. 3.4. A coherent-state wave function moves through the harmonic oscillator potential, between the classical turning points, without dispersion.

a Gaussian wave function. Of course, the probability distribution over the "position" variable q,

$$P(q) = |\psi_\alpha(q)|^2 , \tag{3.48}$$

is also Gaussian.

We now consider the time evolution of a coherent state for a single-mode free field where the Hamiltonian is given by Eq. (2.18). The time-evolving coherent state is given by

$$\begin{aligned} |\alpha, t\rangle &\equiv \exp(-i\hat{H}t/\hbar)\,|\alpha\rangle = e^{-i\omega t/2} e^{-i\omega t\,\hat{n}}\,|\alpha\rangle \\ &= e^{-i\omega t/2} |\alpha\, e^{-i\,\omega t}\rangle, \end{aligned} \tag{3.49}$$

and so the coherent state remains a coherent state under free field evolution. The corresponding time-evolving wave function is

$$\psi_\alpha(q,t) = \left(\frac{\omega}{\pi\,\hbar}\right)^{1/4} e^{-|\alpha|^2/2} e^{\xi^2/2} e^{-(\xi - \alpha\, e^{-i\,\omega t}/\sqrt{2})^2}, \tag{3.50}$$

a Gaussian whose shape does not change with time and whose centroid follows the motion of a classical point particle in a harmonic oscillator potential (the demonstration of this is left as an exercise, e.g. Problem 2 at the end of this chapter). In Fig. 3.4 we illustrate the motion of a coherent-state wave packet in the harmonic oscillator potential. The motional states of laser-cooled trapped atoms or ions can be engineered to have this minimum uncertainty character as we shall see in Chapter 10.

Ultimately, the stability of the wave packet results from the fact that, for the harmonic oscillator, the energy levels are integer spaced. For quantum systems where the energy levels are not integer spaced, such as for the Coulomb problem, formulating coherent states is a bit of a challenge and no such states are truly

stable for all times. Consideration of such states is beyond the scope of this book, but see Reference [7].

3.4 Generation of coherent states

A coherent state may be generated by a classical oscillating current. Let the quantum electromagnetic vector potential be $\hat{\mathbf{A}}(\mathbf{r}, t)$ for a field interacting with a classical current described by the current density $\mathbf{j}(\mathbf{r}, t)$. According to classical electromagnetic theory, the interaction energy $\hat{V}(t)$ is given by

$$\hat{V}(t) = \int d^3\mathbf{r}\, \mathbf{j}\,(\mathbf{r}, t) \cdot \hat{\mathbf{A}}\,(\mathbf{r}, t)\,. \tag{3.51}$$

For a single-mode field, $\hat{\mathbf{A}}(\mathbf{r}, \mathrm{t})$ is given, in the interaction picture, by

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \mathbf{e}\left(\frac{\hbar}{2\omega_0\varepsilon_0 V}\right)^{\frac{1}{2}} \left[\hat{a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \hat{a}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\right] \tag{3.52}$$

where $\hat{a}$ means $\hat{a}(0)$. Substituting this into Eq. (3.51), we have

$$\hat{V}(t) = -\left(\frac{\hbar}{2\omega\varepsilon_0 V}\right)^{\frac{1}{2}} \left[\hat{a}\mathbf{e}\cdot\mathbf{J}\,(\mathbf{k}, t)\, e^{-i\omega t} + \hat{a}^\dagger \mathbf{e}\cdot\mathbf{J}^*\,(\mathbf{k}, t)\; e^{i\omega t}\right] \tag{3.53}$$

where

$$\mathbf{J}\,(\mathbf{k}, t) = \int d^3\mathbf{r}\, \mathbf{j}(\mathbf{r}, t) e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{3.54}$$

Since $\hat{V}(t)$ depends on time, the associated evolution operator is a time-ordered product [8]. But for an infinitesimally short time interval from t to $t + \delta t$, the evolution operator is

$$\begin{aligned} \hat{U}(t + \delta t, t) &\cong \exp[-i\,\hat{V}\,(t)\,\delta t/\hbar] \\ &= \exp\{-\delta t[u(t)\hat{a} - u^*(t)\hat{a}^\dagger]\} \\ &= \hat{D}\,[u(t)\delta t]\,, \end{aligned} \tag{3.55}$$

where

$$u(t) = -\left(\frac{\hbar}{2\omega\mathcal{E}_0 V}\right)^{\frac{1}{2}} \mathbf{e}\cdot\mathbf{J}(\mathbf{k}, t)\, e^{-i\omega t}. \tag{3.56}$$

For a finite time interval, say from 0 to T, the evolution operator may be written as

$$\hat{U}(T, 0) = \lim_{\delta t\to 0} \hat{\mathcal{T}} \prod_{l=0}^{T/\delta t} \hat{D}\,[u(t_l)\delta t] \tag{3.57}$$

where $\hat{\mathcal{T}}$ is the time-ordering operator and where $t_l = l\,\delta t$. Equation (3.57) becomes, by using Eq. (3.42),

$$\begin{aligned} \hat{U}(T, 0) &= \lim_{\delta t\to 0} e^{i\Phi}\hat{D}\left[\sum_{l=0}^{T/\delta t} u\,(t_l)\,\delta t\right] \\ &= e^{i\Phi}\hat{D}\,[\alpha\,(T)]\,, \end{aligned} \tag{3.58}$$

where

$$\alpha(T) = \lim_{\delta t \to 0} \sum_{l=0}^{T/\delta t} u(t_l)\,\delta t = \int_0^T u(t')dt' \tag{3.59}$$

and where Φ is the accumulated overall phase. With the initial state the vacuum, the state at time T is just the coherent state $|\alpha(T)\rangle$ with $\alpha(T)$ given by Eq. (3.59), apart from the irrelevant overall phase.

3.5 More on the properties of coherent states

The number states are orthonormal, $\langle n|n'\rangle = \delta_{nn'}$, and complete, $\sum_{n=0}^{\infty} |n\rangle\langle n| = \hat{I}$, and as such can be used as a basis for the expansion of an arbitrary state vector of the field, i.e. for a given $|\psi\rangle$

$$|\psi\rangle = \sum_n C_n |n\rangle \tag{3.60}$$

where $C_n = \langle n \mid \psi\rangle$, the coherent states being a particular example. But the coherent states themselves are not orthogonal: for $|\alpha\rangle$ and $|\beta\rangle$

$$\begin{aligned}
\langle \beta \mid \alpha\rangle &= e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2} \\
&\quad \times \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{\beta^{*n}\alpha^m}{\sqrt{n!m!}} \langle n \mid m\rangle \\
&= e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2} \sum_{n=0}^{\infty} \frac{(\beta^*\alpha)^n}{n!} \\
&= e^{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \beta^*\alpha} \\
&= \exp\left[\frac{1}{2}(\beta^*\alpha - \beta\alpha^*)\right] \exp\left[-\frac{1}{2}|\beta - \alpha|^2\right].
\end{aligned} \tag{3.61}$$

The first term is just a complex phase so that

$$|\langle \beta \mid \alpha\rangle|^2 = e^{-|\beta-\alpha|^2} \neq 0. \tag{3.62}$$

Thus the coherent states are not orthogonal. Of course, if $|\beta - \alpha|^2$ is large, they are nearly orthogonal.

The completeness relation for the coherent states is given as an integral over the complex α-plane according to

$$\int |\alpha\rangle\langle\alpha| \frac{d^2\alpha}{\pi} = 1 \tag{3.63}$$

where $d^2\alpha = d\mathrm{Re}(\alpha)d\mathrm{Im}(\alpha)$. The proof of this goes as follows: writing

$$\int |\alpha\rangle\langle\alpha|\, d^2\alpha = \int e^{-|\alpha|^2} \sum_n \sum_m \frac{\alpha^n \alpha^{*m}}{\sqrt{n!m!}} |n\rangle\langle m|\, d^2\alpha, \tag{3.64}$$

we transform to polar coordinates setting $\alpha = re^{i\theta}$, $d^2\alpha = r\,dr\,d\theta$ so that

$$\int |\alpha\rangle\langle\alpha|\,d^2\alpha = \sum_n \sum_m \frac{|n\rangle\langle m|}{\sqrt{n!m!}} \int_0^\infty dr e^{-r^2} r^{n+m+1} \int_0^{2\pi} d\theta e^{i(n-m)\theta}. \tag{3.65}$$

But

$$\int_0^{2\pi} d\theta e^{i(n-m)\theta} = 2\pi\delta_{nm}, \tag{3.66}$$

and, with a further change of variables, $r^2 = y$, $2r\,dr = dy$, we have

$$\int |\alpha\rangle\langle\alpha|\,d^2\alpha = \pi \sum_{n=0}^\infty \frac{|n\rangle\langle n|}{n!} \int_0^\infty dy e^{-y} y^n. \tag{3.67}$$

Since

$$\int_0^\infty dy e^{-y} y^n = n! \tag{3.68}$$

we have

$$\int |\alpha\rangle\langle\alpha|\,d^2\alpha = \pi \sum_{n=0}^\infty |n\rangle\langle n| = \pi, \tag{3.69}$$

which completes the proof.

Any state vector $|\psi\rangle$ in the Hilbert space of the quantized single-mode field can be expressed in terms of the coherent states as

$$|\psi\rangle = \int \frac{d^2\alpha}{\pi} |\alpha\rangle\langle\alpha \mid \psi\rangle. \tag{3.70}$$

But suppose the state $|\psi\rangle$ itself is the coherent state $|\beta\rangle$. Then

$$\begin{aligned} |\beta\rangle &= \int \frac{d^2\alpha}{\pi} |\alpha\rangle\langle\alpha \mid \beta\rangle \\ &= \int \frac{d^2\alpha}{\pi} |\alpha\rangle \exp\left[-\frac{1}{2}|\alpha| - \frac{1}{2}|\beta|^2 + \alpha^*\beta\right]. \end{aligned} \tag{3.71}$$

This last equation shows that the coherent states are not linearly independent. The coherent states are said to be "overcomplete", there being more than enough states available to express any state in terms of the coherent states. Note that in Eq. (3.71) the quantity $\langle\alpha \mid \beta\rangle = \exp(-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 + \alpha^*\beta)$ plays the role of a Dirac delta function. It is often referred to as a reproducing kernel.

For arbitrary state $|\psi\rangle$ we may write

$$\begin{aligned} \langle\alpha \mid \psi\rangle &= \exp\left(-\frac{1}{2}|\alpha|^2\right) \sum_{n=0}^\infty \psi_n \frac{(\alpha^*)^n}{\sqrt{n!}} \\ &= \exp\left(-\frac{1}{2}|\alpha|^2\right) \psi(\alpha^*) \end{aligned} \tag{3.72}$$

where $\psi_n = \langle n|\psi\rangle$ and

$$\psi(z) = \sum_{n=0}^{\infty} \psi_n \frac{z^n}{\sqrt{n!}} \tag{3.73}$$

is absolutely convergent anywhere on the complex z-plane, i.e. $\psi(z)$ is an entire function, since $\langle\psi \mid \psi\rangle = \sum_n |\langle n \mid \psi\rangle|^2 = 1$. The functions $\psi(z)$ constitute the Segal–Bargmann [9] space of entire functions. If $|\psi\rangle$ is a number state, $|\psi\rangle = |n\rangle$, then $\psi_n(z) = z^n/\sqrt{n!}$. These functions form an orthonormal basis on the Segal–Bargmann space.

Let $\hat{F}$ be an operator given as a function of $\hat{a}$, and $\hat{a}^\dagger$, $\hat{F} = F(\hat{a}, \hat{a}^\dagger)$. In terms of the number states, $\hat{F}$ may be decomposed as

$$\begin{aligned} \hat{F} &= \sum_n \sum_m |m\rangle \langle m| \hat{F} |n\rangle \langle n| \\ &= \sum_n \sum_m |m\rangle \hat{F}_{mn} \langle n| \end{aligned} \tag{3.74}$$

where the F_{mn} are the matrix elements $\langle m|\hat{F}|n\rangle$.

With coherent states

$$\hat{F} = \frac{1}{\pi^2} \int d^2\beta \int d^2\alpha \, |\beta\rangle \langle\beta| \hat{F} |\alpha\rangle \langle\alpha| \,. \tag{3.75}$$

But

$$\begin{aligned} \langle\beta| \hat{F} |\alpha\rangle &= \sum_n \sum_m F_{mn} \langle\beta \mid m\rangle \langle n \mid \alpha\rangle \\ &= \exp\left[-\frac{1}{2}(|\beta|^2 + |\alpha|^2)\right] F(\beta^*, \alpha) \,, \end{aligned} \tag{3.76}$$

where

$$F(\beta^*, \alpha) = \sum_m \sum_n F_{mn} \frac{(\beta^*)^m (\alpha)^n}{\sqrt{m!n!}} \,. \tag{3.77}$$

Thus

$$\begin{aligned} \hat{F} = \frac{1}{\pi^2} \int d^2\beta \int d^2\alpha \exp\left[-\frac{1}{2}(|\beta|^2 + |\alpha|^2)\right] \\ \times F(\beta^*, \alpha) |\beta\rangle \langle\alpha| \,. \end{aligned} \tag{3.78}$$

Suppose now that $\hat{F}$ is a Hermitian operator with eigenstates $|\lambda\rangle$ such that

$$\hat{F} = \sum_\lambda \lambda \, |\lambda\rangle \langle\lambda| \,. \tag{3.79}$$

Then

$$\langle m| \hat{F} |n\rangle = \sum_\lambda \lambda \langle m \mid \lambda\rangle \langle\lambda \mid n\rangle \,. \tag{3.80}$$

But

$$|\langle m|\hat{F}|n\rangle| \leq \sum_{\lambda} \lambda\, |\langle m \mid \lambda\rangle\, \langle \lambda \mid n\rangle| \leq \sum_{\lambda} \lambda = \mathrm{Tr}\,\hat{F} \tag{3.81}$$

which implies that $|\langle m|\hat{F}|n\rangle|$ has an upper bound. This being the case, it follows that the function $F(\beta^*, \alpha)$ is an entire function in both β^* and α.

The diagonal elements of an operator $\hat{F}$ in a coherent state basis completely determine the operator. From Eqs. (3.76) and (3.77) we have

$$\langle \alpha|\, \hat{F}\, |\alpha\rangle\, e^{\alpha^*\alpha} = \sum_n \sum_m \frac{\alpha^{*m}\alpha^n}{\sqrt{m!n!}}\, \langle m|\, \hat{F}\, |n\rangle\,. \tag{3.82}$$

Treating α and α^* as independent variables it is apparent that

$$\frac{1}{\sqrt{m!n!}} \left[\frac{\partial^{n+m}(\langle\alpha|\hat{F}|\alpha\rangle e^{\alpha^*\alpha})}{\partial\alpha^{*m}\,\partial\alpha^n}\right]\Bigg|_{\substack{\alpha^*=0\\ \alpha=0}} = \langle m|\, \hat{F}\, |n\rangle\ . \tag{3.83}$$

Thus from "diagonal" coherent-state matrix elements of $\hat{F}$ we can obtain all the matrix elements of the operator in the number basis.

3.6 Phase-space pictures of coherent states

It is well known that the concept of phase space in quantum mechanics is problematic owing to the fact that the canonical variables $\hat{x}$ and $\hat{p}$ are incompatible, i.e. they do not commute. Thus the state of a system is not well localized as a point in phase space as it is in classical mechanics. Nevertheless, we have shown that the coherent states minimize the uncertainty relation for the two orthogonal quadrature operators and that the uncertainties of the two quadratures are equal. Recall that $\hat{X}_1 = (\hat{a} + \hat{a}^\dagger)/2$ and $\hat{X}_2 = (\hat{a} - \hat{a}^\dagger)/2i$. These operators are dimensionless scaled position and momentum operators, respectively. Their coherent-state expectation values are $\langle \hat{X}_1\rangle_\alpha = \frac{1}{2}(\alpha + \alpha^*) = \mathrm{Re}\alpha$, $\langle \hat{X}_2\rangle_\alpha = \frac{1}{2i}(\alpha - \alpha^*) = \mathrm{Im}\alpha$. Thus the complex α-plane plays the role of phase space where, up to scale factors, the real and imaginary parts of α are position and momentum variables respectively. A coherent state $|\alpha\rangle$with $\alpha = |\alpha|e^{i\theta}$ may be represented pictorially then as in Fig. 3.5, where the shaded circle represents "area of uncertainty" of the coherent state, the fluctuations being equal in all directions of phase space, the center of the circle located at distance $|\alpha| = \langle\hat{n}\rangle^{1/2}$ from the origin and at angle θ above the position axis. Further, $\Delta\theta$, in a qualitative sense, represents the phase uncertainty of the coherent state and it should be clear that $\Delta\theta$ diminishes for increasing $|\alpha|$, the fluctuations in X_1 and X_2 being independent of α and identical to those of the vacuum. For the vacuum, $|\alpha| = 0$, the phase-space representation is given in Fig. 3.6, where it is evident that uncertainty in the phase is as large as possible, i.e. $\Delta\theta = 2\pi$.

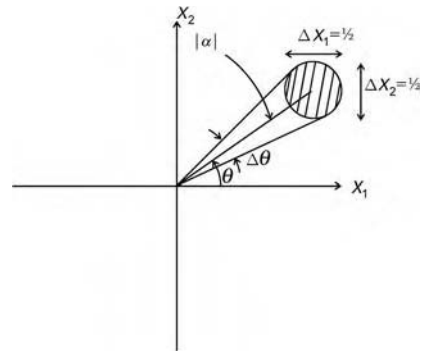


Fig. 3.5. Phase-space portrait of a coherent state of amplitude $|\alpha|$ and phase angle θ. Note the error circle is the same for all coherent states. Note that as $|\alpha|$ increases, the phase uncertainty $\Delta\theta$ decreases, as would be expected in the "classical limit".

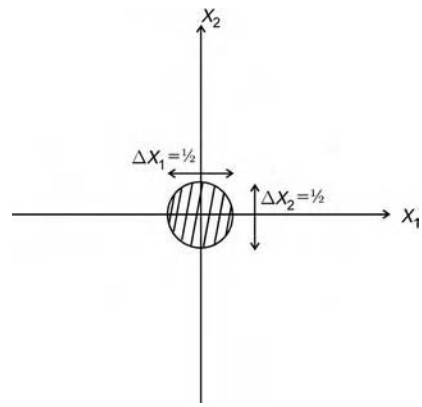


Fig. 3.6. Phase-space portrait of the quantum vacuum state.

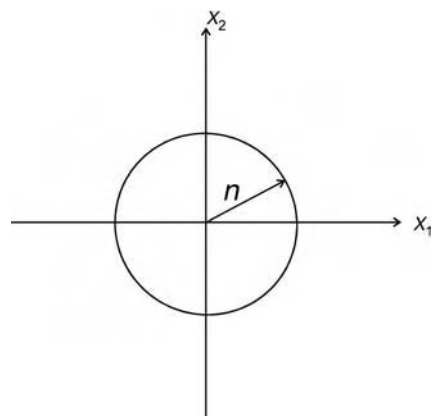


Fig. 3.7. Phase-space portrait of the number state $|n\rangle$. The uncertainty in the photon number is $\Delta n = 0$ while the phase is entirely random.

A number state $|n\rangle$ can be represented in phase space as a circle of radius n, the uncertainty in n being zero and the uncertainty in phase again being 2π, as in Fig. 3.7. It must be understood that these pictures are qualitative in nature but are useful as a graphical way of visualizing the distribution of noise in various quantum states of the field. As most quantum states of the field have no classical

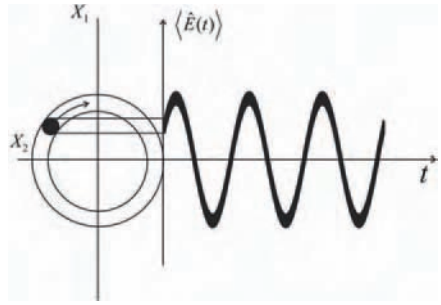


Fig. 3.8. The error circle of a coherent state (the black dot) revolves about the origin of phase space at the oscillator angular frequency ω and the expectation value of the electric field is the projection onto an axis parallel with X_1.

analog, the corresponding phase-space portraits should not be taken too literally. Yet these representations will be quite useful when we discuss the nature of the squeezed states of light in Chapter 7.

Finally in this section, we make one more use of the phase-space diagrams, namely to illustrate the time evolution of quantum states for a non-interacting field. We have seen that, for a noninteracting field, a coherent state $|\alpha\rangle$ evolves to the coherent state $|\alpha e^{-i\omega t}\rangle$. This can be pictured as a clockwise rotation of the error circle in phase since $\langle \alpha e^{-i\omega t}|\hat{X}_1|\alpha e^{-i\omega t}\rangle = \alpha \cos\omega t$, $\langle \alpha e^{-i\omega t}|\hat{X}_2|\alpha e^{-i\omega t}\rangle = -\alpha \sin\omega t$, assuming α real. Because in the Schrödinger picture the electric field operator is given, from Eq. (2.15) and (2.52), as

$$\hat{E}_x = 2\mathcal{E}_0 \sin(kz)\hat{X}_1, \tag{3.84}$$

the expectation value for the coherent state $|\alpha e^{-i\omega t}\rangle$ is

$$\langle \alpha e^{-i\omega t}|\hat{E}_x|\alpha e^{-i\omega t}\rangle = 2\mathcal{E}_0 \sin(kz)\alpha \cos\omega t. \tag{3.85}$$

Thus apart from the scale factor $2\mathcal{E}_0 \sin(kz)$, the time evolution of the electric field and its fluctuations, is given by the projection of the $\langle \hat{X}_1 \rangle$ axis as a function of time as indicated in Fig. 3.8.

The evolution of points within the error circle are shown indicating the uncertainty of the electric field – the "quantum flesh" on the "classical bones", so to speak. Note that the greater the excitation of the field, i.e. α, the more classical the field appears since the fluctuations are independent of α. But the coherent state is the most classical of all the quantum states so it is apparent that for the field in something other than a coherent state, the expectation value of the field may in no way appear classical-like. A number state is a very nonclassical state and, by using representative points from its phase-space portrait, it is easy to see, from Fig. 3.7, that the expectation value of the field is zero. But it is possible to imagine other kinds of states having no vanishing expectation values of the field but where fluctuations may be less than those of a coherent state in one part of the field. These are the squeezed states to be taken up in Chapter 7.

3.7 Density operators and phase-space probability distributions

Recall (see also Appendix A) that for a mixture (or ensemble) of quantum states $|\psi_1\rangle, |\psi_2\rangle, \ldots$, the density operator is given as

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle \, \langle\psi_i| \tag{3.86}$$

where the p_i are the probabilities of finding the system in the ith member of the ensemble and

$$\mathrm{Tr}\,(\hat{\rho}) = \sum_i p_i = 1. \tag{3.87}$$

(For a pure state, $\hat{\rho} = |\psi\rangle\langle\psi|$.) The expectation value of an operator $\hat{O}$ is given by

$$\langle\hat{O}\rangle = \mathrm{Tr}(\hat{O}\hat{\rho}) = \sum_i p_i \, \langle\psi_i| \, \hat{O} \, |\psi_i\rangle \, . \tag{3.88}$$

As in Eq. (3.74), we can resolve unity in terms of the number states on both sides of the density operator to obtain

$$\hat{\rho} = \sum_n \sum_m |m\rangle \, \rho_{mn} \, \langle n| \, . \tag{3.89}$$

All the matrix elements $\rho_{mn} = \langle m|\hat{\rho}|n\rangle$ are required completely to determine the operator $\hat{\rho}$. The diagonal elements, $P_n = \rho_{nn}$, are just the probabilities of finding n photons in the field.

On the other hand, resolving unity with coherent states on both sides of $\hat{\rho}$, as in Eq. (3.75), results in

$$\hat{\rho} = \iint \langle\alpha'|\hat{\rho}|\alpha''\rangle|\alpha'\rangle\langle\alpha''| \frac{d^2\alpha' d^2\alpha''}{\pi^2}. \tag{3.90}$$

But there is yet another way to represent $\hat{\rho}$ in terms of coherent states, namely

$$\hat{\rho} = \int P\,(\alpha) \, |\alpha\rangle \, \langle\alpha| \, d^2\alpha, \tag{3.91}$$

where $P(\alpha)$ is a weight function sometimes known as the Glauber–Sudarshan P function [2,10]. The right-hand side of Eq. (3.91) is the "diagonal" form of the density operator and the P function is analogous to the phase-space distributions of statistical mechanics. Here, the real and imaginary parts of α are the variables of the phase space. $P(\alpha)$ must be real as $\hat{\rho}$ is a Hermitian operator. Also

$$\begin{aligned}
\mathrm{Tr}\hat{\rho} &= \mathrm{Tr} \int P\,(\alpha) \, |\alpha\rangle \, \langle\alpha| \, d^2\alpha \\
&= \int \sum_n P\,(\alpha) \, \langle n|\alpha\rangle \, \langle\alpha|n\rangle \, d^2\alpha \\
&= \int P\,(\alpha) \sum_n \langle\alpha \, |n\rangle \, \langle n| \, \alpha\rangle \, d^2\alpha \\
&= \int P\,(\alpha) \, \langle\alpha \mid \alpha\rangle \, d^2\alpha = \int P\,(\alpha) \, d^2\alpha = 1,
\end{aligned} \tag{3.92}$$

just as we would expect of a phase-space probability distribution.

But for some quantum states of the field, $P(\alpha)$ can have properties quite unlike those of any *true* probability distribution where one would expect to have $P(\alpha) \geq 0$. There are quantum states for which $P(\alpha)$ is negative or highly singular. In these cases, the corresponding quantum states are called "nonclassical". In fact, we may define a nonclassical state as a state for which the corresponding $P(\alpha)$ is negative in some regions of phase space (the α-plane) or is more singular than a delta function (the reason for the second criterion will become clear shortly). To speak of nonclassical states of light might seem an oxymoron. After all, aren't *all* states of light quantum mechanical? Well, yes, *all* states of light *are* quantum mechanical, but, waxing Orwellian, it turns out that some states are more quantum mechanical than others. States for which $P(\alpha)$ is positive or no more singular than a delta function are, in this sense, classical. Coherent states, which we have already shown to be quasi-classical in that they describe states of the field having properties close to what we would expect for classical oscillating coherent fields, have P functions given as delta functions, as we shall show, and are therefore classical in the sense described here. Certain effects, among them being quadrature and amplitude (or number) squeezing, can occur only for states for which the P functions are negative or highly singular. For that reason, the various forms of squeezing are known as distinctly non-classical effects. These effects will be discussed in Chapter 7.

But how do we calculate $P(\alpha)$? A rather general procedure, which we follow here, has been given by Mehta [11]. Starting with Eq. (3.91) and using coherent state $|u\rangle$ and $|-u\rangle$, we have

$$\begin{aligned}\langle -u\,|\hat{\rho}|\,u\rangle &= \int P(\alpha)\,\langle -u\,|\alpha\rangle\,\langle\alpha|\,u\rangle\,d^2\alpha \\ &= \int P(\alpha)\exp\left[-\frac{1}{2}\,|u|^2-\frac{1}{2}\,|\alpha|^2-u^*\alpha\right] \\ &\quad\times\exp\left[-\frac{1}{2}\,|\alpha|^2-\frac{1}{2}\,|u|^2+a^*u\right]d^2\alpha \\ &= e^{-|u|^2}\int P(\alpha)\,e^{-|\alpha|^2}e^{\alpha^*u-\alpha u^*}d^2\alpha. \end{aligned} \tag{3.93}$$

Now let $\alpha = x + iy$ and $u = x' + iy'$ so that $\alpha^*u - \alpha u^* = 2\,i(x'y - xy')$. We then define the Fourier transforms in the complex plane:

$$g(u) = \int f\,(\alpha)\,e^{\alpha^*u-\alpha u^*}d^2\alpha \tag{3.94a}$$

$$f\,(\alpha) = \frac{1}{\pi^2}\int g\,(u)\,e^{u^*\alpha-u\alpha^*}d^2\alpha. \tag{3.94b}$$

With the identifications

$$\begin{aligned} g\,(u) &= e^{|u|^2}\,\langle -u|\,\hat{\rho}\,|u\rangle \\ f\,(\alpha) &= P(\alpha)\,e^{-|u|^2} \end{aligned} \tag{3.95}$$

we then obtain from Eq. (3.94b)

$$P(\alpha) = \frac{e^{|\alpha|^2}}{\pi^2} \int e^{|u|^2} \langle -u| \rho |u\rangle e^{u^*\alpha - u\alpha^*} d^2u. \tag{3.96}$$

Care must be taken in regard to the convergence of the integral since $e^{|u|^2} \to \infty$ as $|u| \to \infty$.

Now let us consider some examples. First we consider the pure coherent state $|\beta\rangle$, where $\hat{\rho} = |\beta\rangle\langle\beta|$. With

$$\begin{aligned} \langle -u| \hat{\rho} |u\rangle &= \langle -u \mid \beta\rangle \langle \beta \mid u\rangle \\ &= e^{-|\beta|^2} e^{-|u|^2} e^{-u^*\beta + \beta^* u}, \end{aligned} \tag{3.97}$$

then

$$P(\alpha) = e^{|\alpha|^2} e^{-|\beta|^2} \frac{1}{\pi^2} \int e^{u^*(\alpha-\beta) - u(\alpha^* - \beta^*)} d^2u. \tag{3.98}$$

But the Fourier integral is just the two-dimensional form of the Dirac delta function:

$$\begin{aligned} \delta^2(\alpha - \beta) &= \delta[\mathrm{Re}(\alpha) - \mathrm{Re}(\beta)]\, \delta[\mathrm{Im}(\alpha) - \mathrm{Im}(\beta)] \\ &= \frac{1}{\pi^2} \int e^{u^*(\alpha-\beta) - u(\alpha^* - \beta^*)} d^2u \end{aligned} \tag{3.99}$$

so that

$$P(\alpha) = \delta^2(\alpha - \beta). \tag{3.100}$$

This is the same as the distribution for a classical harmonic oscillator.

But if a coherent state is a classical-like state, the number state $|n\rangle$ is at the other extreme, a state that cannot at all be described classically. For a pure number state $|n\rangle$, $\hat{\rho} = |n\rangle\langle n|$ and

$$\langle -u| \hat{\rho} |u\rangle = \langle -u \mid n\rangle \langle n \mid u\rangle = e^{-|u|^2} \frac{(-u^*u)^n}{n!}. \tag{3.101}$$

Thus

$$P(\alpha) = \frac{e^{|\alpha|^2}}{n!} \frac{1}{\pi^2} \int (-u^*u)^n e^{u^*\alpha - u\alpha^*} d^2u. \tag{3.102}$$

The integral does not exist in terms of ordinary functions. Formally, we can write

$$\begin{aligned} P(\alpha) &= \frac{e^{|\alpha|^2}}{n!} \frac{\partial^{2n}}{\partial\alpha^n \partial\alpha^{*n}} \frac{1}{\pi^2} \int e^{u^*\alpha - u\alpha^*} d^2u \\ &= \frac{e^{|\alpha|^2}}{n!} \frac{\partial^{2n}}{\partial\alpha^n \partial\alpha^{*n}} \delta^{(2)}(\alpha). \end{aligned} \tag{3.103}$$

The derivative of the delta function, called a tempered distribution, is more singular than a delta function, and has meaning only under the integral sign,

e.g. for some function $F(\alpha, \alpha^*)$

$$\int F(\alpha, \alpha^*) \frac{\partial^{2n}}{\partial\alpha^n \partial\alpha^{*n}} \delta^{(2)}(\alpha) d^2\alpha = \left[\frac{\partial^{2n} F(\alpha, \alpha^*)}{\partial\alpha^n \partial\alpha^{*n}} \right]_{\substack{\alpha^*=0 \\ \alpha=0}} . \tag{3.104}$$

At this point we introduce the optical equivalence theorem of Sudarshan [10]. Suppose we have a "normally ordered" function of the operators $\hat{a}$ and $\hat{a}^\dagger$, $G^{(\mathrm{N})}(\hat{a}, \hat{a}^\dagger)$, where the annihilation operators stand to the right of the creation operators:

$$\hat{G}^{(\mathrm{N})}(\hat{a}, \hat{a}^\dagger) = \sum_n \sum_m C_{nm} (\hat{a}^\dagger)^n \hat{a}^m. \tag{3.105}$$

(It should not be hard to guess what an *antinormally* ordered operator looks like!) The average of this function is

$$\begin{aligned} \langle G^{(\mathrm{N})}(\hat{a}, \hat{a}^\dagger) \rangle &= \mathrm{Tr}\left[\hat{G}^{(\mathrm{N})}(\hat{a}, \hat{a}^\dagger) \hat{\rho} \right] \\ &= \mathrm{Tr} \int P(\alpha) \sum_n \sum_m C_{nm} (\hat{a}^\dagger)^n a^m |\alpha\rangle \langle\alpha| d^2\alpha \\ &= \int P(\alpha) \sum_n \sum_m C_{nm} \langle\alpha| (\hat{a}^\dagger)^n a^m |\alpha\rangle d^2\alpha \\ &= \int P(\alpha) \sum_n \sum_m C_{nm} \alpha^{*n} \alpha^m d^2\alpha \\ &= \int P(\alpha) G^{(\mathrm{N})}(\alpha, \alpha^*) \, d^2\alpha. \end{aligned} \tag{3.106}$$

The last line is the **optical equivalence theorem**: the expectation value of a normally ordered operator is just the P function weighted average of the function obtained from the operator by the replacements $a \to \alpha$ and $a^\dagger \to \alpha^*$.

Normally ordered operators will be important later, particularly in the discussion of the photoelectric detection of light resulting from the absorption of photons. It is useful to introduce the normal ordering operator denoted : :. For an arbitrary function of $\hat{a}$ and $\hat{a}^\dagger$, $O(\hat{a}, \hat{a}^\dagger)$, we have

$$: O(\hat{a}, \hat{a}^\dagger) : \equiv O^{(\mathrm{N})}(\hat{a}, \hat{a}^\dagger) \tag{3.107}$$

where the commutation relations are to be *disregarded*. The number operator $\hat{n} = \hat{a}^\dagger \hat{a}$ is already normally ordered so

$$\langle \hat{n} \rangle = \langle \hat{a}^\dagger \hat{a} \rangle = \int P(\alpha) |\alpha|^2 d^2\alpha. \tag{3.108}$$

But $\hat{n}^2 = \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a}$ is not normally ordered. Thus

$$: \hat{n}^2 : = (\hat{a}^\dagger)^2 \hat{a}^2$$

and

$$\langle : \hat{n}^2 : \rangle = \langle (\hat{a}^\dagger)^2 \hat{a}^2 \rangle = \int P(\alpha) |\alpha|^4 d^2\alpha. \tag{3.109}$$

The utility of the normal ordering operator and the optical equivalence theorem will be evident in Chapter 7.

We can, under suitable conditions, represent other operators besides $\hat{\rho}$ in the "diagonal" coherent state form – sometimes called the P-representation. For operator $\hat{B}$, the P-representation is

$$\hat{B} = \int B_p(\alpha, \alpha^*) \, |\alpha\rangle \langle\alpha| \, d^2\alpha. \tag{3.110}$$

The average of $\hat{B}$ is given by

$$\begin{aligned} \langle \hat{B} \rangle &= \mathrm{Tr}(\hat{B}\hat{\rho}) \\ &= \sum_n \langle n| \int B_p(\alpha, \alpha^*) \, |\alpha\rangle\langle\alpha|\hat{\rho}|n\rangle d^2\alpha \\ &= \int B_p(\alpha, \alpha^*) \, \langle\alpha|\hat{\rho}|\alpha\rangle d^2\alpha. \end{aligned} \tag{3.111}$$

Evidently, the expectation value of the density operator with respect to the coherent state also plays the role of a phase-space probability distribution. This is usually called the Q, or Husimi, function [12]:

$$Q(\alpha) = \langle\alpha| \, \hat{\rho} \, |\alpha\rangle/\pi. \tag{3.112}$$

For $\hat{B} = \hat{I}$ we obtain the normalization condition

$$\int Q(\alpha) d^2\alpha = 1. \tag{3.113}$$

Unlike the P function, the Q function is positive for all quantum states. Of course, we can always define the corresponding Q-representation of the operator $\hat{B}$ as just the expectation value with respect to the coherent state:

$$\begin{aligned} B_Q(\alpha, \alpha^*) &\equiv \langle\alpha| \, \hat{B} \, |\alpha\rangle \\ &= e^{-|\alpha|^2} \sum_n \sum_m \frac{B_{nm}}{(n!m!)^{1/2}} \, (\alpha^*)^n \, (\alpha)^m \end{aligned} \tag{3.114}$$

where $B_{nm} = \langle n|\hat{B}|m\rangle$. Once again, let us calculate $\langle \hat{B} \rangle$, but now we write $\hat{\rho}$ in the P-representation:

$$\begin{aligned} \langle \hat{B} \rangle = \mathrm{Tr}(\hat{B}\hat{\rho}) &= \mathrm{Tr} \int \hat{B} P(\alpha) |\alpha\rangle\langle\alpha| d^2\alpha \\ &= \int P(\alpha) \langle\alpha|\hat{B}|\alpha\rangle d^2\alpha \\ &= \int P(\alpha) B_Q(\alpha, \alpha^*) d^2\alpha. \end{aligned} \tag{3.115}$$

Thus if we use the P-representation of $\hat{\rho}$ we need the Q-representation of $\hat{B}$, or, from Eq. (3.111), if we use the P-representation of $\hat{B}$ we need the Q-representation of $\hat{\rho}$.

The Q function has the character of a probability distribution in the sense of positivity whereas the P function is really a quasi-probability distribution.

There is, in fact, one other important quasi-probability distribution over phase space, namely, the Wigner function. The Wigner function seems to be the earliest introduced of the phase-space quasi-probability distributions, making its debut in 1932 [13]. It is defined, for an arbitrary density operator $\hat{\rho}$, as

$$W(q,p) \equiv \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\left\langle q+\frac{1}{2}x\left|\hat{\rho}\right|q-\frac{1}{2}x\right\rangle e^{ipx/\hbar}dx \tag{3.116}$$

where $|q \pm \frac{1}{2}x\rangle$ are the eigenkets of the position operator. If the state in question is a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$ then

$$W(q,p) \equiv \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\psi^*\left(q-\frac{1}{2}x\right)\psi\left(q+\frac{1}{2}x\right)e^{ipx/\hbar}dx \tag{3.117}$$

where $\langle q+\frac{1}{2}x \mid \psi\rangle = \psi(q+\frac{1}{2}x)$, etc. Integrating over the momentum we obtain

$$\begin{aligned}\int_{-\infty}^{\infty}W(q,p)\,dp &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\psi^*\left(q-\frac{1}{2}x\right)\psi\left(q+\frac{1}{2}x\right)\int_{-\infty}^{\infty}e^{ipx/\hbar}dpdx\\ &= \int_{-\infty}^{\infty}\psi^*\left(q-\frac{1}{2}x\right)\psi\left(q+\frac{1}{2}x\right)\delta(x)dx\\ &= |\psi(q)|^2\end{aligned} \tag{3.118}$$

which is the probability density for position variable q. Likewise, integrating over q we obtain

$$\int_{-\infty}^{\infty}W(q,p)\,dq = |\varphi(p)|^2 \tag{3.119}$$

where $\varphi(p)$ is the momentum space wave related to the coordinate space wave function $\psi(q)$ by a Fourier transform. The right-hand side of Eq. (3.119) is, of course, just the probability density in momentum space. But $W(q,p)$ itself is not a true probability distribution since it can take on negative values for some non-classical states, as we shall show. Like the other distributions, the Wigner function can be used to calculate averages. However, functions of the operators $\hat{q}$ and $\hat{p}$ to be averaged must be Weyl, or symmetrically, ordered in terms of these operators. For example, the classical function qp must be replaced by $(\hat{q}\hat{p}+\hat{p}\hat{q})/2$ and

$$\langle\hat{q}\hat{p}+\hat{p}\hat{q}\rangle = \int(qp+pq)\,W(q,p)dqdp\,. \tag{3.120}$$

In general, if $\{G(\hat{q},\hat{p})\}_{\mathrm{W}}$ is a Weyl-ordered function, where the bracket $\{\ \}_{\mathrm{W}}$ means Weyl, or symmetric, ordering, then

$$\langle\{G(\hat{q},\hat{p})\}_{\mathrm{W}}\rangle = \int\{G(\hat{q},\hat{p})\}_{\mathrm{W}}\,W(q,p)dqdp \tag{3.121}$$

is the corresponding phase-space average.

3.8 Characteristic functions

Consider for a moment a classical random variable that we denote x. Suppose that $\rho(x)$ is a classical probability density associated with the variable X. Then it follows that

$$\rho(x) \geq 0 \tag{3.122}$$

and

$$\int \rho(x)dx = 1. \tag{3.123}$$

The nth moment of x is defined as

$$\langle x^n \rangle = \int dx\, x^n \rho(x). \tag{3.124}$$

If all moments $\langle x^n \rangle$ are known then $\rho(x)$ is completely specified. This can be seen by introducing the characteristic function

$$\begin{aligned} C(k) &= \langle e^{ikx} \rangle = \int dx\, e^{ikx} \rho(x) \\ &= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle . \end{aligned} \tag{3.125}$$

Evidently, the probability density is just the Fourier transform of the characteristic function:

$$\rho(x) = \frac{1}{2\pi} \int dk e^{-ikx} C(k). \tag{3.126}$$

Thus, if all the moments $\langle x^n \rangle$ are known, $C(k)$ is known and thus $\rho(x)$ is known. On the other hand if we are given the characteristic function, we can calculate the moments according to

$$\langle x^n \rangle = \frac{1}{i^n} \left. \frac{d^n C(k)}{dk^n} \right|_{k=0}. \tag{3.127}$$

We now proceed to introduce quantum mechanical characteristic functions. There are, in fact, three such functions, namely

$$C_{\mathrm{W}}(\lambda) = \mathrm{Tr}[\hat{\rho}\, e^{\lambda \hat{a}^\dagger - \lambda^* \hat{a}}] = \mathrm{Tr}[\hat{\rho}\, \hat{D}(\lambda)] \quad \text{(Wigner)} \tag{3.128a}$$

$$C_{\mathrm{N}}(\lambda) = \mathrm{Tr}[\hat{\rho} e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}}] \qquad \text{(normally ordered)} \tag{3.128b}$$

$$C_{\mathrm{A}}(\lambda) = \mathrm{Tr}[\hat{\rho} e^{-\lambda \hat{a}^\dagger} e^{\lambda^* \hat{a}}] \qquad \text{(antinormally ordered).} \tag{3.128c}$$

These functions are related through the disentangling theorem of Eq. (3.32) according to

$$C_{\mathrm{W}}(\lambda) = C_{\mathrm{N}}(\lambda)\, e^{-\frac{1}{2}|\lambda|^2} = C_{\mathrm{A}}(\lambda)\, e^{\frac{1}{2}|\lambda|^2}. \tag{3.129}$$

Furthermore, it is easy to show that

$$\langle(\hat{a}^\dagger)^m\hat{a}^n\rangle = \mathrm{Tr}[\hat{\rho}(\hat{a}^\dagger)^m\hat{a}^n] = \frac{\partial^{(m+n)}}{\partial\lambda^m\partial(-\lambda^*)^n}C_{\mathrm{N}}(\lambda)\bigg|_{\lambda=0} \tag{3.130a}$$

$$\langle\hat{a}^m(\hat{a}^\dagger)^n\rangle = \mathrm{Tr}[\hat{\rho}\,\hat{a}^m(\hat{a}^\dagger)^n] = \frac{\partial^{(m+n)}}{\partial\lambda^n\partial(-\lambda^*)^m}C_{\mathrm{A}}(\lambda)\bigg|_{\lambda=0} \tag{3.130b}$$

$$\langle\{(\hat{a}^\dagger)^m\hat{a}^n\}_{\mathrm{W}}\rangle = \mathrm{Tr}[\hat{\rho}\,\{(\hat{a}^\dagger)^m\hat{a}^n\}_{\mathrm{W}}] = \frac{\partial^{(m+n)}}{\partial\lambda^m\partial(-\lambda^*)^n}C_{\mathrm{W}}(\lambda)\bigg|_{\lambda=0}. \tag{3.130c}$$

Instead of the three different characteristic functions, we can introduce the s-parameterized function of Cahill and Glauber [14]:

$$C(\lambda, s) = \mathrm{Tr}[\hat{\rho}\exp(\lambda\hat{a}^\dagger - \lambda^*\hat{a} + s|\lambda|^2/2)] \tag{3.131}$$

such that $C(\lambda, 0) = C_{\mathrm{W}}(\lambda)$, $C(\lambda, 1) = C_{\mathrm{N}}(\lambda)$, and $C(\lambda, -1) = C_{\mathrm{A}}(\lambda)$.

The connections between these characteristic functions and the various quasi-probability distributions will now be made. For example, the antinormally ordered characteristic function may be written as

$$\begin{aligned} C_{\mathrm{A}}(\lambda) &= \mathrm{Tr}[\hat{\rho}e^{-\lambda^*\hat{a}}e^{\lambda\hat{a}^\dagger}] \\ &= \mathrm{Tr}[e^{\lambda\hat{a}^\dagger}\hat{\rho}\,e^{-\lambda^*\hat{a}}] \\ &= \frac{1}{\pi}\int d^2\alpha\,\langle\alpha|e^{\lambda\hat{a}^\dagger}\hat{\rho}\,e^{-\lambda^*\hat{a}}|\alpha\rangle \\ &= \int d^2\alpha\,Q(\alpha)e^{\lambda\alpha^*-\lambda^*\alpha}, \end{aligned} \tag{3.132}$$

which is just a two-dimensional Fourier transform of the Q function. The inverse Fourier transform yields:

$$Q(\alpha) = \frac{1}{\pi^2}\int C_{\mathrm{A}}(\lambda)e^{\lambda^*\alpha-\lambda\alpha^*}d^2\lambda. \tag{3.133}$$

Now consider the normally ordered characteristic function. Writing $\hat{\rho}$ in the P-representation we have

$$\begin{aligned} C_{\mathrm{N}}(\lambda) &= \mathrm{Tr}[\hat{\rho}\,e^{\lambda\hat{a}^\dagger}e^{-\lambda^*\hat{a}}] \\ &= \int P(\alpha)\langle\alpha|\,e^{\lambda\hat{a}^\dagger}e^{-\lambda^*\hat{a}}|\alpha\rangle d^2\alpha \\ &= \int P(\alpha)\,e^{\lambda\alpha^*-\lambda^*\alpha}d^2\alpha \end{aligned} \tag{3.134}$$

which is just the Fourier transform of the P function. The inverse transform yields

$$P(\alpha) = \frac{1}{\pi^2}\int e^{\lambda^*\alpha-\lambda\,\alpha^*}C_{\mathrm{N}}(\lambda)d^2\lambda. \tag{3.135}$$

And finally, and perhaps to no great surprise, it turns out that the Wigner function may be obtained as the Fourier transform of the Weyl-ordered

characteristic function:

$$W(\alpha) \equiv \frac{1}{\pi^2}\int \exp(\lambda^*\alpha - \lambda\alpha^*)\, C_{\rm W}(\lambda) d^2\lambda$$
$$= \frac{1}{\pi^2}\int \exp(\lambda^*\alpha - \lambda\alpha^*)\, C_{\rm N}(\lambda) e^{-|\lambda|^2/2}\, d^2\lambda. \tag{3.136}$$

This definition is equivalent to the previous one with the proper interpretation of variables.

As an application of the characteristic function, we now derive the P function corresponding to the thermal, or chaotic, state of the field. Recall that the field in this case is a mixed state given by density operator $\hat{\rho}_{\rm Th}$ of Eq. (2.144). We first calculate the Q function for this density operator according to

$$Q(\alpha) = \langle\alpha|\hat{\rho}_{\rm Th}|\alpha\rangle/\pi$$
$$= \frac{1}{\pi}\, e^{-|\alpha|^2} \sum_n \sum_m \langle m|\hat{\rho}_{\rm Th}\,|n\rangle\, \frac{(\alpha^*)^m\, \alpha^n}{(m!n!)^{1/2}}$$
$$= \frac{e^{-|\alpha|^2}}{\pi\,(1+\bar{n})}\, \sum_m \left(\frac{\bar{n}}{1+\bar{n}}\right)^n \frac{(\alpha^*\alpha)^n}{n!}$$
$$= \frac{1}{\pi\,(1+\bar{n})} \exp\left(-\frac{|\alpha|^2}{1+\bar{n}}\right), \tag{3.137}$$

where $\bar{n}$ for the thermal state is given by Eq. (2.141). Then from Eq. (3.132) we have

$$C_{\rm A}(\lambda) = \frac{1}{\pi\,(1+\bar{n})}\int d^2\alpha \exp\left(-\frac{|\alpha|^2}{1+\bar{n}}\right) e^{\lambda\alpha^* - \lambda^*\alpha}. \tag{3.138}$$

Now letting $\alpha = (q+ip)/\sqrt{2}$, $\lambda = (x+iy)/\sqrt{2}$ where $d^2\alpha = dqdp/2$ we have

$$C_{\rm A}(x,y) = \frac{1}{2\pi\,(1+\bar{n})}\int \exp\left[-\frac{(q^2+p^2)}{2(1+\bar{n})}\right]$$
$$\times \exp[i(yq - xp)] dqdp. \tag{3.139}$$

Using the standard Gaussian integral

$$\int e^{-as^2} e^{\pm\beta s} ds = \sqrt{\frac{\pi}{a}}\, e^{\beta^2/4a}, \tag{3.140}$$

we straightforwardly obtain

$$C_{\rm A}(\lambda) = \exp[-(1+\bar{n})|\lambda|^2]. \tag{3.141}$$

But from Eq. (3.129) we have $C_{\rm N}(\lambda) = C_{\rm A}(\lambda)\exp(|\lambda|^2)$ and thus from Eq. (3.135) we finally obtain

$$P(\alpha) = \frac{1}{\pi^2}\int \exp(-\bar{n}|\lambda|^2)\, e^{\lambda^*\alpha - \lambda\alpha^*} d^2\lambda$$
$$= \frac{1}{\pi\bar{n}}\, \exp\left(-\frac{|\alpha|^2}{\bar{n}}\right). \tag{3.142}$$

This is Gaussian so it may be interpreted as a true probability distribution.

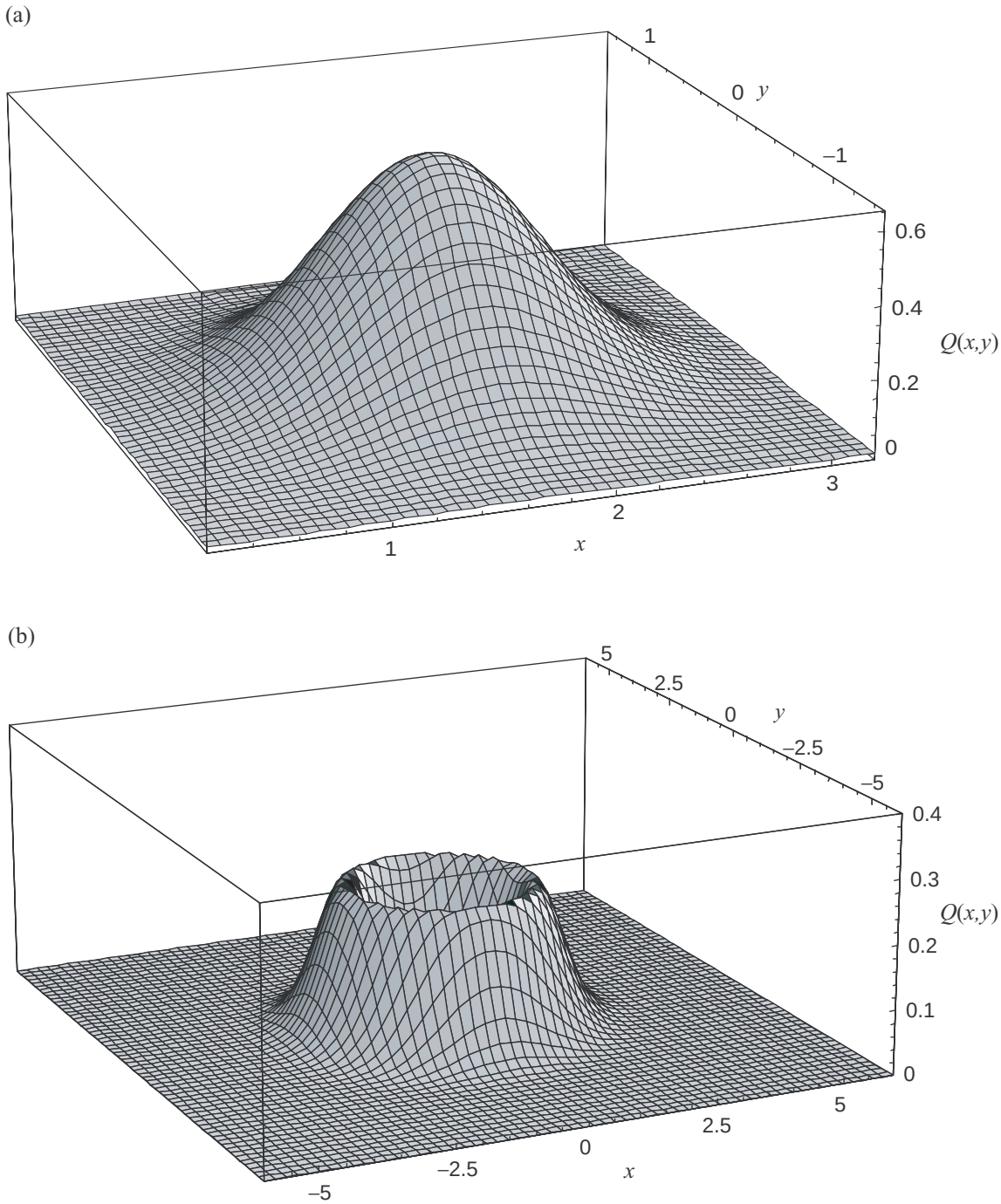


Fig. 3.9. Q function for (a) a coherent state with $\bar{n} = 10$, (b) a number state with $n = 3$.

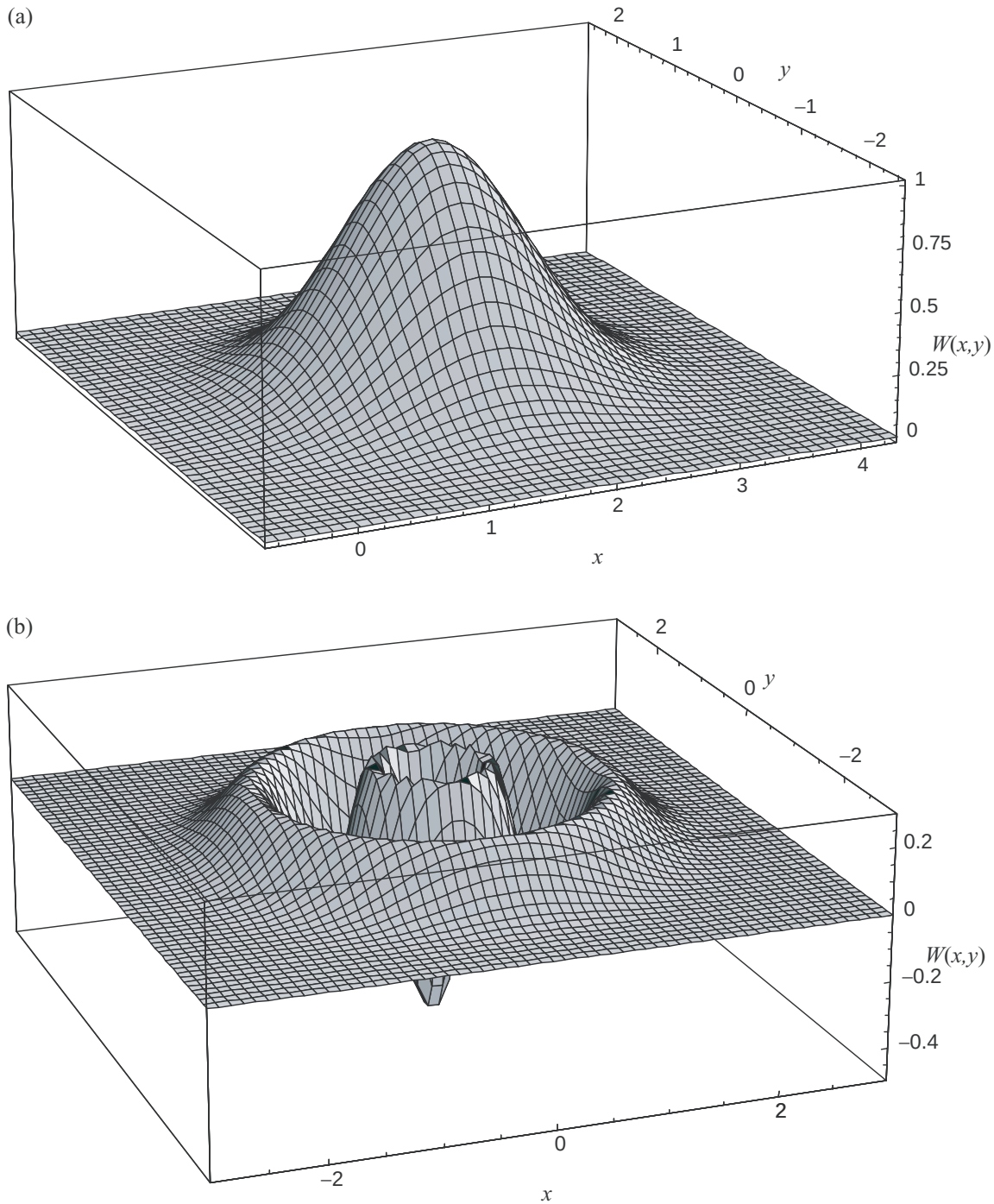


Fig. 3.10. Wigner function for (a) a coherent state with $\bar{n} = 10$, (b) a number state with $n = 3$.

Finally, to conclude this chapter, we examine the Q and Wigner functions for the most classical of the quantum states, the coherent state, and for the most quantum-mechanical state, the number state. For the coherent state $\hat{\rho} = |\beta\rangle\langle\beta|$ one easily has the Q function

$$Q(\alpha) = \frac{1}{\pi}|\langle\alpha \mid \beta\rangle|^2 = \frac{1}{\pi}\exp(-|\alpha - \beta|^2), \tag{3.143}$$

whereas for the number state $\hat{\rho} = |n\rangle\langle n|$,

$$Q(\alpha) = \frac{1}{\pi}|\langle\alpha \mid n\rangle|^2 = \frac{1}{\pi}\exp(-|\alpha|^2)\frac{|\alpha|^{2n}}{n!}. \tag{3.144}$$

Setting $\alpha = x + iy$, we plot these as functions of x and y in Fig. 3.9.

The Q function of the coherent state is just a Gaussian centered at β while the state for a number state is rather annular of center radius $r \sim n$. Note how these functions correlate with the phase-space figures introduced in Section 3.6. The corresponding Wigner functions, obtained from Eq. (3.136), are

$$W(\alpha) = \frac{2}{\pi}\exp(-2|\alpha - \beta|^2) \tag{3.145}$$

for the coherent state $|\beta\rangle$, and for a number state $|n\rangle$

$$W(\alpha) = \frac{2}{\pi}(-1)^n L_n(4|\alpha|^2)\exp(-2|\alpha|^2), \tag{3.146}$$

where $L_n(\zeta)$ is a Laguerre polynomial. (The derivation of these functions is left as an exercise. See Problem 3.12.) We plot these functions in Fig. 3.10, again with $\alpha = x + iy$. Evidently, the Q and Wigner functions for the coherent state are identical apart from an overall scale factor. But for the number state we see that the Wigner function oscillates and becomes negative over a wide region of phase space. The Q function, of course, can never become negative. It is always a probability distribution. But the Wigner function is not always positive, the Wigner function of the number state being a case in point, in which case it is not a probability distribution. A state whose Wigner function takes on negative values over some region of phase space is *nonclassical.* However, the converse is not necessarily true. A state can be nonclassical yet have a non-negative Wigner function. As we said earlier, for nonclassical states, the P function becomes negative or more singular than a delta function over some region of phase space. The squeezed states are strongly nonclassical in this sense, as we shall discuss in Chapter 7, yet their Wigner functions are always positive. Nevertheless, for these and other nonclassical states, the Wigner function is still of primary interest as the P function generally cannot be written down as a function in the usual sense, whereas this can always be done for the Wigner function. Furthermore, the Wigner function can be more sensitive to the quantum nature of some states than can the Q function, as we have seen for the number state. More importantly, as we shall see in Chapter 7 (and below in Problem 12), the Wigner function can display the *totality* of interference effects associated with a quantum state [15]. Finally, it happens that it is possible to reconstruct the Wigner from experimental

data, a procedure known as quantum-state tomography [16]. A discussion of this procedure would take us beyond the scope of this book, but we refer the reader to the references and bibliography for this chapter.

Problems

1. Investigate the possible existence of *right* eigenstates of the creation operator $\hat{a}^\dagger$.
2. Show, perhaps with the help of phase-space graphics, for a coherent state $|\alpha\rangle$, of average photon number $\bar{n} = |\alpha|^2$, that the phase uncertainty for the state is $\Delta\phi = 1/(2\bar{n})^{1/2}$ if $\bar{n} \gg 1$.
3. Complete the derivation of the wave function for the coherent state given by Eq. (3.47).
4. Prove the following identities:

$$\hat{a}^\dagger |\alpha\rangle \langle\alpha| = \left(\alpha^* + \frac{\partial}{\partial\alpha}\right) |\alpha\rangle \langle\alpha| ,$$
$$|\alpha\rangle \langle\alpha| \hat{a} = \left(\alpha + \frac{\partial}{\partial\alpha^*}\right) |\alpha\rangle \langle\alpha| .$$

5. Verify Eq. (3.16), that the quantum fluctuations of the field quadrature operators are the same as for the vacuum when the field is in a coherent state.
6. For a coherent state $|\alpha\rangle$, evaluate the factorial moments

$$\langle \hat{n}\,(\hat{n}-1)\,(\hat{n}-2)\ldots(\hat{n}-r+1)\rangle .$$

7. For a coherent state $|\alpha\rangle$ evaluate the sine and cosine operators of Eqs. (2.211) and their squares. (You will not get closed forms.) Examine the limit where the average photon number $\bar{n} = |\alpha|^2 \gg 1$. Then examine the uncertainty products of Eqs. (2.215) and (2.216) in this limit.
8. Consider again the exponential phase operator $\hat{E} = (\hat{n}+1)^{-1/2}\hat{a}$. (a) Seek normalized right eigenstates $|z\rangle$ of this operator satisfying the equation $\hat{E}|z\rangle = z|z\rangle$ where z is a complex number. Is there a restriction on the range of $|z|$? (b) Can one resolve unity with these states? (c) Examine the states for the case where $|z| \to 1$. How is this special case related to the phase states $|\phi\rangle$ of Eq. (2.221)? Is there a conflict with Eq. (2.223)? (d) Obtain the photon number distribution for this state and express it in terms of the average photon number $\bar{n}$ for the state. Does it look like anything you may have already seen in this book? (e) Obtain the phase distribution for $|z\rangle$.
9. Work out the normally ordered variance of the photon number operator, $\langle : (\Delta\hat{n})^2 : \rangle \equiv \langle : \hat{n}^2 : \rangle - \langle : \hat{n} : \rangle^2$, where $: \hat{n} := \hat{n} = \hat{a}^\dagger\hat{a}$ is already normally ordered, in terms of the P function. Show that for a coherent state $\langle : (\Delta\hat{n})^2 : \rangle = 0$. Suppose now that for some quantum state we find that $\langle : (\Delta\hat{n})^2 : \rangle < 0$. What does this say about the P function for the state? Can such a state be considered a classical state?
10. Work out the normally ordered variances $\langle : (\Delta\hat{X})_i^2 : \rangle$, $i = 1, 2$, of the quadrature operators in terms of the P function. Show that these variances vanish for a coherent state. Examine the conditions under which some quantum state might yield $\langle : (\Delta\hat{X})_i^2 : \rangle < 0$ for either $i = 1$ or 2. Can this condition hold simultaneously for both quadratures?

11. Show that the two expressions given for the Wigner function, Eqs. (3.116) and (3.136), are equivalent.
12. Carry through the derivation of the Wigner functions of Eqs. (3.145) and (3.146) for the coherent state and the number state, respectively.
13. Consider the superposition state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\beta\rangle + |-\beta\rangle\right)$$

where the $|\pm\beta\rangle$ are coherent states. (a) Show that this state is normalized for the case where $|\beta|^2 \gg 1$. (b) Obtain the photon number probability distribution. (c) Obtain the phase distribution. (d) Obtain the Q and Wigner functions for this state and display them as three-dimensional plots. Is $|\psi\rangle$ a classical state?

14. Show that for some state $|\psi\rangle$, the Wigner function can be written in the form

$$W(\alpha) = \frac{2}{\pi}\sum_{n=0}^{\infty}(-1)^n \langle\psi|\,\hat{D}(\alpha)\,|n\rangle\langle n|\,\hat{D}^{\dagger}(\alpha)\,|\psi\rangle\,.$$

References

[1] J. J. Sakurai, *Advanced Quantum Mechanics* (Reading: Addison-Wesley, 1967), p. 35.

[2] R. J. Glauber, *Phys. Rev*., **131** (1963), 2766; J. R. Klauder, *J. Math. Phys*., **4** (1963), 1055.

[3] E. Merzbacher, *Quantum Mechanics*, 2nd edition (New York: Wiley, 1970), pp. 158–161.

[4] See C. Aragone, E. Chalbaud and S. Salamó, *J. Math. Phys*., **17** (1976), 1963.

[5] See S. M. Barnett and D. T. Pegg, *J. Mod. Opt*., **36** (1989), 7.

[6] The wave functions for the energy eigenstates for the one-dimensional harmonic oscillator can be found in any standard quantum-mechanics textbook.

[7] See, for example, M. Nauenberg, *Phys. Rev. A*, **40** (1989), 1133; Z. Dačić Gaeta and C. R. Stroud, Jr., *Phys. Rev. A*, **42** (1990), 6308.

[8] See L. W. Ryder, *Quantum Field Theory*, 2nd edition (Cambridge: Cambridge University Press, 1996), p. 178.

[9] I. Segal, *Illinois J. Math*., **6** (1962), 500; V. Bargmann, *Commun. Pure Appl. Math*., **14** (1961), 187.

[10] E. C. G. Sudarshan, *Phys. Rev. Lett.*, **10** (1963), 277.

[11] C. L. Mehta, *Phys. Rev. Lett*., **18** (1967), 752.

[12] Y. Kano, *J. Math. Phys*., **6** (1965), 1913; C. L. Mehta and E. C. G. Sudarshan, *Phys. Rev*., **138** (1965), B274; R. J. Glauber, in *Quantum Optics and Electronics*, edited by C. Dewitt, A. Blandin and C. Cohen-Tannoudji (New York: Gordon and Breach, 1965), p. 65.

[13] E. P. Wigner, *Phys. Rev*., **40** (1932), 794.

[14] K. E. Cahill and R. J. Glauber, *Phys. Rev.*, **177** (1969), 1882.

[15] G. S. Agarwal, *Found. Phys*., **25** (1995), 219.

[16] See K. Vogel and H. Risken, *Phys. Rev. A*, **40** (1989), 2847; D. T. Smithey, M. Beck, M. G. Raymer and A. Faridani, *Phys. Rev. Lett*., **70** (1993), 1244; K. Banaszek, C. Radzewicz, K. Wodkiewicz and J. S. Krasinski, *Phys. Rev. A*, **60** (1999), 674.

Bibliography

Coherent states

J. R. Klauder and B.-S. Skagerstam (editors), *Coherent States*: *Applications in Physics and Mathematical Physics* (Singapore: World Scientific, 1985).

W.-M. Zhang, D. H. Feng and R. Gilmore, *Rev. Mod. Phys*., **62** (1990), 867.

Quasi-probability distributions

G. S. Agarwal and E. Wolf, *Phys. Rev. D*, **2** (1970), 2161.

M. Hillery, R. F. O'Connell, M. O. Scully and E. P. Wigner, *Phys. Rep.*, **106** (1984), 121.

Quantum-state tomography

M. G. Raymer, *Contemp. Phys*., **38** (1997), 343.

U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge: Cambridge University Press, 1997).

Chapter 4

Emission and absorption of radiation by atoms

In this chapter, we first discuss atom–field interactions using quantum mechanical perturbation theory for a classical driving field and then for a quantum mechanical field. In the latter case, spontaneous emission appears as a fully quantum-mechanical phenomenon. We then examine the so-called Rabi model, a model of a "two-level" atom interacting with a strong near-resonant classical field, introduce the rotating wave approximation, and then introduce the fully quantum mechanical Rabi model, better known as the Jaynes–Cummings model. Differences in the evolution predicted by the two-models are drawn out, where we show that the Jaynes–Cummings model predicts behavior that has no semiclassical analog and that depends entirely on the discreteness of photons. Finally, we study an extension of the Jaynes–Cummings model, the dispersive model, for the case where the quantized field is far out-of-resonance with the atomic transition frequency. This will eventually allow us to describe how superpositions of different coherent states of the field, states known as Schrödinger cat states, can be generated from atom–field interactions.

4.1 Atom–field interactions

To begin, let us suppose that the Hamiltonian of an electron bound to an atom in the absence of external fields, in the configuration representation, is given by

$$\hat{H}_0 = \frac{1}{2m}\hat{\mathbf{P}}^2 + V(r) \tag{4.1}$$

where $V(r)$ is the usual Coulomb interaction binding the electron to the nucleus and $r = |\mathbf{r}|$. In the configuration space representation $\hat{\mathbf{P}} = -i\nabla$, $\hat{\mathbf{r}}|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle$ and the wave functions are given by $\psi(\mathbf{r}) = \langle \mathbf{r} \mid \psi \rangle$. We assume that energy eigenstates $|k\rangle$ of $\hat{H}_0$, satisfying the time-independent Schrödinger equation

$$\hat{H}_0 \psi_k^{(0)}(\mathbf{r}) = E_k \psi_k^{(0)}(\mathbf{r}), \tag{4.2}$$

where $\langle \mathbf{r} \mid k \rangle = \psi_k^{(0)}(\mathbf{r})$, are known. In the presence of external fields the Hamiltonian is modified to

$$\hat{H}(\mathbf{r}, t) = \frac{1}{2m}[\hat{\mathbf{P}} + e\mathbf{A}(\mathbf{r}, t)]^2 - e\,\Phi(\mathbf{r}, t) + V(r) \tag{4.3}$$

where $\mathbf{A}(\mathbf{r}, t)$ and $\Phi(\mathbf{r}, t)$ are the vector and scalar potentials respectively of the external field and where $-e$ is the electron charge, e taken to be positive. The fields themselves are given by

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= -\nabla\Phi(\mathbf{r}, t) - \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}, \\ \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t), \end{aligned} \tag{4.4}$$

and are invariant under the gauge transformations

$$\begin{aligned} \Phi'(\mathbf{r}, t) &= \Phi(\mathbf{r}, t) - \frac{\partial \chi(\mathbf{r}, t)}{\partial t}, \\ \mathbf{A}'(\mathbf{r}, t) &= \mathbf{A}(\mathbf{r}, t) + \nabla\chi(\mathbf{r}, t). \end{aligned} \tag{4.5}$$

The time-dependent Schrödinger equation is

$$\hat{H}(\mathbf{r}, t)\Psi(\mathbf{r}, t) = i\,\hbar\frac{\partial \Psi(\mathbf{r}, t)}{\partial t}. \tag{4.6}$$

In order eventually to simplify the form of the atom–field interaction, we define a unitary operator $\hat{R}$ such that $\Psi'(\mathbf{r}, t) \equiv \hat{R}\,\Psi(\mathbf{r}, t)$. We have

$$\hat{H}'\Psi'(\mathbf{r}, t) = i\,\hbar\frac{\partial \Psi'(\mathbf{r}, t)}{\partial t} \tag{4.7}$$

where

$$\hat{H}' = \hat{R}\hat{H}\hat{R}^{\dagger} + i\,\hbar\frac{\partial \hat{R}}{\partial t}\hat{R}^{\dagger}. \tag{4.8}$$

We now choose $\hat{R} = \exp(-ie\chi(\mathbf{r}, t)/\hbar)$ so that (using $\hat{\mathbf{P}} = -i\,\hbar\nabla$)

$$\hat{H}' = \frac{1}{2m}[\hat{\mathbf{P}} + e\mathbf{A}']^2 - e\,\Phi' + V(r) \tag{4.9}$$

where $\mathbf{A}'$ and Φ' are given by Eq. (4.5). At this point we make a definite choice of gauge, namely the Coulomb (or radiation) gauge, for which $\Phi = 0$ and $\mathbf{A}$ satisfies the transversality condition $\nabla \cdot \mathbf{A} = 0$. The vector potential $\mathbf{A}$, for no sources near the atom, satisfies the wave equation

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = 0. \tag{4.10}$$

This choice of gauge is not relativistically invariant, in contrast to the Lorentz gauge, but the domain of quantum optics is, for the most part, nonrelativistic so that no inconsistency will be introduced. The Coulomb gauge has the advantage that the radiation field is completely described by the vector potential, as is obvious from Eq. (4.3), which, in this gauge, reads

$$\begin{aligned} \hat{H}(\mathbf{r}, t) &= \frac{1}{2m}[\hat{\mathbf{P}} + e\mathbf{A}(\mathbf{r}, t)]^2 + V(r) \\ &= \frac{\hat{\mathbf{P}}^2}{2m} + \frac{e}{m}\,\mathbf{A}\cdot\hat{\mathbf{P}} + \frac{e^2}{2m}\mathbf{A}^2 + V(r). \end{aligned} \tag{4.11}$$

Equation (4.9) now reads

$$\hat{H}'(\mathbf{r},t) = \frac{1}{2m}[\hat{\mathbf{P}} + e(\mathbf{A} + \nabla\chi)]^2 + e\frac{\partial\chi}{\partial t} + V(r). \tag{4.12}$$

The solution of the wave equation (4.10) has the form

$$\mathbf{A} = \mathbf{A}_0\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c.c. \tag{4.13}$$

where $|\mathbf{k}| = 2\pi/\lambda$ is the wave vector of the radiation. For $|\mathbf{r}|$ of typical atomic dimensions (a few Ångströms) and λ of typical optical wavelengths (a few hundred nanometers in the range 400–700 nm), $\mathbf{k}\cdot\mathbf{r} \ll 1$ so that over the extent of an atom, the vector potential is spatially uniform, $\mathbf{A}(\mathbf{r},t) \simeq \mathbf{A}(t)$. This is the so-called dipole approximation. We now choose the gauge function $\chi(\mathbf{r},t) = -\mathbf{A}(t)\cdot\mathbf{r}$. With this choice,

$$\begin{aligned} \nabla\chi(\mathbf{r},t) &= -\mathbf{A}(t), \\ \frac{\partial\chi}{\partial t}(\mathbf{r},t) &= -\mathbf{r}\cdot\frac{\partial\mathbf{A}}{\partial t} = -\mathbf{r}\cdot\mathbf{E}(t), \end{aligned} \tag{4.14}$$

and thus

$$\hat{H}' = \frac{\hat{\mathbf{P}}^2}{2m} + V(r) + e\,\mathbf{r}\cdot\mathbf{E}(t). \tag{4.15}$$

This equation contains only one interaction term (within the dipole approximation) as opposed to the two terms of Eq. (4.11). We shall work with $\hat{H}'$ in all that follows, the interaction being in what is sometimes called the "length" gauge. The quantity $-e\mathbf{r}$ is the dipole moment: $\mathbf{d} = -e\mathbf{r}$. In general, i.e. for an unspecified representation, the dipole moment is an operator, $\hat{\mathbf{d}}$. We shall denote it as such in what follows. Thus we write

$$\hat{H}' = \hat{H}_0 - \hat{\mathbf{d}}\cdot\mathbf{E}(t) \tag{4.16}$$

where $\hat{H}_0$ is given by Eq. (4.1).

4.2 Interaction of an atom with a classical field

So far, we have not specified the nature of the interacting field and have not even stated whether we consider the field to be classical or quantum mechanical. The derivation leading to Eq. (4.16) is valid for both classical and quantum fields. But eventually we want to demonstrate differences in the way an atom behaves when interacting with classical or quantum fields. With that in mind, we first turn to the case when an atom is driven by a classical sinusoidal electric field.

We assume that the field has the form $\mathbf{E}(t) = \mathbf{E}_0\cos(\omega t)$, ω being the frequency of the radiation, and that this field is abruptly turned on at $t = 0$. The

dipole approximation, where we assume that $\mathbf{k}\cdot\mathbf{r} \ll 1$ over the atom, has already been made. We further assume that the initial state of the atom is $|i\rangle$ where $\hat{H}_0|i\rangle = E_i|i\rangle$. For times $t > 0$ we expand the state vector $|\psi(t)\rangle$ in terms of the complete set of uncoupled atomic states $|k\rangle$:

$$|\psi(t)\rangle = \sum_k C_k(t) e^{-iE_k t/\hbar}|k\rangle, \tag{4.17}$$

where the time-dependent amplitudes $C_k(t)$ satisfy the normalization requirement

$$\sum_k |C_k(t)|^2 = 1. \tag{4.18}$$

Substituting this expansion into the time-dependent Schrödinger equation

$$i\hbar\frac{\partial|\psi(t)\rangle}{\partial t} = \left(\hat{H}_0 + \hat{H}^{(\mathrm{I})}\right)|\psi(t)\rangle, \tag{4.19}$$

where $H^{(\mathrm{I})} = -\hat{\mathbf{d}}\cdot\mathbf{E}(t)$, then multiplying from the left by $\langle l|e^{iE_l t/\hbar}$ leads to the set of coupled first-order differential equations for the amplitudes

$$\dot{C}_l(t) = -\frac{i}{\hbar}\sum_k C_k(t)\langle l|\hat{H}^{(\mathrm{I})}|k\rangle e^{i\omega_{lk}t} \tag{4.20}$$

where the $\omega_{lk} = (E_l - E_k)/\hbar$ are the transition frequencies between levels l and k. These equations are, so far, exact and need to be solved subject to the initial condition $C_i(0) = 1$, only state $|i\rangle$ being initially populated. As time goes forward, population will be lost from state $|i\rangle$ and increased in some initially unpopulated state $|f\rangle$, i.e. the amplitude $C_f(t)$ increases. The probability for the atom to make a transition from state $|i\rangle$ to state $|f\rangle$ in time t is given by

$$P_{i\to f}(t) = C_f^*(t)C_f(t) = |C_f(t)|^2. \tag{4.21}$$

The equations for the amplitudes are solvable only for very simple cases. These days, of course, one might solve the set of differential equations numerically, but in the case of a driving field that in some sense is "weak", we can use a time-dependent perturbation theory [1] approach to the problem. "Weak" in this case means that $|\mathbf{E}_0|$ is small, or actually, that $|\langle f|\hat{\mathbf{d}}\cdot\mathbf{E}_0|i\rangle|$ is small. As a matter of bookkeeping, we write the interaction Hamiltonian as $\lambda\hat{H}^{(\mathrm{I})}$, where λ is treated as a number in the range $0 \le \lambda \le 1$. (At the end of the calculations we will always take $\lambda \to 1$.) We then expand the probability amplitude for, say, state $|l\rangle$ in the power series

$$C_l(t) = C_l^{(0)}(t) + \lambda C_l^{(1)}(t) + \lambda^2 C_l^{(2)}(t) + \cdots \tag{4.22}$$

Inserting such expansions into Eqs. (4.20) and equating like powers of λ we obtain, up to second order,

$$\dot{C}_l^{(0)} = 0, \tag{4.23}$$

$$\dot{C}_l^{(1)} = -\frac{i}{\hbar}\sum_k C_k^{(0)} H_{lk}^{(\mathrm{I})}(t) e^{i\omega_{lk}t}, \tag{4.24}$$

$$\dot{C}_l^{(2)} = -\frac{i}{\hbar}\sum_k C_k^{(1)} H_{lk}^{(\mathrm{I})}(t) e^{i\omega_{lk}t}, \tag{4.25}$$

where $H_{lk}^{(\mathrm{I})}(t) \equiv \langle l|\hat{H}^{(\mathrm{I})}(t)|k\rangle$. Note the general pattern that relates the nth order to the $(n-1)$th order:

$$\dot{C}_l^{(n)} = -\frac{i}{\hbar}\sum_k C_k^{(n-1)}(t) H_{lk}^{(\mathrm{I})}(t) e^{i\omega_{lk}t}. \tag{4.26}$$

The essential assumption underlying the perturbation-theory approach is that the driving field is so weak that the atomic populations change very little. That is, if $C_i(0) = 1,\ C_f(0) = 0\ (f \neq i)$ then for $t > 0$, to a good approximation $C_i(t) \approx 1, |C_f(t)| \ll 1\ (f \neq i)$. Thus, in the first-order equation (4.24), the only term surviving the sum on the right-hand side is for $k = i$ yielding

$$\dot{C}_f^{(1)}(t) = -\frac{i}{\hbar} H_{fi}^{(\mathrm{I})}(t) e^{i\omega_{fi}t} C_i^{(0)}(t) \tag{4.27}$$

or

$$C_f^{(1)}(t) = -\frac{i}{\hbar}\int_0^t dt' H_{fi}^{(\mathrm{I})}(t') e^{i\omega_{fi}t'} C_i^{(0)}(t'). \tag{4.28}$$

Inserting this result into the second-order equation, Eq. (4.25), we obtain

$$\begin{aligned} C_f^{(2)}(t) &= -\frac{i}{\hbar}\sum_l \int_0^t dt' H_{fl}^{(\mathrm{I})}(t') e^{i\omega_{fl}t'} C_l^{(1)}(t') \\ &= \left(-\frac{i}{\hbar}\right)^2 \sum_l \int_0^t dt' \int_0^{t'} dt'' H_{fl}^{(\mathrm{I})}(t') e^{i\omega_{fl}t'} \\ &\quad \times H_{li}^{(\mathrm{I})}(t'') e^{i\omega_{li}t''} C_i^{(0)}(t''). \end{aligned} \tag{4.29}$$

Equation (4.28) gives the amplitude for a transition from state $|i\rangle$ to state $|f\rangle$ while Eq. (4.29) gives the amplitude for a transition from state $|i\rangle$ to states $\{|l\rangle\}$ then to state $|f\rangle$. The total transition probability from state $|i\rangle$ to state $|f\rangle$ is

$$P_{i\to f}(t) = \left|C_f^{(0)}(t) + C_f^{(1)}(t) + C_f^{(2)}(t) + \cdots\right|^2. \tag{4.30}$$

Now the dipole moment operator $\hat{\mathbf{d}}$ has nonvanishing matrix elements only between states of opposite parity. Thus the first-order correction to the amplitude

of the initial state vanishes:

$$C_i^{(1)}(t) = -\frac{i}{\hbar}\int_0^t dt' H_{ii}^{(\mathrm{I})}(t') C_i^{(0)}(t') = 0 \tag{4.31}$$

because $H_{ii}^{(\mathrm{I})}(t) = 0$. Therefore, to first order $C_i(t) = C_i^{(0)}(t) = 1$ so that

$$C_f^{(1)}(t) = -\frac{i}{\hbar}\int_0^t dt' H_{fi}^{(\mathrm{I})}(t')\, e^{i\omega_{fi}t'}. \tag{4.32}$$

With $H^{(\mathrm{I})} = -\hat{\mathbf{d}}\cdot\mathbf{E}_0\cos\omega t$, and by expanding the cosine in terms of exponentials, this integrates to

$$C_f^{(1)}(t) = \frac{1}{2\hbar}(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi} \times\left\{\frac{\left(e^{i(\omega+\omega_{fi})t}-1\right)}{(\omega+\omega_{fi})} - \frac{\left(e^{-i(\omega-\omega_{fi})t}-1\right)}{(\omega-\omega_{fi})}\right\}, \tag{4.33}$$

where $(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi} = \langle f|\hat{\mathbf{d}}\cdot\mathbf{E}_0|i\rangle$. If the frequency of the radiation, ω, is near resonance with the atomic transition frequency ω_{fi}, the second term clearly dominates the first. Therefore we may drop the "antiresonant" first term, making the so-called "rotating wave approximation" (RWA), familiar in the context of magnetic resonance [2]. Thus we have to first order the transition probability

$$P_{i\to f}^{(1)}(t) = \left|C_f^{(1)}(t)\right|^2 = \frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{\hbar^2}\frac{\sin^2(\Delta t/2)}{\Delta^2} \tag{4.34}$$

where $\Delta = \omega - \omega_{fi}$ is the "detuning" between the radiation field and the atomic transition. When $\Delta \neq 0$, $P_{i\to f}^{(1)}(t)$ maximizes at

$$\left(P_{i\to f}^{(1)}\right)_{\max} = \frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{\hbar^2}\frac{1}{\Delta^2}. \tag{4.35}$$

For the case of exact resonance, $\Delta = 0$,

$$\left(P_{i\to f}^{(1)}\right)_{\max} = \frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{4\hbar^2}t^2\,. \tag{4.36}$$

For the perturbation expansion to be valid we must have $(P_{i\to f}^{(1)})_{\max} \ll 1$. For the off-resonance case, this places conditions on both $|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|$ and Δ. For the resonant case, Eq. (4.36) is valid only for very short times. In Fig. 4.1 we plot the

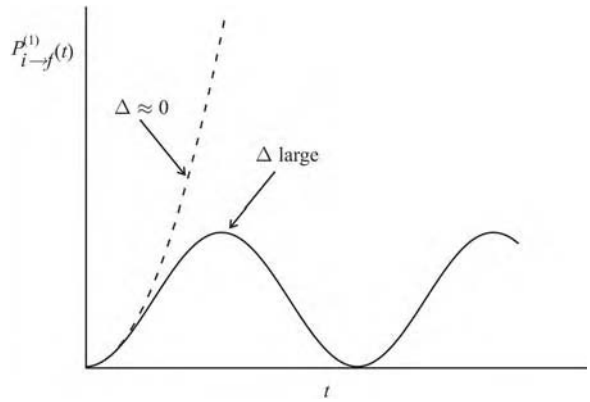


Fig. 4.1. A plot of $P^{(1)}_{i\to f}(t)$ versus time for small and large detunings Δ.

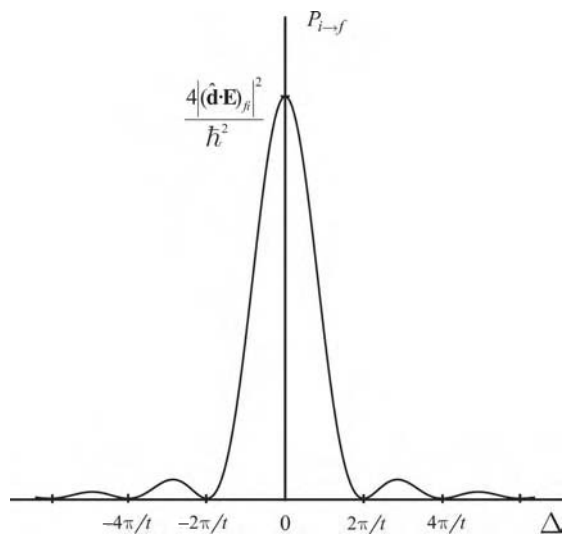


Fig. 4.2. The transition probability $P^{(1)}_{i\to f}(t)$ as a function of Δ.

evolution of the probability distribution $P^{(1)}_{i\to f}(t)$ for both small detuning ($\Delta \approx 0$) and large detuning. The latter is periodic. The transition probability $P^{(1)}_{i\to f}(t)$ is a sharply peaked function of the detuning at $\Delta = 0$ as shown in Fig. 4.2. The width of the peak is proportional to t^{-1} while the height is proportional to t^2. Thus the area under the peak is proportional to t. In fact

$$\int_{-\infty}^{\infty} \frac{\sin^2(\Delta t/2)}{\Delta^2} d(\Delta) = \frac{\pi}{2} t. \tag{4.37}$$

Furthermore, in the limit when $\Delta \approx 0$ and $t \gg 2\pi/\omega_{fi}$, the function in the integrand of Eq. (4.37) may be approximated by a Dirac delta function:

$$\lim_{t\to\infty} \frac{\sin^2(\Delta t/2)}{\Delta^2} = \frac{\pi}{2} t\, \delta(\Delta), \tag{4.38}$$

although the limit $t \to \infty$ is, in fact, actually constrained by the requirement that the right-hand side of Eq. (4.35) be $\ll 1$. In this case the transition probability is

$$P^{(1)}_{i\to f}(t) = \frac{\pi}{2}\frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{\hbar^2} t\,\delta(\omega-\omega_{fi}). \tag{4.39}$$

We define the time-independent transition probability *rate* as

$$W_{i\to f} = \frac{P^{(1)}_{i\to f}}{t} = \frac{\pi}{2}\frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{\hbar^2}\,\delta(\omega-\omega_{fi}). \tag{4.40}$$

In practice, there will be a broad range of final states $|f\rangle$ accessible from the initial state, and the driving field will not be monochromatic so that a range of frequencies will have to be summed or integrated over to obtain the total transition rate. If $[f]$ represents a set of accessible final states, then the transition rate for a monochromatic field is

$$W_{i\to[f]} = \frac{\pi}{2}\sum_{[f]}\frac{|(\hat{\mathbf{d}}\cdot\mathbf{E}_0)_{fi}|^2}{\hbar^2}\,\delta(\omega-\omega_{fi}). \tag{4.41}$$

This expression is often famously called Fermi's Golden Rule [3].

Now suppose that the light irradiating the atom is from a lamp emitting a broad range of frequencies where there is no phase relationship between the different frequency components. The amplitude of the light will now be frequency dependent so that the transition probability rate induced by all the frequency components must be

$$\frac{P^{(1)}_{i\to f}(t)}{t} = \frac{1}{\hbar^2}\int d\omega \frac{\sin^2(\Delta t/2)}{\Delta^2}F(\omega) \tag{4.42}$$

where

$$F(\omega) \equiv |\langle f|\hat{\mathbf{d}}\cdot\mathbf{E}_0(\omega)|i\rangle|^2. \tag{4.43}$$

If the function $F(\omega)$ is broadband and varies slowly with ω in comparison to $(\sin^2(\Delta t/2)/\Delta^2)$, then $F(\omega)$ can be replaced by its resonance value $F(\omega_{fi})$ and taken outside the integral so that

$$P^{(1)}_{i\to f}(t) = \frac{\pi}{2\hbar^2}F(\omega_{fi})\,t. \tag{4.44}$$

Thus the transition rate is

$$W_{i\to f} = \frac{\pi}{2\hbar^2}F(\omega_{fi}). \tag{4.45}$$

The spread of frequencies results in the washing out, or dephasing, of the oscillations seen in Fig. 4.1. This happens because of the lack of phase relations between the various frequency components – the light is incoherent. If the atom is driven by a coherent light field, such as from a laser, dephasing does not occur and

the perturbative time-independent transition rates above generally do not adequately describe the dynamics. We postpone a discussion of driving the atom by a coherent laser field until Section 4.4.

4.3 Interaction of an atom with a quantized field

In the preceding discussion, we made no assumption regarding the relative positions of the energy levels i and f. Transitions between the levels occur with some nonzero probability as long as $\mathbf{E}_0 \neq 0$ whether $E_i < E_f$ or $E_i > E_f$. As we will show, when the field is quantized, transitions will occur for the case $E_i > E_f$ even when no photons are present – the so-called spontaneous emission. This is only one of several differences that will appear in the atom–field dynamics in the comparison between cases when the field is quantized and when it is not.

We consider a single free-space field mode of the form given by Eq. (2.130)

$$\hat{\mathbf{E}}(t) = i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{1/2} \mathbf{e}[\hat{a}e^{-i\omega t} - \hat{a}e^{i\omega t}] \tag{4.46}$$

where the dipole approximation has been made. This operator is in the Heisenberg picture but we are going to be working in the Schrödinger picture where the field operator is

$$\hat{\mathbf{E}} = i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{1/2} \mathbf{e}(\hat{a} - \hat{a}^\dagger). \tag{4.47}$$

The free Hamiltonian $\hat{H}_0$ now must be

$$\hat{H}_0 = \hat{H}_{\text{atom}} + \hat{H}_{\text{field}} \tag{4.48}$$

where $\hat{H}_{\text{atom}}$ is just the free-atom Hamiltonian as before and $\hat{H}_{\text{field}}$ is the free-field Hamiltonian $\hbar\omega\,\hat{a}^\dagger\hat{a}$, where the zero-point term has been dropped as it does not contribute to the dynamics. The interaction Hamiltonian becomes

$$\begin{aligned}\hat{H}^{(\mathrm{I})} &= -\hat{\mathbf{d}}\cdot\hat{\mathbf{E}} = -i\left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{1/2}(\hat{\mathbf{d}}\cdot\mathbf{e})(\hat{a} - \hat{a}^\dagger) \\ &= -\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}_0(\hat{a} - \hat{a}^\dagger)\end{aligned} \tag{4.49}$$

where $\boldsymbol{\mathcal{E}}_0 = i(\hbar\omega/2\varepsilon_0 V)^{1/2}\mathbf{e}$.

Because both atomic and field systems are now quantized, the states of the combined system will involve products of states of both systems. Suppose the initial state of the atom–field system is $|i\rangle = |a\rangle|n\rangle$ where $|a\rangle$ is the initial state of the atom and the field contains n photons. The perturbation interaction of the quantized field causes a transition to the state $|f_1\rangle = |b\rangle|n-1\rangle$, where $|b\rangle$ is another atomic state, by the absorption of a photon or to the state

$|f_2\rangle = |b\rangle|n+1\rangle$ by the emission of a photon. The energies of these states are,

$$\text{for} \quad |i\rangle = |a\rangle|n\rangle, \qquad E_i = E_a + n\,\hbar\omega, \tag{4.50a}$$

$$\text{for} \quad |f_1\rangle = |b\rangle|n-1\rangle, \qquad E_{f_1} = E_b + (n-1)\,\hbar\omega, \tag{4.50b}$$

$$\text{for} \quad |f_2\rangle = |b\rangle|n+1\rangle, \qquad E_{f_2} = E_b + (n+1)\,\hbar\omega, \tag{4.50c}$$

where E_a and E_b are the energies of the atomic states $|a\rangle$ and $|b\rangle$ respectively.

The perturbation given by Eq. (4.49) is now time independent. The matrix elements of the interaction are as follows:

$$\begin{aligned}\langle f_1|\hat{H}^{(\mathrm{I})}|i\rangle &= \langle b, n-1|\hat{H}^{(\mathrm{I})}|a, n\rangle \\ &= -(\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}_0)_{ba}\sqrt{n}, \qquad \text{(absorption)}\end{aligned} \tag{4.51}$$

and

$$\begin{aligned}\langle f_2|\hat{H}^{(\mathrm{I})}|i\rangle &= \langle b, n+1|\hat{H}^{(\mathrm{I})}|a, n\rangle \\ &= (\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}_0)_{ba}\sqrt{n+1}, \qquad \text{(emission)}\end{aligned} \tag{4.52}$$

where

$$(\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}_0)_{ab} = \langle a|\hat{\mathbf{d}}|b\rangle\cdot\boldsymbol{\mathcal{E}}_0 \equiv \mathbf{d}_{ab}\cdot\boldsymbol{\mathcal{E}}_0 \tag{4.53}$$

the factor $\langle a|\hat{\mathbf{d}}|b\rangle = \mathbf{d}_{ab}$ being the dipole matrix element between states $|a\rangle$ and $|b\rangle$. In comparison with the semiclassical case, two things are noteworthy here. The absence of photons ($n = 0$) precludes absorption, just as one might expect. This is obviously in agreement with the case of a classical driving field – no field, no transitions. But in the case of emission, according to Eq. (4.52), transitions may occur even when no photons are present. This is spontaneous emission and it has no semiclassical counterpart. If $n > 0$, the emission of an additional photon is called stimulated emission, this process being the essential one for the operation of the laser (or LASER: Light Amplification by Stimulated Emission of Radiation). The rates of emission and absorption are proportional to the moduli squared of the above matrix elements for the respective processes. The ratio of these rates is

$$\frac{\left|\langle f_2|\hat{H}^{(\mathrm{I})}|i\rangle\right|^2}{\left|\langle f_1|\hat{H}^{(\mathrm{I})}|i\rangle\right|^2} = \frac{n+1}{n}, \tag{4.54}$$

a result to be used shortly.

The perturbation method developed previously can still be used with appropriate modifications to accommodate the fact that the field is now quantized. The Schrödinger equation still has the form of Eq. (4.19) but where now $\hat{H}_0$ is given by Eq. (4.48) and $\hat{H}^{(\mathrm{I})}$ by Eq. (4.49). Ignoring all other atomic states except $|a\rangle$

and $|b\rangle$, the state vector can be written as

$$\begin{aligned}|\psi(t)\rangle = & C_i(t)|a\rangle|n\rangle e^{-iE_at/\hbar}e^{-in\omega t} \\ & + C_{f_1}(t)|b\rangle|n-1\rangle e^{-iE_bt/\hbar}e^{-i(n-1)\omega t} \\ & + C_{f_2}(t)|b\rangle|n+1\rangle e^{-iE_bt/\hbar}e^{-i(n+1)\omega t}\end{aligned} \tag{4.55}$$

where, assuming that $|\psi(0)\rangle = |a\rangle|n\rangle$, $C_i(0) = 1$ and $C_{f1}(0) = C_{f2}(0) = 0$. Following the perturbative method used before, we obtain the first-order correction for the amplitudes C_{f1} and C_{f2} associated with the atom being in state $|b\rangle$:

$$\begin{aligned}C_{f1}^{(1)}(t) &= -\frac{i}{\hbar}\int_0^t dt'\langle f_1|\hat{H}^{(\mathrm{I})}|i\rangle e^{i(E_{f1}-E_i)t/\hbar}, \\ C_{f2}^{(1)}(t) &= -\frac{i}{\hbar}\int_0^t dt'\langle f_2|\hat{H}^{(\mathrm{I})}|i\rangle e^{i(E_{f2}-E_i)t/\hbar},\end{aligned} \tag{4.56}$$

where the former is associated with absorption and the latter with emission. The amplitude associated with the atom being in the state $|b\rangle$, regardless of how it got there, is just the sum of the amplitudes in Eqs. (4.56), i.e. $C_f^{(1)} = C_{f1}^{(1)} + C_{f2}^{(1)}$. From Eqs. (4.50) we obtain

$$C_f^{(1)}(t) = \frac{i}{\hbar}(\hat{\mathbf{d}}\cdot\boldsymbol{\mathcal{E}}_0)_{ab}\left\{(n+1)^{1/2}\frac{\left[e^{i(\omega+\omega_{ba})t}-1\right]}{(\omega+\omega_{ba})} - n^{1/2}\frac{\left[e^{i(\omega-\omega_{ba})t}-1\right]}{(\omega-\omega_{ba})}\right\} \tag{4.57}$$

where $\omega_{ba} = (E_b - E_a)/\hbar$ and where the first term is due to emission and the second to absorption. If the number of photons is large, $n \gg 1$, then we can replace $\sqrt{n+1}$ by $\sqrt{n}$ in the first term and then Eqs. (4.57) and (4.33) are essentially the same, there being the correspondence between the classical and quantum field amplitudes $(\mathbf{E}_0)_{\mathrm{cl}} \leftrightarrow (2i\mathcal{E}_0\sqrt{n})_{\mathrm{quantum}}$. This correspondence between quantum and classical fields has its limits, one being the case when $n = 0$, as already discussed.

If $|b\rangle$ is the excited state then $\omega_{ba} > 0$ so if $\omega \sim \omega_{ba}$ then the first term of Eq. (4.51) can be dropped, which is again the rotating wave approximation. Of course, if $|a\rangle$ is the excited state, then $\omega_{ba} < 0$ and if $\omega \sim -\omega_{ba}$, the second term of Eq. (4.57) can be dropped, and we notice that the remaining term does not vanish even when $n = 0$, the transition between $|a\rangle$ and $|b\rangle$ taking place by spontaneous emission. Thus the rotating wave approximation carries over into the case where the field and the atom are quantized. It can be shown that Fermi's Golden Rule carries over in a similar fashion.

We conclude this section with a field-theoretic derivation of the Planck distribution law. Suppose we have a collection of atoms interacting resonantly with a quantized field of frequency $\omega = (E_a - E_b)/\hbar$, where $|a\rangle$ and $|b\rangle$ are the atomic states with $E_a > E_b$. We let N_a and N_b represent the populations of atoms in states $|a\rangle$ and $|b\rangle$ respectively. Further, we let W_{emis} represent the transition rate due to photon emission and W_{abs} the transition rate due to photon absorption.

Because the atoms are constantly emitting and absorbing photons, the atomic populations change with time according to

$$\begin{aligned} \frac{dN_a}{dt} &= -N_a W_{\text{emis}} + N_b W_{\text{abs}} \\ \frac{dN_b}{dt} &= -N_b W_{\text{abs}} + N_a W_{\text{emis}}. \end{aligned} \tag{4.58}$$

At thermal equilibrium, we have

$$\frac{dN_a}{dt} = 0 = \frac{dN_b}{dt} \tag{4.59}$$

and thus we obtain

$$N_a W_{\text{emis}} = N_b W_{\text{abs}}. \tag{4.60}$$

But, according to Boltzmann,

$$\frac{N_b}{N_a} = \exp[(E_a - E_b)/kT] = \exp(\hbar\omega/kT), \tag{4.61}$$

and from Eq. (4.57) we have that

$$\frac{N_b}{N_a} = \frac{W_{\text{emis}}}{W_{\text{abs}}} = \frac{n+1}{n}. \tag{4.62}$$

Thus from Eqs. (4.61) and (4.62) it follows that

$$n = \frac{1}{\exp(\hbar\omega/kT) - 1}, \tag{4.63}$$

in agreement with Eq. (2.141) if we replace n by $\bar{n}$ in Eq. (4.63) to incorporate the fact that we cannot assume a definite number of photons.

Let us compare this derivation to the one given by Einstein [4] before quantum electrodynamics was even invented. His derivation is similar to the above but explicitly makes the distinction between spontaneous and stimulated emission. He introduced the coefficients A, B, and C having the following meanings: AN_a is the growth rate of the population in state $|b\rangle$ owing to spontaneous emission from state $|b\rangle$ (A being the rate of spontaneous emission); $BU(\omega)N_a$ is the growth rate of the population of state $|b\rangle$ owing to stimulated emission from $|a\rangle$, $U(\omega)$ being the spectral energy density of the field; and $CU(\omega)N_b$ is the rate of growth of state $|a\rangle$ as the result of absorption by atoms in state $|b\rangle$. Note that the spontaneous

emission term is independent of $U(\omega)$. The rate equations for the populations are now

$$\begin{aligned} \frac{dN_a}{dt} &= -[A + BU(\omega)]N_a + CU(\omega)N_b \\ \frac{dN_b}{dt} &= -CU(\omega)N_b + [A + BU(\omega)]N_a. \end{aligned} \tag{4.64}$$

At long times, the populations reach a steady state and the derivatives on the left-hand sides vanish to yield

$$[A + BU(\omega)]N_a = CU(\omega)N_b. \tag{4.65}$$

Using once again the relation in Eq. (4.61), we are led to

$$U(\omega) = \frac{A}{C\exp(\hbar\omega/kT) - B}. \tag{4.66}$$

But for a thermal field, we must have, by comparison to Eq. (2.151), that $C = B$ and that

$$\frac{A}{B} = \frac{\hbar\omega^3}{\pi^2 c^3}. \tag{4.67}$$

It is worth comparing the rates of spontaneous emission A and stimulated emission $BU(\omega)$:

$$\frac{A}{BU(\omega)} = \exp(\hbar\omega/kT - 1). \tag{4.68}$$

For a natural source such as the sun, which we approximate as a black body with surface temperature $T \approx 6000$ K, this ratio is about 400 for $\lambda = 400$ nm and about 30 for $\lambda = 700$ nm [5]. Thus at both ends of the visible spectrum spontaneous emission dominates stimulated emission. In the range of visible light, it is only in "unnatural" sources, where there exists a population inversion (i.e. more atoms in the excited state than in the ground), such as in lasers, where stimulated emission dominates spontaneous emission.

Spontaneous emission is a complex phenomenon and we have discussed only its most essential features. For atoms in free space there is an infinity of modes into which the atoms may radiate. Spontaneous emission in this case is well described by the Weisskopf–Wigner theory [6] as an irreversible decay process. A discussion of that theory is outside the scope of this book. However, under certain circumstances, where there may be only one mode into which an atom can radiate, such as in the case of an atom in a cavity, spontaneous emission can

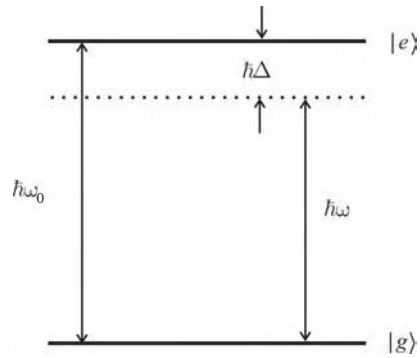


Fig. 4.3. Energy level diagram for a two-level atom acting with a near resonant classical driving field of frequency ω. The resonant frequency between the two atomic levels is ω_0 and the detuning $\Delta = \omega_0 - \omega$.

be reversible, that is, the atom can re-absorb the emitted photon. Such behavior is discussed in Section 4.5.

4.4 The Rabi model

The perturbation theory approach to atom–field interactions assumes that the initial atomic state population is essentially unchanged, that is, the probability amplitude for the atom being in any other state remains small. On the other hand, a strong laser field of frequency near resonance with a pair of atomic levels (assumed of opposite parity) will cause a large population transfer to the near-resonant state but not to any other. In such a case, perturbation theory must be abandoned. Only the two dominant states will be retained and the problem will be solved more "exactly". This is the Rabi model [7], so named because of its original setting in magnetic resonance as studied by Rabi long ago. We study the semiclassical case first.

For definiteness and to follow convention, we label our two atomic states $|g\rangle$ (for ground) and $|e\rangle$ (for excited). The energy difference between these states is characterized by the transition frequency $\omega_0 = (E_e - E_g)/\hbar$. This frequency is close to the frequency ω of the driving laser field as shown in Fig. 4.3. The interaction Hamiltonian we write as

$$\hat{H}^{(\mathrm{I})}(t) = \hat{V}_0 \cos\omega t \tag{4.69}$$

where $\hat{V}_0 = -\hat{\mathbf{d}} \cdot \mathbf{E}_0$. We write the state vector as

$$|\psi(t)\rangle = C_g(t)e^{-iE_g t/\hbar}|g\rangle + C_e(t)e^{-iE_e t/\hbar}|e\rangle. \tag{4.70}$$

From the Schrödinger equation

$$i\hbar\frac{\partial|\psi(t)\rangle}{\partial t} = \hat{H}(t)\,|\psi(t)\rangle, \tag{4.71}$$

where

$$\hat{H} = \hat{H}_0 + \hat{V}_0 \cos\omega t, \tag{4.72}$$

we arrive at the coupled set of equations for the amplitudes C_g and C_e:

$$\begin{aligned} \dot{C}_g &= -\frac{i}{\hbar}\mathcal{V}\cos\omega t e^{-i\omega_0 t}\, C_e \\ \dot{C}_e &= -\frac{i}{\hbar}\mathcal{V}\cos\omega t e^{i\omega_0 t}\, C_g \end{aligned} \tag{4.73}$$

where $\mathcal{V} = \langle e|\hat{V}_0|g\rangle = -\hat{\mathbf{d}}_{eg}\cdot\mathbf{E}_0$, which we have taken to be real. As an initial condition we take all the population to be in the ground state: $C_g(0) = 1$ and $C_e(0) = 0$. In Eqs. (4.73) we expand $\cos\omega t$ in exponentials and retain only those terms oscillating at the frequency $\omega_0 - \omega$ to obtain

$$\begin{aligned} \dot{C}_g &= -\frac{i}{2\hbar}\mathcal{V}\exp[i(\omega-\omega_0)t]C_e \\ \dot{C}_e &= -\frac{i}{2\hbar}\mathcal{V}\exp[-i(\omega-\omega_0)t]C_g. \end{aligned} \tag{4.74}$$

Dropping the terms oscillating at $\omega_0 + \omega$, of course, constitutes the RWA. Eliminating C_g, we have for C_e:

$$\ddot{C}_e + i(\omega-\omega_0)\dot{C}_e + \frac{1}{4}\,\frac{\mathcal{V}^2}{\hbar^2}C_e = 0. \tag{4.75}$$

As a trial solution we set

$$C_e(t) = e^{i\lambda t} \tag{4.76}$$

which leads to the two roots

$$\lambda_\pm = \frac{1}{2}\{\Delta \pm [\Delta^2 + \mathcal{V}^2/\hbar^2]^{1/2}\} \tag{4.77}$$

where $\Delta = \omega_0 - \omega$ is the detuning of the atomic transition frequency and the laser field. Thus the general solution is of the form

$$C_e(t) = A_+ e^{i\lambda_+ t} + A_- e^{i\lambda_- t}, \tag{4.78}$$

where from the initial conditions we must have

$$A_\pm = \mp\frac{1}{2\hbar}\,\mathcal{V}[\Delta^2 + \mathcal{V}^2/\hbar^2]^{-1/2}. \tag{4.79}$$

Finally, then, our solution is

$$\begin{aligned} C_e(t) &= i\frac{\mathcal{V}}{\Omega_R\hbar}\, e^{i\Delta t/2}\sin(\Omega_R t/2) \\ C_g(t) &= e^{i\Delta t/2}\left\{\cos(\Omega_R t/2) - i\frac{\Delta}{\Omega_R}\sin(\Omega_R t/2)\right\}, \end{aligned} \tag{4.80}$$

where

$$\Omega_R = [\Delta^2 + \mathcal{V}^2/\hbar^2]^{1/2} \tag{4.81}$$

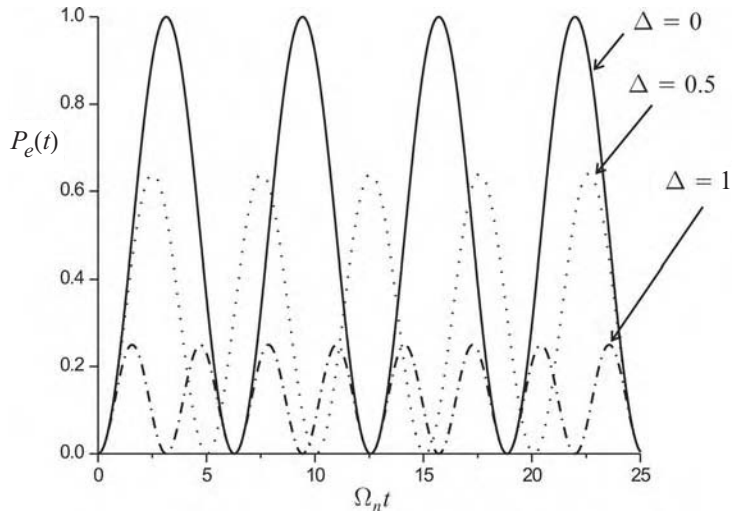


Fig. 4.4. Plots of $P_e(t)$ versus t for various detunings Δ.

is the so-called Rabi frequency. The probability that the atom is in state $|e\rangle$ is

$$\begin{aligned} P_e(t) &= |C_e(t)|^2 \\ &= \frac{\mathcal{V}^2}{\Omega_{\mathrm{R}}^2 \hbar^2} \sin^2(\Omega_{\mathrm{R}} t/2), \end{aligned} \tag{4.82}$$

which is plotted in Fig. 4.4 for various values of Δ. For the case of exact resonance, $\Delta = 0$, we have

$$P_e(t) = \sin^2\left(\frac{\mathcal{V}t}{2\hbar}\right), \tag{4.83}$$

and at the time $t = \pi\hbar/\mathcal{V}$ all the atomic population has been transferred to the excited state.

It is frequently convenient to consider the quantity known as the atomic inversion $W(t)$ defined as the difference in the excited- and ground-state populations:

$$W(t) = P_e(t) - P_g(t), \tag{4.84}$$

which, for the resonant case and with the atom initially in the ground state, is

$$\begin{aligned} W(t) &= \sin^2\left(\frac{\mathcal{V}t}{2\hbar}\right) - \cos^2\left(\frac{\mathcal{V}t}{2\hbar}\right) \\ &= -\cos(\mathcal{V}t/\hbar). \end{aligned} \tag{4.85}$$

Note that for $\Delta = 0$ the Rabi frequency is just $\Omega_{\mathrm{R}} = \mathcal{V}/\hbar$, the oscillation frequency of the atomic inversion. Again, for $t = \pi\hbar/\mathcal{V}$ all the population is transferred to the excited state: $W(\pi\hbar/\mathcal{V}) = 1$. In the parlance of NMR experiments

[8], such transfers are called π-pulses. On the other hand, if $t = \pi\hbar/2\mathcal{V}$ then $W(\pi\hbar/2\mathcal{V}) = 0$ and the population is shared coherently between the excited and ground states with

$$\begin{aligned} C_e(\pi\hbar/2\mathcal{V}) &= \frac{i}{\sqrt{2}}, \\ C_g(\pi\hbar/2\mathcal{V}) &= \frac{1}{\sqrt{2}}, \end{aligned} \tag{4.86}$$

so that

$$|\psi(\pi\hbar/2\mathcal{V})\rangle = \frac{1}{\sqrt{2}}(|g\rangle + i|e\rangle). \tag{4.87}$$

For obvious reasons, the transfer of population from the ground state to that state of Eq. (4.87) is called a $\pi/2$-pulse. Such transfers of populations by π- or $\pi/2$-pulses are standard procedures for manipulating not only spin states in NMR experiments [2] but have become routine for manipulating atomic or ionic states in laser spectroscopy experiments [9].

The results of the pertubation theory may be recovered from the Rabi model either when the size of $\mathcal{V}/2\hbar$ is so small compared with the detuning Δ that it can be neglected in Ω_R or if the radiation field acts only for such a short time that the term $\sin^2(\mathcal{V}t/2\hbar)$ can legitimately be represented by the first term in its expansion. In both cases the depletion of the initial atomic population is small and the perturbation-theory approach remains valid.

4.5 Fully quantum-mechanical model; the Jaynes–Cummings model

We now turn to the quantum electrodynamic version of the Rabi model. In our previous pertubation discussion of an atom interacting with a quantized electromagnetic field, we assumed the field to be a single-mode free field (plane wave). As we just discussed above, a free atom interacts with an infinite number of modes and thus the dynamics is not well described assuming only a single-mode field. On the other hand, it has recently become possible to manufacture environments where the density of modes is significantly different than in free space. We have in mind here small microwave cavities, or in some cases, optical cavities, capable of supporting only a single mode or maybe a few widely spaced (in frequency) modes. Thus in some cases, the ideal single-mode interaction can be realized in the laboratory. We shall discuss some specific examples in Chapter 10 but for now we consider an atom, with levels $|g\rangle$ and $|e\rangle$ as before, interacting with a single-mode cavity field of the form

$$\hat{\mathbf{E}} = \mathbf{e}\left(\frac{\hbar\omega}{\varepsilon_0 V}\right)^{1/2}(\hat{a} + \hat{a}^\dagger)\sin(kz) \tag{4.88}$$

where $\mathbf{e}$ is an arbitrarily oriented polarization vector.

The interaction Hamiltonian is now

$$\begin{aligned}\hat{H}^{(\mathrm{I})} &= -\hat{\mathbf{d}}\cdot\hat{\mathbf{E}} \\ &= \hat{d}g(\hat{a}+\hat{a}^{\dagger})\end{aligned} \tag{4.89}$$

where

$$g = -\left(\frac{\hbar\omega}{\varepsilon_0 V}\right)^{1/2}\sin(kz) \tag{4.90}$$

and where $\hat{d} = \hat{\mathbf{d}}\cdot\mathbf{e}$.

At this point it is convenient to introduce the so-called atomic transition operators

$$\hat{\sigma}_+ = |e\rangle\langle g|, \qquad \hat{\sigma}_- = |g\rangle\langle e| = \hat{\sigma}_+^{\dagger}, \tag{4.91}$$

and the inversion operator

$$\hat{\sigma}_3 = |e\rangle\langle e| - |g\rangle\langle g|. \tag{4.92}$$

These operators obey the Pauli spin algebra

$$\begin{aligned}[\hat{\sigma}_+, \hat{\sigma}_-] &= \hat{\sigma}_3 \\ [\hat{\sigma}_3, \hat{\sigma}_\pm] &= 2\hat{\sigma}_\pm.\end{aligned} \tag{4.93}$$

Only the off-diagonal elements of the dipole operator are nonzero, since by parity consideration $\langle e|\hat{d}|e\rangle = 0 = \langle g|\hat{d}|g\rangle$, so that we may write

$$\begin{aligned}\hat{d} &= d|g\rangle\langle e| + d^*|e\rangle\langle g| \\ &= d\hat{\sigma}_- + d^*\hat{\sigma}_+ = d(\hat{\sigma}_+ + \hat{\sigma}_-)\end{aligned} \tag{4.94}$$

where we have set $\langle e|\hat{d}|g\rangle = d$ and have assumed, without loss of generality, that d is real. Thus the interaction Hamiltonian is

$$\hat{H}^{(\mathrm{I})} = \hbar\lambda(\hat{\sigma}_+ + \hat{\sigma}_-)(\hat{a}+\hat{a}^{\dagger}) \tag{4.95}$$

where $\lambda = dg/\hbar$.

If we define the level of the energy to be zero halfway between the states $|g\rangle$ and $|e\rangle$ as in Fig. 4.5, then the free atomic Hamiltonian may be written as

$$\hat{H}_{\mathrm{A}} = \frac{1}{2}(E_e - E_g)\hat{\sigma}_3 = \frac{1}{2}\hbar\omega_0\hat{\sigma}_3, \tag{4.96}$$

where $E_e = -E_g = \frac{1}{2}\hbar\omega_0$. The free-field Hamiltonian is, after dropping the zero-point energy term,

$$\hat{H}_{\mathrm{F}} = \hbar\omega\hat{a}^{\dagger}\hat{a}. \tag{4.97}$$

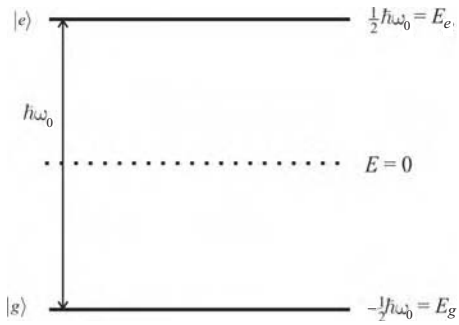


Fig. 4.5. Atomic energy level diagram where the $E = 0$ level is taken halfway between the two levels.

Thus the total Hamiltonian is

$$\begin{aligned}\hat{H} &= \hat{H}_{\mathrm{A}} + \hat{H}_{\mathrm{F}} + \hat{H}^{(\mathrm{I})} \\ &= \frac{1}{2}\hbar\omega_0\hat{\sigma}_3 + \hbar\omega\hat{a}^\dagger\hat{a} + \hbar\lambda(\hat{\sigma}_+ + \hat{\sigma}_-)(\hat{a} + \hat{a}^\dagger).\end{aligned} \tag{4.98}$$

In the free-field case, as we have already shown, the operators $\hat{a}$ and $\hat{a}^+$ evolve as

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t}, \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}. \tag{4.99}$$

One can show similarly that for the free-atomic case

$$\hat{\sigma}_\pm(t) = \hat{\sigma}_\pm(0)e^{\pm i\omega_0 t}. \tag{4.100}$$

Thus we can see that the approximate time dependences of the operator products in Eq. (4.98) are as follows:

$$\begin{aligned}\hat{\sigma}_+\hat{a} &\sim e^{i(\omega_0-\omega)t} \\ \hat{\sigma}_-\hat{a}^\dagger &\sim e^{-i(\omega_0-\omega)t} \\ \hat{\sigma}_+\hat{a}^\dagger &\sim e^{i(\omega+\omega_0)t} \\ \hat{\sigma}_-\hat{a} &\sim e^{-i(\omega+\omega_0)t}.\end{aligned} \tag{4.101}$$

For $\omega_0 \approx \omega$ the last two terms vary much more rapidly than the first two. Furthermore, the last two terms do not conserve energy in contrast to the first two. The term $\hat{\sigma}_+\hat{a}^\dagger$ corresponds to the emission of a photon as the atom goes from the ground to the excited state, whereas $\hat{\sigma}_-\hat{a}$ corresponds to the absorption of a photon as the atom goes from the excited to the ground state. Integrating the time-dependent Schrödinger equation, as in the perturbative case, will lead, for the last two terms, to denominators containing $\omega_0 + \omega$ as compared with $\omega_0 - \omega$ for the first two terms. The reader will not be surprised to learn that we are going to drop the non-energy conserving terms, making the RWA again, so that our Hamiltonian in this approximation is

$$\hat{H} = \frac{1}{2}\hbar\omega_0\hat{\sigma}_3 + \hbar\omega\hat{a}^\dagger\hat{a} + \hbar\lambda(\hat{\sigma}_+\hat{a} + \hat{\sigma}_-\hat{a}^\dagger). \tag{4.102}$$

The interaction described by this Hamiltonian is widely referred to as the Jaynes–Cummings model [10].

Before solving for the dynamics for any specific cases we take note of certain constants of the motion. An obvious one is the electron "number"

$$\hat{P}_{\mathrm{E}} = |e\rangle\langle e| + |g\rangle\langle g| = 1, \quad [\hat{H}, \hat{P}_{\mathrm{E}}] = 0 \tag{4.103}$$

valid when no other atomic states can become populated. Another is the excitation number

$$\hat{N}_{\mathrm{e}} = \hat{a}^{\dagger}\hat{a} + |e\rangle\langle e|, \quad [\hat{H}, \hat{N}_{\mathrm{e}}] = 0. \tag{4.104}$$

Using these constants of the motion we may break the Hamiltonian Eq. (4.102) into two commuting parts:

$$\hat{H} = \hat{H}_{\mathrm{I}} + \hat{H}_{\mathrm{II}} \tag{4.105}$$

where

$$\begin{aligned} \hat{H}_{\mathrm{I}} &= \hbar\omega\hat{N}_{\mathrm{e}} + \hbar\left(\frac{\omega_0}{2} - \omega\right)\hat{P}_{\mathrm{E}}, \\ \hat{H}_{\mathrm{II}} &= -\hbar\Delta + \hbar\lambda(\hat{\sigma}_{+}\hat{a} + \hat{\sigma}_{-}\hat{a}^{\dagger}), \end{aligned} \tag{4.106}$$

such that $[\hat{H}_{\mathrm{I}}, \hat{H}_{\mathrm{II}}] = 0$. Clearly, all the essential dynamics is contained in $\hat{H}_{\mathrm{II}}$ whereas $\hat{H}_{\mathrm{I}}$ contributes only overall irrelevant phase factors.

Let us now consider a simple example, with $\Delta = 0$, where the atom is initially in the excited state $|e\rangle$ and the field is initially in the number state $|n\rangle$. The initial state of the atom–field system is then $|i\rangle = |e\rangle|n\rangle$ and is of energy $E_i = \frac{1}{2}\hbar\omega + n\hbar\omega$. State $|i\rangle$ is coupled to (and only to) the state $|f\rangle = |g\rangle|n+1\rangle$ with energy $E_f = -\frac{1}{2}\hbar\omega + (n+1)\hbar\omega$. Note that $E_i = E_f$. We write the state vector as

$$|\psi(t)\rangle = C_i(t)|i\rangle + C_f(t)|f\rangle \tag{4.107}$$

where $C_i(0) = 1$ and $C_f(0) = 0$. Following standard procedures we obtain, from the interaction picture Schrödinger equation $i\hbar d|\psi(t)\rangle/dt = \hat{H}_{\mathrm{II}}|\psi(t)\rangle$ the equations for the coefficients

$$\begin{aligned} \dot{C}_i &= -i\lambda\sqrt{n+1}\,C_f, \\ \dot{C}_f &= -i\lambda\sqrt{n+1}\,C_i. \end{aligned} \tag{4.108}$$

Eliminating C_f we obtain

$$\ddot{C}_i + \lambda^2(n+1)\,C_i = 0. \tag{4.109}$$

The solution matching the initial conditions is

$$C_i(t) = \cos(\lambda t\sqrt{n+1}). \tag{4.110}$$

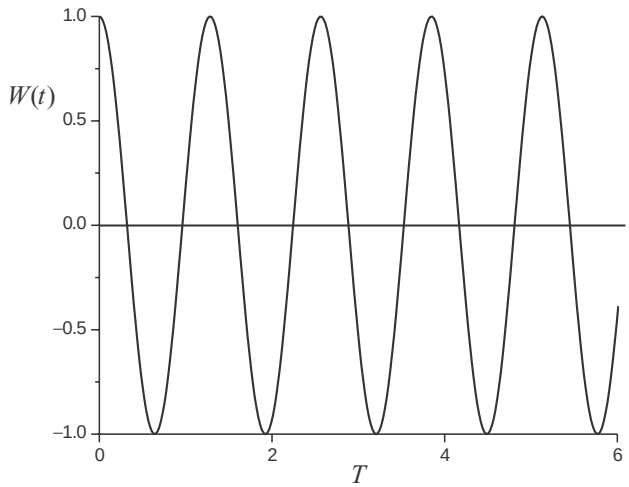


Fig. 4.6. Periodic atomic inversion with the field initially in a number state $|n\rangle$ with $n = 5$ photons.

From Eq. (4.108) we obtain

$$C_f(t) = -i \sin(\lambda t\sqrt{n+1}). \tag{4.111}$$

Thus our solution is

$$\begin{aligned} |\psi(t)\rangle &= \cos(\lambda t\sqrt{n+1})\,|e\rangle|n\rangle \\ &\quad -i \sin(\lambda t\sqrt{n+1})\,|g\rangle|n+1\rangle. \end{aligned} \tag{4.112}$$

The probability that the system remains in the initial state is

$$P_i(t) = |C_i(t)|^2 = \cos^2(\lambda t\sqrt{n+1}) \tag{4.113}$$

while the probability that it makes a transition to the state $|f\rangle$ is

$$P_f(t) = |C_f(t)|^2 = \sin^2(\lambda t\sqrt{n+1}). \tag{4.114}$$

The atomic inversion is given by

$$\begin{aligned} W(t) &= \langle\psi(t)|\hat{\sigma}_3|\psi(t)\rangle \\ &= P_i(t) - P_f(t) \\ &= \cos(2\lambda t\sqrt{n+1}). \end{aligned} \tag{4.115}$$

We may define a quantum electrodynamic Rabi frequency $\Omega(n) = 2\lambda\sqrt{n+1}$ so that

$$W(t) = \cos[\Omega(n)t]. \tag{4.116}$$

Clearly, the atomic inversion for the field initially in a number state is strictly periodic (Fig. 4.6), just as in the semiclassical case of Eq. (4.85) (apart from the minus sign due only to a different initial atomic state) and except for the fact that in

the classical case there must always be a field present initially. But in the quantum-mechanical case there are Rabi oscillations even for the case when $n = 0$. These are the vacuum-field Rabi oscillations [11] and, of course, they have no classical counterpart. They are the result of the atom spontaneously emitting a photon then re-absorbing it, re-emitting it, etc.: an example of reversible spontaneous emission. Such effects can be observed if atoms interact with fields in very high Q cavities. But aside from this, overall, the behavior of the atomic dynamics for a definite number of photons is very much like the semiclassical Rabi model, i.e. it is periodic and regular. Perhaps this is a bit counterintuitive since a number state is the most nonclassical of all the field states. Intuition might suggest that when the field is initially in a coherent state, we should recover the semiclassical, periodic and regular, Rabi oscillations. As we are about to demonstrate, intuition, in this case, fails.

Let us now consider a more general (pure state) solution of the dynamics. We assume the atom is initially in a superposition of states $|e\rangle$ and $|g\rangle$:

$$|\psi(0)\rangle_{\text{atom}} = C_g|g\rangle + C_e|e\rangle, \tag{4.117}$$

and the field is initially in the state

$$|\psi(0)\rangle_{\text{field}} = \sum_{n=0}^{\infty} C_n|n\rangle, \tag{4.118}$$

such that the initial atom–field state is

$$|\psi(0)\rangle = |\psi(0)\rangle_{\text{atom}} \otimes |\psi(0)\rangle_{\text{field}}. \tag{4.119}$$

The solution of the Schrödinger equation is now

$$\begin{aligned} |\psi(t)\rangle = \sum_{n=0}^{\infty} \{&[C_e C_n \cos(\lambda t\sqrt{n+1}) - iC_g C_{n+1}\sin(\lambda t\sqrt{n+1})]\,|e\rangle \\ &+ [-iC_e C_{n-1}\sin(\lambda t\sqrt{n}) + C_g C_n \cos(\lambda t\sqrt{n})]|g\rangle\}\,|n\rangle. \end{aligned} \tag{4.120}$$

In general, this is an entangled state.

For the case of the atom initially in the excited state, where $C_e = 1$ and $C_g = 0$, we may write the solution as

$$|\psi(t)\rangle = |\psi_g(t)\rangle|g\rangle + |\psi_e(t)\rangle|e\rangle \tag{4.121}$$

where $|\psi_g(t)\rangle$ and $|\psi_e(t)\rangle$ are the field components of $|\psi(t)\rangle$ given by

$$\begin{aligned} |\psi_g(t)\rangle &= -i\sum_{n=0}^{\infty} C_n \sin(\lambda t\sqrt{n+1})|n+1\rangle, \\ |\psi_e(t)\rangle &= \sum_{n=0}^{\infty} C_n \cos(\lambda t\sqrt{n+1})|n\rangle. \end{aligned} \tag{4.122}$$

The atomic inversion is

$$\begin{aligned} W(t) &= \langle\psi(t)|\hat{\sigma}_3|\psi(t)\rangle \\ &= \langle\psi_e(t)|\psi_e(t)\rangle - \langle\psi_g(t)|\psi_g(t)\rangle \\ &= \sum_{n=0}^{\infty} |C_n|^2 \cos(2\lambda t\sqrt{n+1}). \end{aligned} \tag{4.123}$$

The result is just the sum of n-photon inversions of Eq. (4.115) weighted with the photon number distribution of the initial field state.

For the coherent state, again that most classical of all quantum states, we have

$$C_n = e^{-|\alpha|^2/2}\frac{\alpha^n}{\sqrt{n!}} \tag{4.124}$$

and the inversion is

$$W(t) = e^{-\bar{n}}\sum_{n=0}^{\infty}\frac{\bar{n}^n}{n!}\cos(2\lambda t\sqrt{n+1}). \tag{4.125}$$

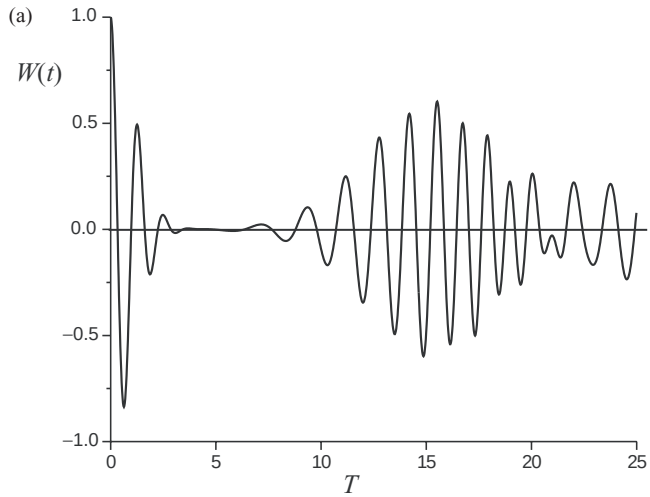
A plot of $W(t)$ versus the scaled time $T = \lambda\, t$ in Fig. 4.7 reveals significant discrepancies between the fully quantized and semiclassical Rabi oscillations. We note first that the Rabi oscillations initially appear to damp out, or collapse. The collapse of the Rabi oscillations was noted fairly early in the study of this "idealized" model interaction [12]. Several years later, perhaps by executing longer runs of a computer program, it was found that after a period of quiescence following the collapse, the Rabi oscillations start to revive [13], although not completely. At longer times one finds a sequence of collapses and revivals, the revivals becoming less distinct as time increases. This collapse and revival behavior of the Rabi oscillations in the fully quantized model is strikingly different than in the semiclassical case where the oscillations have constant amplitude. We must now explain this difference.

First we consider the collapse. The average photon number is $\bar{n} = |\alpha|^2$ so the dominant Rabi frequency is

$$\Omega(\bar{n}) = 2\lambda\sqrt{\bar{n}+1} \approx 2\lambda\sqrt{\bar{n}}, \quad \bar{n} \gg 1. \tag{4.126}$$

But there will be a range of "important" frequencies as a result of the spread of the probabilities $|C_n|^2$ about $\bar{n}$ for photon numbers in the range $\bar{n} \pm \Delta n$; i.e. the frequencies in the range $\Omega(\bar{n} - \Delta n)$ to $\Omega(\bar{n} + \Delta n)$. The collapse time t_c may be estimated from the time-frequency "uncertainty" relation

$$t_c[\Omega(\bar{n} + \Delta n) - \Omega(\bar{n} - \Delta n)] \simeq 1 \tag{4.127}$$



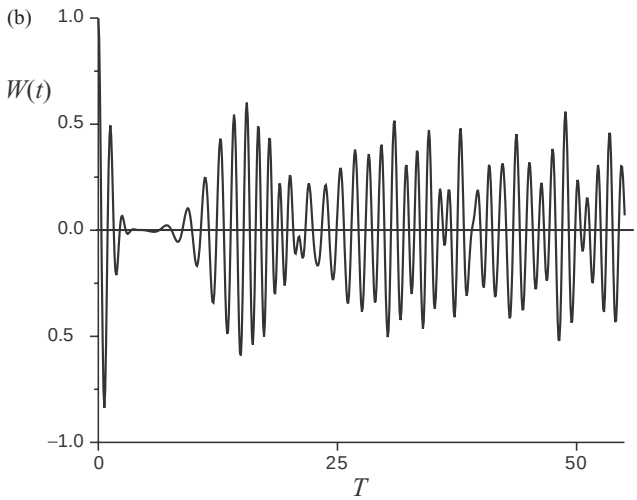


Fig. 4.7. (a) Atomic inversion with the field initially in a coherent state $\bar{n} = 5$. (b) Same as (a) but showing the evolution for a longer time, beyond the first revival. Here, T is the scaled time λt.

where the spread of frequencies is responsible for the "dephasing" of the Rabi oscillations. For the coherent state, $\Delta n = \bar{n}^{1/2}$, and with

$$\begin{aligned}\Omega(\bar{n} \pm \bar{n}^{1/2}) &\simeq 2\lambda[\bar{n} \pm \bar{n}^{1/2}]^{1/2} \\ &= 2\lambda\bar{n}^{1/2}\left[1 \pm \frac{1}{\bar{n}^{1/2}}\right]^{1/2} \\ &\simeq 2\lambda\bar{n}^{1/2}\left(1 \pm \frac{1}{2\bar{n}^{1/2}}\right) \\ &= 2\lambda\bar{n}^{1/2} \pm \lambda \end{aligned} \tag{4.128}$$

it follows that

$$t_c[\Omega(\bar{n}+\bar{n}^{1/2})-\Omega(\bar{n}-\bar{n}^{1/2})] \simeq t_c 2\lambda \simeq 1 \tag{4.129}$$

and thus $t_c \simeq (2\lambda)^{-1}$, which is independent of $\bar{n}$.

The preceding "derivation" of the collapse time is not very rigorous. We shall now give a more rigorous derivation. We expand $[n+1]^{1/2}$ about $\bar{n}$ as

$$[n+1]^{1/2} = [\bar{n}+1]^{1/2} + \frac{1}{2(\bar{n}+1)^{1/2}}(n-\bar{n}) + \cdots \tag{4.130}$$

so that we may approximate the inversion as

$$W(t) \simeq \frac{1}{2}e^{-\bar{n}} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{n!} \Big[e^{2i\lambda t(\bar{n}+1)^{1/2}} e^{i\lambda n t/(\bar{n}+1)^{1/2}} e^{-i\lambda t\bar{n}/(\bar{n}+1)^{1/2}} + e^{-2i\lambda t(\bar{n}+1)^{1/2}} e^{-i\lambda n t/(\bar{n}+1)^{1/2}} e^{i\lambda t\bar{n}/(\bar{n}+1)^{1/2}} \Big]. \tag{4.131}$$

Note that

$$\sum_{n=0}^{\infty} \frac{\bar{n}^n}{n!} e^{in\lambda t/(\bar{n}+1)^{1/2}} = \exp\left[\bar{n}\, e^{i\lambda t/(\bar{n}+1)^{1/2}}\right]. \tag{4.132}$$

For short times t

$$e^{i\lambda t/(\bar{n}+1)^{1/2}} \simeq 1 + i\lambda t/(\bar{n}+1)^{1/2} - \frac{\lambda^2 t^2}{2(\bar{n}+1)} \tag{4.133}$$

and thus

$$e^{\bar{n}e^{i\lambda t/(\bar{n}+1)^{1/2}}} \simeq e^{\bar{n}} e^{i\lambda t\bar{n}/(\bar{n}+1)} e^{-\frac{1}{2}\frac{\lambda^2 t^2}{(\bar{n}+1)}\bar{n}}. \tag{4.134}$$

Putting all of this together we arrive at

$$W(t) \simeq \cos\{2\lambda t(\bar{n}+1)^{1/2}\} \exp\left\{-\frac{1}{2}\frac{\lambda^2 t^2 \bar{n}}{\bar{n}+1}\right\}, \tag{4.135}$$

valid for a short time. The inversion evidently exhibits a Gaussian decay law with a decay time given by

$$t_c = \frac{\sqrt{2}}{\lambda}\sqrt{\frac{\bar{n}+1}{\bar{n}}} \simeq \frac{\sqrt{2}}{\lambda}, \qquad \bar{n} \gg 1 \tag{4.136}$$

which, apart from a numerical constant of the order of unity, agrees with our previous estimate.

Now let us examine the phenomenon of the revivals. $W(t)$ obviously consists of a sum of oscillating terms, each term oscillating at a particular Rabi frequency $\Omega(n) = 2\lambda\sqrt{n+1}$. If two neighboring terms are oscillating 180° out of phase with each other we expect at least an approximate cancellation of these terms. On the other hand, if the neighboring terms are in phase with each other we expect a constructive interference. In fact, this should be so whenever neighboring phases

differ by some multiple of 2π. Since only those important frequencies around $\bar{n}$ will contribute, revivals should occur for times $t = t_R$ such that

$$[\Omega(\bar{n}+1) - \Omega(\bar{n})]\, t_R = 2\pi k, \quad k = 0, 1, 2, \ldots \tag{4.137}$$

holds. Expanding $\Omega(\bar{n})$ and $\Omega(\bar{n}+1)$ we easily arrive at $t_R = (2\pi/\lambda)\bar{n}^{1/2}k$, $(\bar{n} \gg 1)$. More rigorously, using Eq. (4.118) we have

$$W(t) \simeq \cos\left[2\lambda t(\bar{n}+1)^{1/2} + \lambda t\bar{n}/(\bar{n}+1)^{-1/2} - \bar{n}\sin\left(\frac{\lambda t}{(\bar{n}+1)^{1/2}}\right)\right]$$
$$\times \exp\left\{-\bar{n}\left[1 - \cos\left(\frac{\lambda t}{(\bar{n}+1)^{1/2}}\right)\right]\right\}. \tag{4.138}$$

Obviously, the amplitude will be a maximum wherever the time $t = t_R = (2\pi/\lambda)(\bar{n}+1)^{1/2}k \approx (2\pi/\lambda)^{1/2}\bar{n}^{1/2}k$, $(\bar{n} \gg 1)$, in agreement with the previous analysis. In Chapter 10 we discuss two experiments, one in the context of cavity quantum electrodynamics and the other in the context of the center-of-mass motion of a trapped ion, where the predicted collapse and revival of the Rabi oscillations have been observed.

4.6 The dressed states

There are many ways to solve for the dynamics of the Jaynes–Cummings model (see the reviews [14]). In Section 4.5 we have solved the time-dependent Schrödinger equation first for a field containing n photons and then, by simple extrapolation, for the case of a field in a superposition of the number states. Another important way to obtain the dynamics is first to find the stationary states of the Jaynes–Cummings Hamiltonian. For reasons that should become clear shortly, these eigenstates are called the "dressed" states [15].

Consider once again the Jaynes–Cummings model Hamiltonian

$$\hat{H} = \frac{1}{2}\hbar\omega_0\hat{\sigma}_3 + \hbar\omega\hat{a}^\dagger\hat{a} + \hbar\lambda(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-) \tag{4.139}$$

where we have *not* assumed the resonance condition $\omega = \omega_0$ at this point. In terms of the field number states, the interaction term in $\hat{H}$ causes only transitions of the type

$$|e\rangle|n\rangle \leftrightarrow |g\rangle|n+1\rangle \tag{4.140}$$

or

$$|e\rangle|n-1\rangle \leftrightarrow |g\rangle|n\rangle. \tag{4.141}$$

The product states $|e\rangle|n-1\rangle$, $|g\rangle|n\rangle$, etc., are sometimes referred to as the "bare" states of the Jaynes–Cummings model; they are product states of the unperturbed atom and field. For a fixed n, the dynamics is completely confined to the

two-dimensional space of product states, either ($|e\rangle|n-1\rangle$, $|g\rangle|n\rangle$) or ($|e\rangle|n\rangle$, $|g\rangle|n-1\rangle$). We define the following product states for a given n:

$$\begin{aligned} |\psi_{1n}\rangle &= |e\rangle|n\rangle \\ |\psi_{2n}\rangle &= |g\rangle|n+1\rangle. \end{aligned} \tag{4.142}$$

Obviously $\langle\psi_{1n} \mid \psi_{2n}\rangle = 0$. Using this basis we obtain the matrix elements of $\hat{H}$, $H_{ij}^{(n)} = \langle\psi_{in}|\hat{H}|\psi_{jn}\rangle$, which are

$$\begin{aligned} H_{11}^{(n)} &= \hbar\left[n\omega + \tfrac{1}{2}\omega_0\right], \\ H_{22}^{(n)} &= \hbar\left[(n+1)\omega - \tfrac{1}{2}\omega_0\right], \\ H_{12}^{(n)} &= \hbar\lambda\sqrt{n+1} = H_{21}^{(n)}. \end{aligned} \tag{4.143}$$

Thus in the 2×2 subspace of Eq. (4.142) we obtain the matrix representation of $\hat{H}$:

$$\boldsymbol{H}^{(n)} = \begin{bmatrix} n\omega + \frac{1}{2}\hbar\omega_0 & \hbar\lambda\sqrt{n+1} \\ \hbar\lambda\sqrt{n+1} & (n+1)\omega - \frac{1}{2}\omega_0 \end{bmatrix}. \tag{4.144}$$

This matrix is "self-contained" since, as we have said, the dynamics connects only those states for which the photon number changes by ± 1. For a given n, the energy eigenvalues of $\boldsymbol{H}^{(n)}$ are as follows:

$$E_{\pm}(n) = \left(n + \frac{1}{2}\right)\hbar\omega \pm \hbar\Omega_n(\Delta) \tag{4.145}$$

where

$$\Omega_n(\Delta) = [\Delta^2 + 4\lambda^2(n+1)]^{1/2} \quad (\Delta = \omega_0 - \omega) \tag{4.146}$$

is the Rabi frequency which now includes the effects of the detuning Δ. Obviously, for $\Delta = 0$ we obtain $\Omega_n(0) = 2\lambda\sqrt{n+1}$, the same quantum electrodynamic Rabi frequencies seen earlier. The eigenstates $|n, \pm\rangle$ associated with the energy eigenvalues are given by

$$\begin{aligned} |n,+\rangle &= \cos(\Phi_n/2)|\psi_{1n}\rangle + \sin(\Phi_n/2)|\psi_{2n}\rangle \\ |n,-\rangle &= -\sin(\Phi_n/2)|\psi_{1n}\rangle + \cos(\Phi_n/2)|\psi_{2n}\rangle \end{aligned} \tag{4.147}$$

where the angle Φ_n is defined through

$$\Phi_n = \tan^{-1}\left(\frac{2\lambda\sqrt{n+1}}{\Delta}\right) = \tan^{-1}\left(\frac{\Omega_n(0)}{\Delta}\right) \tag{4.148}$$

and where

$$\begin{aligned} \sin(\Phi_n/2) &= \frac{1}{\sqrt{2}}\left[\frac{\Omega_n(\Delta) - \Delta}{\Omega_n(\Delta)}\right]^{1/2} \\ \cos(\Phi_n/2) &= \frac{1}{\sqrt{2}}\left[\frac{\Omega_n(\Delta) + \Delta}{\Omega_n(\Delta)}\right]^{1/2}. \end{aligned} \tag{4.149}$$

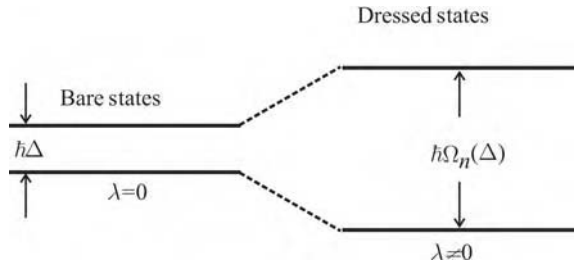


Fig. 4.8. Energy-level splitting due to the interaction of the atom with a quantized field. The split levels on the right are the energy levels of the dressed states.

The states $|n,\pm\rangle$ are often referred to as "dressed states" or as the Jaynes–Cummings doublet. The bare states $|\psi_{1n}\rangle$ and $|\psi_{2n}\rangle$, of energies $E_{1n} = \hbar(\omega_0/2 + \hbar\omega)$ and $E_{2n} = \hbar[-\omega_0/2 + (n+1)\omega]$ respectively, are each further split in energy owing to the interaction as indicated in Fig. 4.8. The splitting of the bare states into the dressed states is a kind of Stark shift, often called the AC, or dynamic, Stark shift. Note that in the limit of exact resonance, $\Delta = 0$, the bare states are degenerate but the splitting of the dressed states of course, remains. In this limit, the dressed states are related to the bare states according to

$$\begin{aligned} |n,+\rangle &= \frac{1}{\sqrt{2}}(|e\rangle|n\rangle + |g\rangle|n+1\rangle) \\ |n,-\rangle &= \frac{1}{\sqrt{2}}(-|e\rangle|n\rangle + |g\rangle|n+1\rangle). \end{aligned} \tag{4.150}$$

To see how the dressed states can be used to obtain the dynamics for rather general initial states, let us consider the specific case of a field prepared in some superposition of number states

$$|\psi_f(0)\rangle = \sum_n C_n|n\rangle \tag{4.151}$$

where an atom, prepared in state $|e\rangle$, is injected into the field. Thus the initial state of the atom–field system is

$$\begin{aligned} |\psi_{\rm af}(0)\rangle &= |\psi_f(0)\rangle|e\rangle \\ &= \sum_n C_n|n\rangle|e\rangle = \sum_n C_n|\psi_{1n}\rangle. \end{aligned} \tag{4.152}$$

From Eqs. (4.147), we obtain $|\psi_{1n}\rangle$ in terms of the dressed states $|n,\pm\rangle$ as

$$|\psi_{1n}\rangle = \cos(\Phi_n/2)|n,+\rangle - \sin(\Phi_n/2)|n,-\rangle \tag{4.153}$$

and thus

$$|\psi_{\rm af}(0)\rangle = \sum_n C_n[\cos(\Phi_n/2)|n,+\rangle - \sin(\Phi_n/2)|n,-\rangle]. \tag{4.154}$$

Since the dressed states $|n, \pm\rangle$ are stationary states of the atom–field system, then the state vector for times $t > 0$ is just given by

$$|\psi_{\rm af}(t)\rangle = \exp\left[-\frac{i}{\hbar}\hat{H}t\right]|\psi_{af}(0)\rangle$$
$$= \sum_n C_n\Big[\cos(\Phi_n/2)|n,+\rangle e^{-iE_+(n)t/\hbar} - \sin(\Phi_n/2)|n,-\rangle e^{-iE_-(n)t/\hbar}\Big]. \tag{4.155}$$

Of course, the entire result may now be recast back into the more familiar "bare" state basis by simply substituting $|n, \pm\rangle$ from Eqs. (4.147). In the limit $\Delta = 0$ we will recover the previous result of Eq. (4.120). The demonstration of this is left as an exercise.

4.7 Density-operator approach: application to thermal states

So far, we have considered only cases where the field and the atom are initially in pure states. In general though, one or both of the subsystems may initially be in mixed states which requires us to seek a solution in terms of the density operator. For example, the field may initially be in a thermal state described by the density operator of Eq. (2.144). In studying the case of a two-level atom interacting with a thermal state described by a density operator, we are afforded yet another way to solve for the dynamics of the Jaynes–Cummings model.

We shall work in the interaction picture and once again assume the resonance condition so that the dynamics is driven by

$$\hat{H}_{\rm I} = \hbar\lambda(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-). \tag{4.156}$$

If $\hat{\rho}(t)$ is the density operator of the atom–field system at time t, the evolution of the system is given by

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}_{\rm I}, \hat{\rho}] \tag{4.157}$$

whose solution may be written as

$$\hat{\rho}(t) = \hat{U}_{\rm I}(t)\hat{\rho}(0)\hat{U}_{\rm I}^\dagger(t) \tag{4.158}$$

where

$$\hat{U}_{\rm I}(t) = \exp[-i\hat{H}_{\rm I}t/\hbar]$$
$$= \exp[-i\lambda t(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-)]. \tag{4.159}$$

In the two-dimensional atomic subspace, the operators $\hat{\sigma}_\pm$ and $\hat{\sigma}_3$ have the matrix representations

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.160}$$

where we have used the convention

$$\sigma_j = \begin{pmatrix} \langle e|\hat{\sigma}_j|e\rangle & \langle e|\hat{\sigma}_j|g\rangle \\ \langle g|\hat{\sigma}_j|e\rangle & \langle g|\hat{\sigma}_j|g\rangle \end{pmatrix} \quad j = \pm,\ 3. \tag{4.161}$$

In this two-dimensional subspace the evolution operator $\hat{U}_{\rm I}(t)$ may be expanded as

$$\hat{\boldsymbol{U}}_{\rm I}(t) = \begin{pmatrix} \hat{C}(t) & \hat{S}'(t) \\ \hat{S}(t) & \hat{C}'(t) \end{pmatrix} \tag{4.162}$$

where

$$\hat{C}(t) = \cos(\lambda t\sqrt{\hat{a}\hat{a}^\dagger}) \tag{4.163}$$

$$\hat{S}(t) = -i\hat{a}^\dagger \frac{\sin(\lambda t\sqrt{\hat{a}\hat{a}^+})}{\sqrt{\hat{a}\hat{a}^\dagger}} \tag{4.164}$$

$$\hat{C}'(t) = \cos(\lambda t\sqrt{\hat{a}^\dagger\hat{a}}) \tag{4.165}$$

$$\hat{S}'(t) = -i\hat{a}\frac{\sin(\lambda t\sqrt{\hat{a}^\dagger\hat{a}})}{\sqrt{\hat{a}^\dagger\hat{a}}}. \tag{4.166}$$

(The operators $\hat{C}$, $\hat{S}$, etc., here are not to be confused with the cosine and sine operators of the phase introduced in Chapter 2.) The Hermitian adjoint of $\hat{U}_{\rm I}(t)$ of Eq. (4.162) is just

$$\hat{U}_{\rm I}^\dagger(t) = \hat{U}_{\rm I}(-t) = \begin{pmatrix} \hat{C}(t) & -\hat{S}'(t) \\ -\hat{S}(t) & \hat{C}'(t) \end{pmatrix}. \tag{4.167}$$

We now suppose that at $t = 0$ the density operator for the atom–field system factors into field and atomic parts:

$$\hat{\rho}(0) = \hat{\rho}^{\rm F}(0) \otimes \hat{\rho}^{\rm A}(0). \tag{4.168}$$

We further suppose (to work out a particular example) that the atom is initially in the excited state $|e\rangle$ such that $\hat{\rho}^{\rm A}(0) = |e\rangle\langle e|$. The corresponding density matrix for the atom (using the convention of Eq. (4.161)) is

$$\rho^{\rm A}(0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.169}$$

and thus we may write for the system

$$\hat{\rho}(0) = \begin{pmatrix} \hat{\rho}^{\rm F}(0) & 0 \\ 0 & 0 \end{pmatrix} = \hat{\rho}^{\rm F}(0) \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.170}$$

Using Eqs. (4.162), (4.167) and (4.169) in Eq. (4.158) we find that

$$\hat{\rho}(t) = \begin{pmatrix} \hat{C}(t)\hat{\rho}(0)\hat{C}(t) & -\hat{C}(t)\hat{\rho}^{\rm F}(0)\hat{S}'(t) \\ \hat{S}(t)\hat{\rho}^{\rm F}(0)\hat{C}(t) & -\hat{S}(t)\hat{\rho}^{\rm F}(0)\hat{S}'(t) \end{pmatrix}. \tag{4.171}$$

The reduced density operator of the field is found by tracing over the atomic states and thus

$$\hat{\rho}^{\mathrm{F}}(t) = \mathrm{Tr}_{\mathrm{A}}\hat{\rho}(t) = \hat{C}(t)\hat{\rho}^{\mathrm{F}}(0)\hat{C}(t) - \hat{S}(t)\hat{\rho}^{\mathrm{F}}(0)\hat{S}'(t). \tag{4.172}$$

The density matrix elements for the field are

$$\begin{aligned}\hat{\rho}^{\mathrm{F}}_{nm}(t) &\equiv \langle n|\hat{\rho}^{\mathrm{F}}(t)|m\rangle \\ &= \langle n|\hat{C}(t)\hat{\rho}^{\mathrm{F}}(0)\hat{C}(t)|m\rangle - \langle n|\hat{S}(t)\hat{\rho}^{\mathrm{F}}(0)\hat{S}'(t)|m\rangle.\end{aligned} \tag{4.173}$$

On the other hand, tracing over the field states we obtain the reduced density operator of the atom:

$$\hat{\rho}^{\mathrm{A}}(t) = \mathrm{Tr}_{\mathrm{F}}\hat{\rho}(t) = \sum_{n=0}^{\infty}\langle n|\hat{\rho}(t)|n\rangle. \tag{4.174}$$

The density operator matrix elements are given by

$$\langle i|\hat{\rho}^{\mathrm{A}}(t)|j\rangle = \sum_{n=0}^{\infty}\langle i,n|\hat{\rho}(t)|j,n\rangle = \rho^{\mathrm{A}}_{ij}(t) \tag{4.175}$$

where $i, j = e, g$. The diagonal elements $\rho^{\mathrm{A}}_{ee}(t)$ and $\rho^{\mathrm{A}}_{gg}(t)$ are the populations of the excited and ground states, respectively, and satisfy the condition

$$\rho^{\mathrm{A}}_{gg}(t) + \rho^{\mathrm{A}}_{ee}(t) \ = 1. \tag{4.176}$$

The atomic inversion is given by

$$W(t) = \rho^{\mathrm{A}}_{ee}(t) - \rho^{\mathrm{A}}_{gg}(t) \ = 2\rho^{\mathrm{A}}_{ee}(t) - 1. \tag{4.177}$$

From Eqs. (4.171) and (4.175) we find that

$$\begin{aligned}\rho^{\mathrm{A}}_{ee}(t) &\equiv \sum_{n=0}^{\infty}\langle n|\hat{C}(t)\hat{\rho}^{\mathrm{F}}(0)\hat{C}(t)|n\rangle \\ &= \sum_{n=0}^{\infty}\langle n|\,\hat{\rho}^{\mathrm{F}}(0)|n\rangle\cos^2(\lambda t\sqrt{n+1}).\end{aligned} \tag{4.178}$$

If the field is initially in a pure state

$$|\psi_{\mathrm{F}}\rangle = \sum_{n=0}^{\infty}C_n|n\rangle \tag{4.179}$$

then

$$\hat{\rho}^{\mathrm{F}}(0) = |\psi_{\mathrm{F}}\rangle\langle\psi_{\mathrm{F}}| \tag{4.180}$$

and thus

$$\rho^{\mathrm{A}}_{ee}(t) = \sum_{n=0}^{\infty}|C_n|^2\cos^2(\lambda t\sqrt{n+1}) \tag{4.181}$$

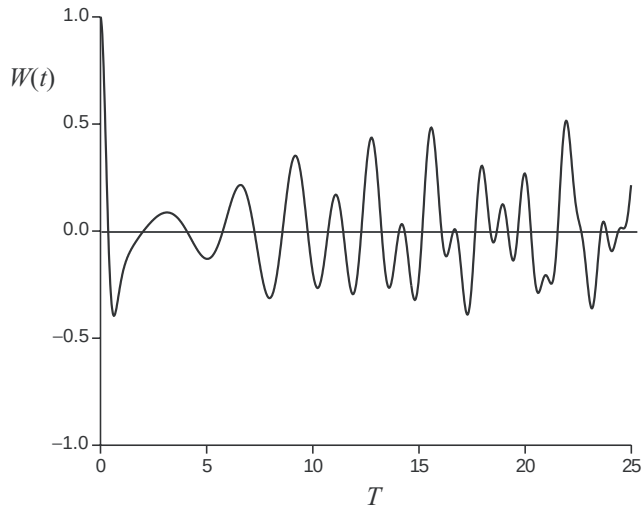


Fig. 4.9. The atomic inversion versus time for the atom initially in the excited state and the field initially in a thermal state with $\bar{n} = 2$. Again, $T = \lambda t$.

which, through Eq. (4.177), yields the atomic inversion found in Eq. (4.123). But suppose the field is initially in a thermal state (a mixed state) where

$$\hat{\rho}^{\mathrm{F}}(0) = \hat{\rho}_{\mathrm{Th}} = \sum_n P_n |n\rangle\langle n| \tag{4.182}$$

where P_n is given by Eq. (2.145). From Eq. (4.178) we ultimately obtain the atomic inversion for an atom resonantly interacting with a thermal field as [16]

$$W(t) = \sum_{n=0}^{\infty} P_n \cos(2\lambda t\sqrt{n+1}). \tag{4.183}$$

In Fig. 4.9 we plot $W(t)$ versus λt for a thermal field containing an average photon number of $\bar{n} = 2$. We leave to the reader an analysis of the observed behavior.

4.8 The Jaynes–Cummings model with large detuning: a dispersive interaction

In the foregoing we have mostly assumed that the detuning $\Delta = \omega_0 - \omega = 0$. An important variation on the original Jaynes–Cummings model is the situation in which the detuning is large enough such that direct atomic transitions do not occur but where nevertheless "dispersive" interactions between a single atom and a cavity field do occur [17]. This version of the Jaynes–Cummings model is important in a number of applications related to fundamental tests of quantum mechanics, some of which are discussed in Chapter 10.

As is shown in Appendix C, the effective atom–field interaction Hamiltonian in the case of large detuning is given by

$$\hat{H}_{\mathrm{eff}} = \hbar\chi[\hat{\sigma}_+\hat{\sigma}_- + \hat{a}^\dagger\hat{a}\,\hat{\sigma}_3], \tag{4.184}$$

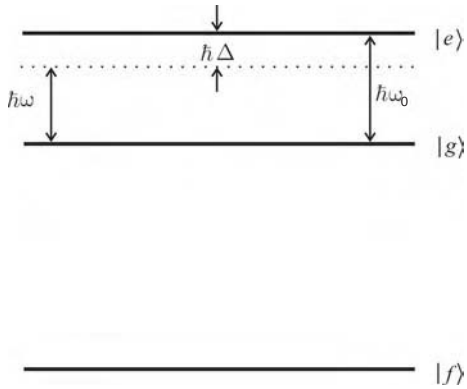


Fig. 4.10. The levels $|e\rangle$ and $|g\rangle$ are out of resonance with the field enough so that there are no direct transitions between them and only the dispersive interaction occurs. The state $|f\rangle$ is so far out of resonance with state $|g\rangle$ and the field that not even a dispersive interaction is present.

where $\chi = \lambda^2/\Delta$. Note that $\hat{\sigma}_+\hat{\sigma}_- = |e\rangle\langle e|$. Suppose the initial state of the atom–field system is $|\psi(0)\rangle = |g\rangle|n\rangle$, that is, the atom is in the ground state and the field is in a number state. Then, according to the interaction Hamiltonian of Eq. (4.184), the state at time $t > 0$ is

$$|\psi(t)\rangle = e^{-i\hat{H}_{\text{eff}}t/\hbar}|\psi(0)\rangle = e^{i\chi nt}|g\rangle|n\rangle \tag{4.185}$$

while for the initial state $|\psi(0)\rangle = |e\rangle|n\rangle$ we have

$$|\psi(t)\rangle = e^{-i\hat{H}_{\text{eff}}t/\hbar}|\psi(0)\rangle = e^{i\chi(n+1)t}|e\rangle|n\rangle. \tag{4.186}$$

Evidently, nothing very interesting happens, just the production of unmeasurable phase factors. On the other hand, for initial coherent states of the field we have, for $|\psi(0)\rangle = |g\rangle|\alpha\rangle$,

$$|\psi(t)\rangle = e^{-i\hat{H}_{\text{eff}}t/\hbar}|\psi(0)\rangle = |g\rangle|\alpha\, e^{i\chi t}\rangle, \tag{4.187}$$

and for $|\psi(0)\rangle = |e\rangle|\alpha\rangle$ we have

$$|\psi(t)\rangle = e^{-i\hat{H}_{\text{eff}}t/\hbar}|\psi(0)\rangle = e^{-i\chi t}|e\rangle|\alpha\, e^{-i\chi t}\rangle. \tag{4.188}$$

We notice that the coherent-state amplitude is rotated in phase space by the angle χt but that the direction of the rotation depends on the state of the atom. Suppose now that the atom is prepared in a superposition of the ground and excited states. For simplicity we take this to be a "balanced" state of the form $|\psi_{\text{atom}}\rangle = (|g\rangle + e^{i\phi}|e\rangle)/\sqrt{2}$, where ϕ is some phase. With the initial state $|\psi(0)\rangle = |\psi_{\text{atom}}\rangle|\alpha\rangle$ we have

$$|\psi(t)\rangle = e^{-i\hat{H}_{\text{eff}}t/\hbar}|\psi(0)\rangle = \frac{1}{\sqrt{2}}\Big(|g\rangle|\alpha\, e^{i\chi t}\rangle + e^{-i(\chi t-\phi)}|e\rangle|\alpha\, e^{-i\chi t}\rangle\Big). \tag{4.189}$$

This state is a bit more interesting, in fact, much more interesting as, in general, there is entanglement between the atom and the field. If we take $\chi t = \pi/2$ we then have

$$\left|\psi\left(\frac{\pi}{2\chi}\right)\right\rangle = \frac{1}{\sqrt{2}}(|g\rangle|i\,\alpha\,\rangle - i\, e^{i\phi}|e\rangle|-i\,\alpha\,\rangle). \tag{4.190}$$

Notice that, in terms of our phase-space pictures, the two coherent states in Eq. (4.190) are separated by 180°, the maximal separation. Coherent states differing in phase by 180° are also maximally *distinguishable* in the sense that there is essentially no overlap between the two states, at least if $|\alpha|$ is large enough. In fact, this is the case even with $|\alpha|$ as low as $\sqrt{2}$. With very large values of $|\alpha|$, the two coherent states are said to be *macroscopically* distinguishable and, for moderate values, *mesoscopically* distinguishable. The entangled state of Eq. (4.190) might bring to mind the tale of Schrödinger's ill-fated cat [18], suspended in a state of limbo, suspended in an entanglement between life and death and a non-decayed or decayed radioactive *microscopic* atom. Symbolically, the entangled state in Schrödinger's famous "paradox" is thus:

$$|\psi_{\text{atom}-\text{cat}}\rangle = \frac{1}{\sqrt{2}}[|\text{atom not decayed}\rangle|\text{cat alive}\rangle + |\text{atom decayed}\rangle|\text{cat dead}\rangle]. \quad (4.191)$$

The parallel between this state and that of Eq. (4.190) is obvious; the states of the two-level atom play the role of the radioactive atom and the two phase-separated coherent field states that of the cat.

Finally, there is another initial atomic state that is often considered. Suppose there is another atomic state, that we denote $|f\rangle$, of energy $E_f \ll E_g$ as pictured in Fig. 4.10, and of parity opposite that of $|g\rangle$. The cavity is assumed to support no mode resonant with the $f \leftrightarrow g$ transition and we further assume that the state $|f\rangle$ is so far out of resonance with the available cavity field mode that there is no discernable dispersive interaction either. Thus with the atom initially prepared in state $|f\rangle$ and with the field initially in the coherent state $|\alpha\rangle$, the initial product state $|f\rangle|\alpha\rangle$ remains as such; i.e. it does not evolve. Now suppose the atom is prepared in a superposition of the form $|\psi_{\text{atom}}\rangle = (|g\rangle + e^{i\phi}|f\rangle)/\sqrt{2}$. The initial-state atom–field state $|\psi(0)\rangle = |\psi_{\text{atom}}\rangle|\alpha\rangle$ is easily seen to evolve into

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}(|g\rangle|\alpha\, e^{i\chi t}\rangle + e^{i\phi}|f\rangle|\alpha\rangle). \quad (4.192)$$

In this case, only one component of the initial atomic superposition causes a phase shift in the coherent state, an advantage in certain applications. For $\chi t = \pi/2$ we have

$$\left|\psi\left(\frac{\pi}{2\chi}\right)\right\rangle = \frac{1}{\sqrt{2}}(|g\rangle|-\alpha\rangle + e^{i\phi}|f\rangle|\alpha\rangle), \quad (4.193)$$

another version of the Schrödinger-cat state.

In Chapters 7, 8, and 10 we shall further elaborate on issues related to Schrödinger's cat.

4.9 Extensions of the Jaynes–Cummings model

There are many possible extensions of the original Jaynes–Cummings model involving various types of alternative interactions. Among them are models

involving two-photon transitions, multimode and multilevel models, Raman coupled models, two-channel models, etc. We shall not discuss these models here but we do refer to the various review articles that have appeared [14, 19] and references therein. Some of these extensions will appear in the context of homework problems. Further, it turns out that Jaynes–Cummings types of interaction also occur in the context of the vibrational motion of an ion in an electromagnetic trap. The simplest example of these will be discussed in Chapter 10.

4.10 Schmidt decomposition and von Neumann entropy for the Jaynes–Cummings model

We finish this chapter with a discussion of the Schmidt decomposition and the related von Neumann entropy as they pertain to the Jaynes–Cummings model. As this system is bipartite, a Schmidt decomposition is assured, as discussed in Appendix A. We have already presented the solution of the time-dependent Schrödinger equation in Eq. (4.120), which we rewrite here as

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} [a_n(t)|g\rangle|n\rangle + b_n(t)|e\rangle|n\rangle], \tag{4.194}$$

where

$$\begin{aligned} a_n(t) &= C_g C_n \cos(\lambda t\sqrt{n}) - iC_e C_{n-1} \sin(\lambda t\sqrt{n}), \\ b_n(t) &= C_e C_n \cos(\lambda t\sqrt{n+1}) - iC_g C_{n+1} \sin(\lambda t\sqrt{n+1}). \end{aligned} \tag{4.195}$$

But according to the Schmidt decomposition, for any instant in time t we can always find bases $\{|u_i(t)\rangle\}$ for the atom and $\{|v_i(t)\rangle\}$ for the field such that the pure state of the system can be written as

$$|\Psi(t)\rangle = g_1(t)|u_1(t)\rangle|v_1(t)\rangle + g_2(t)|u_2(t)\rangle|v_2(t)\rangle. \tag{4.196}$$

The reduced density matrices of the atom and field in these bases are identical:

$$\begin{aligned} \rho_u(t) &= \begin{pmatrix} |g_1(t)|^2 & 0 \\ 0 & |g_2(t)|^2 \end{pmatrix} \\ \rho_v(t) &= \begin{pmatrix} |g_1(t)|^2 & 0 \\ 0 & |g_2(t)|^2 \end{pmatrix}. \end{aligned} \tag{4.197}$$

In order to find the coefficients $g_1(t)$ and $g_2(t)$ and the eigenvectors $|u_i\rangle$ and $|v_i\rangle$ we first calculate the reduced density operator of the atom in the bare basis specified by $|e\rangle$ and $|g\rangle$ and obtain

$$\rho_u = \begin{pmatrix} \sum_{n=0}^{\infty} |a_n(t)|^2 & \sum_{n=0}^{\infty} a_n(t) b_n^*(t) \\ \sum_{n=0}^{\infty} b_n(t) a_n^*(t) & \sum_{n=0}^{\infty} |b_n(t)|^2 \end{pmatrix}. \tag{4.198}$$

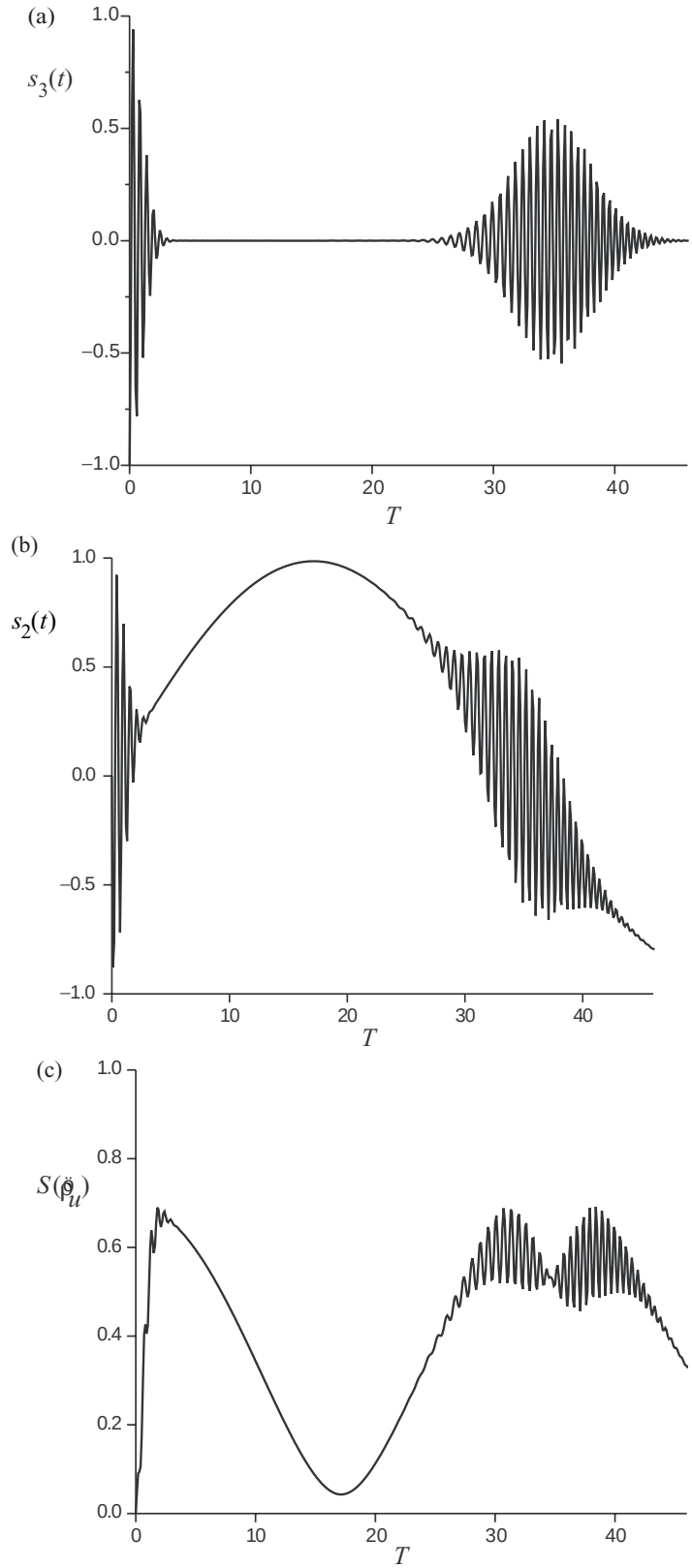


Fig. 4.11. Plots of s_3, s_2 and $S(\hat{\rho}_u)$ against the scaled time $\lambda\, t$. Again $T = \lambda t$.

Then we can use the parametrization of the density matrix in terms of the components of the Bloch vector as in Eq. (A25) to obtain

$$\begin{aligned} s_1(t) &= \sum_{n=0}^{\infty}\Big[a_n(t)b_n^*(t) + b_n(t)a_n^*(t)\Big], \\ s_2(t) &= -i\sum_{n=0}^{\infty}\Big[a_n(t)b_n^*(t) - b_n(t)a_n^*(t)\Big], \\ s_3(t) &= \sum_{n=0}^{\infty}[|a_n(t)|^2 - |b_n(t)|^2]. \end{aligned} \tag{4.199}$$

Then, as discussed in Appendix A, the coefficients $g_1(t)$ and $g_2(t)$ can be expressed in terms of the length of the Bloch vector according to

$$g_1(t) = \frac{1}{2}[1 + |\mathbf{s}(t)|], \quad g_2(t) = \frac{1}{2}[1 - |\mathbf{s}(t)|]. \tag{4.200}$$

Notice that the field mode, previously described in an infinite-dimensional Fock-state basis, is reduced to a "two-level" system in the Schmidt basis. The length of the Bloch vector is a measure of the purity of the atom–field system. For the atom and field in pure states the length of the Bloch vector is unity.

As discussed in Appendix A, the utility of the Schmidt decomposition is that it is easy to obtain an expression for von Neumann entropy, which for each of the subsystems of the Jaynes–Cummings model is

$$S(\hat{\rho}_u) = -g_1(t)\ln g_1(t) - g_2(t)\ln g_2(t) = S(\hat{\rho}_v). \tag{4.201}$$

Whenever $|\mathbf{s}| = 1$ we have $S(\hat{\rho}_u) = 0 = S(\hat{\rho}_v)$. In Fig. 4.11, for an atom initially in the excited state and for the field initially in a coherent state with $\alpha = \sqrt{30}$, we plot (a) the s_3 component of the Bloch vector, (b) the s_2 component, and (c) the von Neumann entropy $S(\hat{\rho}_u)$, all against the scaled time $T = \lambda t$. With our particular initial condition, $s_1 = 0$ for all times. We notice that s_3 undergoes the collapse and revival of the atomic inversion and that s_2 goes close to unity at one point during the quiescent period of the collapse of the former, indicating that the atom and field are nearly in pure states at that time. The von Neumann entropy, of course, is a minimum at that point. This result is perhaps a bit surprising and counter-intuitive. This behavior was first noticed, through a different type of calculation, by Gea-Banacloche [20] and was further examined by Phoenix and Knight [21].

Problems

1. Consider the semiclassical Rabi model (i.e. a two-level atom with a prescribed classical field) in the RWA as described by Eqs. (4.74). Obtain the solution assuming the atom initially in the excited state. Calculate the atomic inversion as a function of time.
2. Using the result of the previous problem, obtain the time-dependent expectation value of the atomic dipole moment operator $\hat{d} = d(\hat{\sigma}_+ + \hat{\sigma}_-)$ for the case of exact resonance. Compare the evolution of the dipole moment with the atomic inversion for the same resonance condition and comment on any similarities or differences.

3. In the fully quantized model of a two-level atom interacting with a quantized field within the RWA, the Jaynes–Cummings model, we have obtained the exact resonance solution for the initial state where the atom is excited and where the field is in a number state $|n\rangle$ (see Eqs. (4.107) to (4.116)). Recall that the Rabi oscillations of the atomic inversion are periodic for this case, just as in the semiclassical model. Use that solution to obtain the expectation value of the atomic dipole moment operator and compare with the result of the previous question. Is the time evolution of the atomic dipole moment in any way similar to that obtained in the semiclassical case? Comment on the result.
4. Obtain the expectation value of the atomic dipole moment as given by the Jaynes–Cummings model in the case where the field is initially in a coherent state. How does the result compare with the two previous cases? You should make a plot of the expectation value of the dipole moment as a function of time.
5. In the text, we obtained the dynamics of the Jaynes–Cummings model assuming exact resonance, $\Delta = 0$. Reconsider the problem for the case where $\Delta \neq 0$. Obtain plots of the atomic inversion and note the effect of the nonzero detuning on the collapse and revivals of the Rabi oscillations. Perform an analysis to obtain the effect of the nonzero detuning on the collapse and revival times.
6. Consider the resonant Jaynes–Cummings model for the initial thermal state as in Section 4.7. Assume the atom is initially in the excited state. Analyse the collapse of the Rabi oscillations and determine the dependence of the collapse time on the average photon number of the thermal field.
7. Consider a simple model of degenerate Raman scattering, pictured in Fig. 4.12 (where $E_g = E_e$), and described by the interaction Hamiltonian $\hat{H}_{\rm I} = \hbar\lambda\hat{a}^\dagger\hat{a}(\sigma_+ + \sigma_-)$, where, as usual, $\sigma_+ = |e\rangle\langle g|$ and $\sigma_- = |g\rangle\langle e|$.
   (a) Obtain the dressed states for this model.
   (b) Assuming initially the field in a coherent state and the atom in the ground state, obtain the atomic inversion and show that the revivals of the Rabi oscillations are regular and complete.
   (c) Obtain the atomic inversion for an initial thermal state.
8. A resonant two-photon extension of the Jaynes–Cummings model is described by the effective Hamiltonian $H_{\rm eff} = \hbar\eta(\hat{a}^2\hat{\sigma}_+ + \hat{a}^{2\dagger}\hat{\sigma}_-)$, where, for the sake of simplicity, a small Stark shift term has been ignored. This Hamiltonian represents two-photon

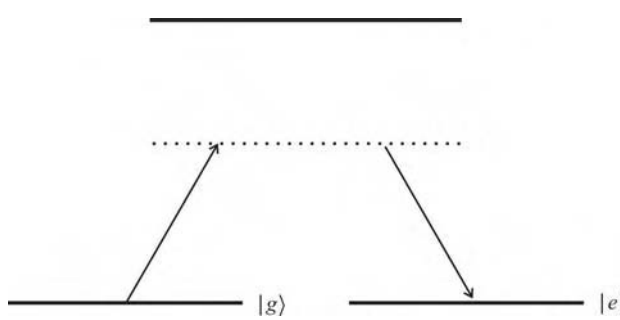


Fig. 4.12. Energy-level diagram for the degenerate Raman coupled model. The broken line represents a "virtual" intermediate state, too far off-resonance from a real level, the upper solid line, to become populated.

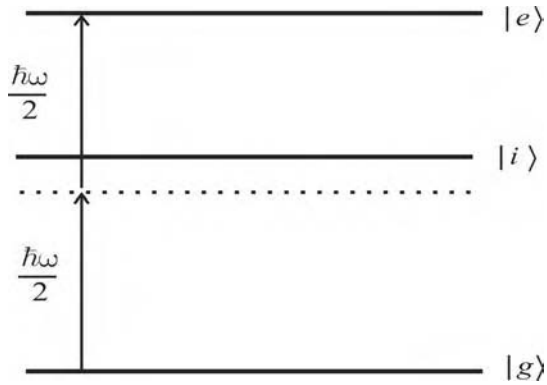


Fig. 4.13. Energy-level diagram for the resonant two-photon process. States $|e\rangle$ and $|g\rangle$ are of like parity whereas the intermediate state $|i\rangle$ is of opposite parity. The broken line represents a virtual atomic level, detuned from state $|i\rangle$.

absorption and emission between atomic levels of like parity. The process is represented by Fig. 4.13, where the broken line represents a virtual intermediate state of opposite parity.

(a) Obtain the dressed states for this system.

(b) Obtain the atomic inversion for this model assuming the atom initially in the ground state and that the field is initially in a number state. Repeat for a coherent state. Comment on the nature of the collapse and revival phenomena for these states.

(c) Obtain the atomic inversion for an initial thermal state.

9. A two-mode variation on the two-photon model of the previous problem is described by the Hamiltonian $H_{\text{eff}} = \hbar\eta(\hat{a}\,\hat{b}\hat{\sigma}_+ + \hat{a}^\dagger\hat{b}^\dagger\hat{\sigma}_-)$, that is, a photon from each mode is absorbed or emitted. Obtain the atomic inversion for the case where both modes are initially in coherent states. Analyse the collapse and revival phenomena.

11. Consider the resonant Jaynes–Cummings model with the atom initially in the excited state and the field initially in a coherent state with $\alpha = \sqrt{30}$.

(a) By taking the trace of the density operator over either the atomic or field states and thus obtaining a reduced density operator (see Appendix A), determine the degree to which the atom and field are entangled as a function of time. Pay particular attention to what happens at about midpoint between the initial collapse of the Rabi oscillations and the first revival. Is the behavior surprising? What connection, if any, is there between the result you have obtained and the results discussed in Section 4.10.

(b) Obtain plots (either contour or 3D) of the Q function (Eq. (3.112)) of the field for various times between time $t = 0$ and the time of the first revival of the Rabi oscillations.

12. Consider the atom–field state of Eq. (4.193) and assume that $\phi = 0$. Suppose that there is some way to determine the state of the atom (such is possible via field ionization, as will be discussed in Chapter 10). For example, if the atom is detected in state $|g\rangle$ then the field state will be reduced to $|-\alpha\rangle = \langle g \mid \psi(\pi)\rangle / |\langle g \mid \psi(\pi)\rangle|$ with a similar result for detection of the atomic state $|f\rangle$. (See Appendix D.) But suppose it were somehow

possible to detect the atom in the superposition states $|S\pm\rangle = (|g\rangle \pm |f\rangle)/\sqrt{2}$. What field states are generated in these cases?

13. Reconsider the problem discussed in Section 4.10 but with the initial state of the atom taken to be the "balanced" superposition $|\psi(0)\rangle_{\text{atom}} = (|e\rangle + |g\rangle)/\sqrt{2}$. Obtain relevant plots of the components of the Bloch vector and of the von Neumann entropy. Contrast the results obtained in this case to the case where the atom is initially in the excited state.

References

[1] J. J. Sakurai, *Modern Quantum Mechanics*, revised edition (Reading: Addison-Wesley, 1994), p. 316.

[2] Two standard references on nuclear magnetic resonance, wherein the rotating wave approximation is discussed, are: A. Abragam, *Principles of Nuclear Magnetism* (Oxford: Oxford University Press, 1961); C. P. Slichter, *Principles of Magnetic Resonance*, 3rd edition (Berlin: Springer, 1992).

[3] See Reference [1], p. 332.

[4] A. Einstein, *Verh. Deutsch. Phys. Ges*, **18** (1916), 318; *Mitt. Phys. Ges. Zürich*, **16** (1916), 47; *Phys. Zeitschr.*, **18** (1917), 121. This last article appears in an English translation from the German in D. ter Haar, *The Old Quantum Theory* (Oxford: Pergamon, 1967) and is reprinted in P. L. Knight and L. Allen, *Concepts in Quantum Optics* (Oxford: Pergamon, 1983). See also, A. Pais *'Subtle is the Lord . . .' The Science and Life of Albert Einstein* (Oxford: Oxford University Press, 1982), pp. 402–415.

[5] P. W. Milonni, *Am. J. Phys*., **52** (1984), 340.

[6] V. F. Weisskopf and E. Wigner, *Z. Phys*., **63** (1930), 54. See also, M. Sargent, III, M. O. Scully and W. E. Lamb, Jr., *Laser Physics* (Reading: Addison-Wesley, 1974), p. 236.

[7] I. I. Rabi, *Phys. Rev*., **51** (1937), 652.

[8] See References [2] and F. Bloch, *Phys. Rev*., **70** (1946), 460; E. L. Hahn, *Phys. Rev.*, **80** (1950), 580.

[9] See, for example, B. W. Shore, *The Theory of Coherent Atomic Excitation* (New York: Wiley, 1990), and references therein.

[10] E. T. Jaynes and F. W. Cummings, *Proc. IEEE*, **51** (1963), 89.

[11] J. J. Sanchez-Mondragon, N. B. Narozhny and J. H. Eberly, *Phys. Rev. Lett*., **51** (1983), 550; G. S. Agarwal, *J. Opt. Soc. Am. B*, **2** (1985), 480.

[12] F. W. Cummings, *Phys. Rev.*, **149** (1965), A1051; S. Stenholm, *Phys. Rep.*, **6** (1973), 1; P. Meystre, E. Geneux, A. Quattropani and A. Faust, *Nuovo Cim. B*, **25** (1975), 521; *J. Phys. A***8**, 95 (1975).

[13] J. H. Eberly, N. B. Narozhny and J. J. Sanchez-Mondragon, *Phys. Rev. Lett*., **44** (1980), 1323; N. B. Narozhny, J. J. Sanchez-Mondragon and J. H. Eberly, *Phys. Rev. A*, **23** (1981), 236; P. L. Knight and P. M. Radmore, *Phys. Rev. A*, **26** (1982), 676.

[14] H. J. Yoo and J. H. Eberly, *Phys. Rep.*, **118** (1985), 239; B. W. Shore and P. L. Knight, *J. Mod. Opt.*, **40** (1993), 1195.

[15] P. L. Knight and P. W. Milonni, *Phys. Rep.*, **66** (1980), 21.

[16] P. L. Knight and P. M. Radmore, *Phys. Lett*. A, **90** (1982), 342. Note that in this paper, the atom is assumed to be initially in the ground state.

[17] C. M. Savage, S. L. Braunstein and D. F. Walls, *Opt. Lett.*, **15** (1990), 628.
[18] E. Schrödinger, *Naturwissenshaften*, **23** (1935), 807, 823, 844. An English translation can be found in *Quantum Theory of Measurement*, edited by J. Wheeler and W. H. Zurek (Princeton: Princeton University Press, 1983).
[19] F. Le Kien and A. S. Shumovsky, *Intl. J. Mod. Phys. B*, **5** (1991), 2287; A. Messina, S. Maniscalco and A. Napoli, *J. Mod. Opt*., **50** (2003), 1.
[20] J. Gea-Banacloche, *Phys. Rev. Lett.*, **65** (1990), 3385.
[21] S. J. D. Phoenix and P. L. Knight, *Phys. Rev. A*, **44** (1991), 6023. See also S. J. D. Phoenix and P. L. Knight, *Ann. Phys. (N. Y.)*, **186** (1988), 381.

Bibliography

The following, in addition those already cited, is a selection of useful books on atom–field interactions, discussing both classically prescribed fields and quantum fields, with an orientation towards an audience of quantum opticians.

W. H. Louisell, *Quantum Statistical Properties of Radiation* (New York: Wiley, 1973).
L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (New York: Wiley, 1975; Mineola: Dover, 1987).
R. Loudon, *The Quantum Theory of Light*, 3rd edition (Oxford: Oxford University Press, 2000).
M. Weissbluth, *Photon–Atom Interactions* (New York: Academic Press, 1989).
P. Meystre and M. Sargent, III, *Elements of Quantum Optics*, 3rd edition (Berlin: Springer, 1999).
C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Atom–Photon Interactions* (New York: Wiley, 1992).
L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press, 1995).

Chapter 5
Quantum coherence functions

In this chapter, we discuss the classical and quantum theories of coherence. Beginning with the example of Young's interference, we review the notions of classical first-order coherence theory and the classical first-order coherence functions. We then proceed to introduce the quantum-mechanical first-order coherence functions and present a fully quantum-mechanical formulation of Young's interference. We next introduce the second-order coherence functions known also as the intensity–intensity correlation functions, and their connection with the Hanbury Brown and Twiss experiment, one of the pioneering developments of quantum optics. We finish the chapter by introducing all the higher-order coherence functions, using them to provide a criterion for full coherence of a quantized radiation field.

5.1 Classical coherence functions

We begin with a brief review of classical coherence [1] and we motivate this with Young's two-slit experiment as sketched in Fig. 5.1. Under certain conditions, interference fringes will appear on the screen. If the source has a bandwidth of $\Delta\omega$ and if $\Delta s = |s_1 - s_2|$ is the path difference, interference will occur if $\Delta s \leq c/\Delta\omega$. The quantity $\Delta s_{\text{coh}} = c/\Delta\omega$ is called the coherence length. The quantity $\Delta t_{\text{coh}} = \Delta s_{\text{coh}}/c = 1/\Delta\omega$ is called the coherence time. Interference fringes will be visible for $\Delta t_{\text{coh}}\Delta\omega \simeq 1$.

The field on the screen, or at the detector, at time t can be written as a linear superposition of the fields at the earlier times $t_1 = t - s_1/c$ and $t_2 = t - s_2/c$, i.e.

$$E(\mathbf{r}, t) = K_1 E(\mathbf{r}_1, t_1) + K_2 E(\mathbf{r}_2, t_2) \tag{5.1}$$

where $E(\mathbf{r}_i, t_i)$ is the complex field on the screen arriving from the ith slit and the quantities K_1 and K_2 are complex geometric factors that depend on the distances s_1 and s_2 respectively. For simplicity we are treating the fields as scalars here, equivalent to assuming that the fields have the same polarization. Diffraction effects associated with the slits are ignored. Optical light detectors have long

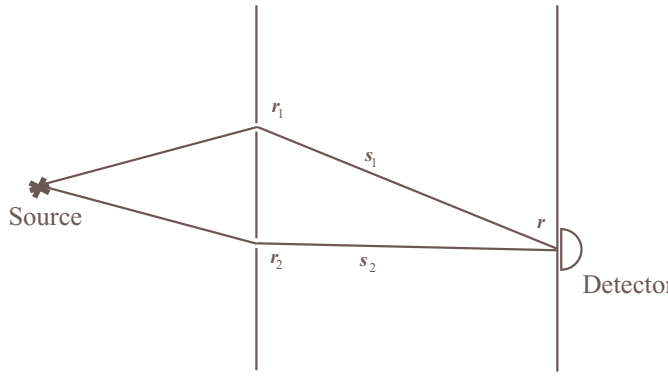


Fig. 5.1. Sketch of the standard setup for Young's double-slit interference experiment.

response times and are capable only of measuring the average light intensity. Thus the intensity of light on the detector is given by

$$I(\mathbf{r}) = \langle |E(\mathbf{r}, t)|^2 \rangle \tag{5.2}$$

where here the angular bracket means time average:

$$\langle f(t) \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t)\, dt. \tag{5.3}$$

We assume that this average is *stationary*, i.e. independent of the origin of the time axis. By the ergodic hypothesis, the time average will then be equivalent to an ensemble average. Using Eq. (5.1) we then have

$$\begin{aligned} I(\mathbf{r}) = {} & |K_1|^2 \langle |E(\mathbf{r}_1, t_1)|^2 \rangle \\ & + |K_2|^2 \langle |E(\mathbf{r}_2, t_2)|^2 \rangle \\ & + 2\mathrm{Re}\Big[K_1^* K_2 \langle E^*(\mathbf{r}_1, t_1) E(\mathbf{r}_2, t_2) \rangle \Big]. \end{aligned} \tag{5.4}$$

The first two terms are just the intensities associated with the fields from each of the slits while the third term gives rise to the interference. We set

$$\begin{aligned} I_1 &= |K_1|^2 \langle |E(\mathbf{r}_1, t_1)|^2 \rangle \\ I_2 &= |K_2|^2 \langle |E(\mathbf{r}_2, t_2)|^2 \rangle. \end{aligned} \tag{5.5}$$

We now introduce the first-order normalized mutual coherence function

$$\gamma^{(1)}(x_1, x_2) = \frac{\langle E^*(x_1) E(x_2) \rangle}{\sqrt{\langle |E(x_1)|^2 \rangle \langle |E(x_2)|^2 \rangle}}, \tag{5.6}$$

where we have followed the standard convention of setting $x_i = \mathbf{r}_i, t_i$. Thus we can write the intensity measured at the screen in terms of component intensities plus a coherence term:

$$I(\mathbf{r}) = I_1 + I_2 + 2\sqrt{I_1 I_2}\, \mathrm{Re}\Big[K_1 K_2 \gamma^{(1)}(x_1, x_2) \Big]. \tag{5.7}$$

If we now set $K_i = |K_i| \exp(i\psi_i)$ and write

$$\gamma^{(1)}(x_1, x_2) = |\gamma^{(1)}(x_1, x_2)| \exp(i\Phi_{12}) \tag{5.8}$$

then we have

$$I(\mathbf{r}) = I_1 + I_2 + 2\sqrt{I_1 I_2}|\gamma^{(1)}(x_1, x_2)| \cos(\Phi_{12} - \psi) \tag{5.9}$$

where $\psi = \psi_1 - \psi_2$ is a phase difference arising from the difference in path length (assumed smaller than the coherence length). Interference will occur for $|\gamma^{(1)}(x_1, x_2)| \neq 0$. We may discern three types of coherence:

$$\begin{aligned} &\left|\gamma^{(1)}(x_1, x_2)\right| = 1, && \text{complete coherence} \\ &0 < \left|\gamma^{(1)}(x_1, x_2)\right| < 1, && \text{partial coherence} \\ &\left|\gamma^{(1)}(x_1, x_2)\right| = 0, && \text{complete incoherence.} \end{aligned} \tag{5.10}$$

It is convenient to introduce Rayleigh's definition of fringe visibility:

$$\mathcal{V} = (I_{\max} - I_{\min})/(I_{\max} + I_{\min}), \tag{5.11}$$

where

$$I_{\substack{\max \\ \min}} = I_1 + I_2 \pm 2\sqrt{I_1 I_2}|\gamma^{(1)}(x_1, x_2)| \tag{5.12}$$

so that

$$\mathcal{V} = \frac{2\sqrt{I_1 I_2}|\gamma^{(1)}(x_1, x_2)|}{I_1 + I_2}\,. \tag{5.13}$$

Clearly, for complete coherence, the visibility (or contrast) for the fringes is a maximum:

$$\mathcal{V} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}\,, \tag{5.14}$$

whereas, for complete incoherence, there is zero visibility, $\mathcal{V} = 0$. We should further note an important property: going back to Eq. (5.6), it will be evident that if the numerator factorizes according to

$$\langle E^*(x_1)E(x_2)\rangle = \langle E^*(x_1)\rangle\langle E(x_2)\rangle \tag{5.15}$$

we then have $|\gamma^{(1)}(x_1, x_2)| = 1$. Thus the factorization property may be used as the criterion for complete optical coherence. The quantum analog of this criterion, the factorization of the expectation value of a product of field operators, will give rise to a criterion for the complete coherence in a quantum-mechanical light field, as we shall see.

We now consider some examples of the classical first-order coherence functions. We consider first the *temporal* coherence of stationary light fields at a fixed

position. For example, a monochromatic light field propagating in the z-direction is given at the position z at times t and $t + \tau$ as

$$\begin{aligned} E(z, t) &= E_0 e^{i(kz-\omega t)} \\ E(z, t+\tau) &= E_0 e^{i[kz-\omega(t+\tau)]}. \end{aligned} \tag{5.16}$$

These are the quantities $E(x_1)$ and $E(x_2)$ respectively from which we obtain

$$\langle E^*(z, t)E(z, t+\tau)\rangle = E_0^2 e^{-i\omega\tau}. \tag{5.17}$$

This is called the autocorrelation function and, dropping the position variable, is often written as $\langle E^*(t)E(t+\tau)\rangle$. It is clear in this case that

$$\gamma^{(1)}(x_1, x_2) = \gamma^{(1)}(\tau) = \gamma^{(1)*}(-\tau) = e^{-i\omega\tau} \tag{5.18}$$

and thus $|\gamma^{(1)}(\tau)| = 1$ so we have complete temporal coherence. But perfectly monochromatic sources of light do not exist. A more realistic model of a "monochromatic" source should take into account the random processes by which light is emitted from the decay of excited atoms in the source. These "monochromatic" sources will emit light as wave trains of finite length, where the trains are separated by a discontinuous change of phase. The average wave train time, τ_0, of a given source is called the coherence time of the source. This coherence time is inversely related to the natural line width of the spectral lines of the radiating atoms in the source. We can then define the coherence length ℓ_{coh} of a wave train as simply $\ell_{\text{coh}} = c\tau_0$. So we now consider the field

$$E(z, t) = E_0 e^{i(kz-\omega t)} e^{i\psi(t)} \tag{5.19}$$

where $\psi(t)$ is a random step function in the range $(0, 2\pi)$ with period τ_0 as shown in Fig. 5.2a. The autocorrelation function in this case is

$$\langle E^*(t)E(t+\tau)\rangle = E_0^2 e^{-i\omega\tau} \left\langle e^{i[\psi(t+\tau)-\psi(t)]} \right\rangle \tag{5.20}$$

so that

$$\gamma^{(1)}(\tau) = e^{-i\omega\tau} \lim_{T\to\infty} \int_0^T e^{i[\psi(t+\tau)-\psi(t)]}\, dt, \tag{5.21}$$

where the difference in the phases is shown in Fig. 5.2b.

Now for $0 < t < \tau_0$, $\psi(t+\tau) - \psi(t) = 0$ for $0 < t < \tau_0 - \tau$. But for $\tau_0 - \tau < t < \tau_0$, $\psi(t+\tau) - \psi(t)$ is a random number in the range $(0, 2\pi)$. The same holds for subsequent coherence time intervals. The integral can be evaluated (for the details see References [1]) to give

$$\begin{aligned} \gamma^{(1)}(\tau) &= \left(1 - \frac{\tau}{\tau_0}\right) e^{-i\omega\tau}, \quad \tau < \tau_0 \\ &= 0 \qquad\qquad\qquad\quad , \quad \tau \geq \tau_0, \end{aligned} \tag{5.22}$$

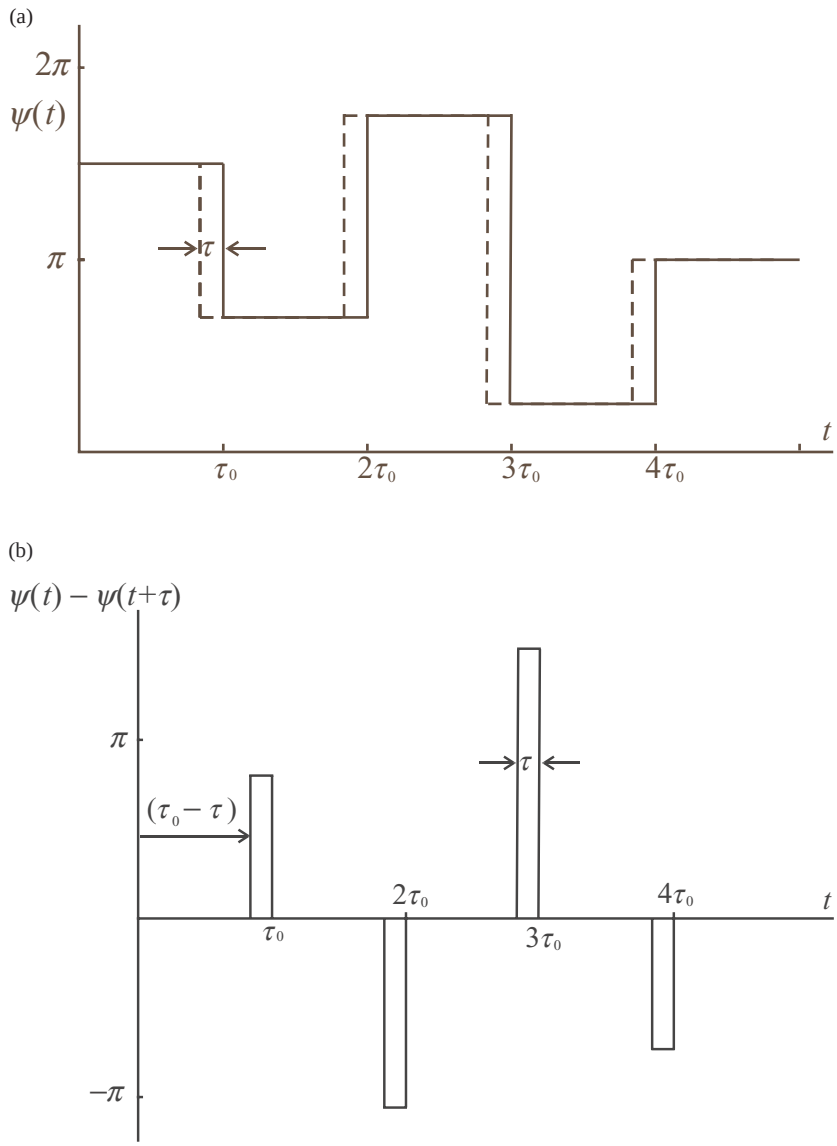


Fig. 5.2. (a) Random phase of period τ_0 $\psi(t)$ (solid line), $\psi(t+\tau)$ (dashed line), and (b) the phase difference $\psi(t) - \psi(t+\tau)$.

and so we have

$$\begin{aligned} \left|\gamma^{(1)}(\tau)\right| &= 1 - \frac{\tau}{\tau_0} \quad , \qquad \tau < \tau_0 \\ &= 0 \qquad\quad , \qquad \tau \geq \tau_0. \end{aligned} \tag{5.23}$$

Thus if the time delay τ is longer than the coherence time τ_0, there is no coherence.

A more realistic model of collision-broadened light is one for which the radiation has a Lorentzian power spectrum centered at frequency ω_0. In such a case the autocorrelation function is (assuming again propagation along the z-direction)

$$\langle E^*(t)\, E(t+\tau)\rangle = E_0^2 e^{-i\omega_0 t - |\tau|/\tau_0} \tag{5.24}$$

where τ_0 here may be interpreted as the average time between collisions. This results in

$$\begin{aligned} &\gamma^{(1)}(\tau) = e^{-i\omega_0\tau - |\tau|/\tau_0}, \\ &0 \le \left|\gamma^{(1)}(\tau)\right| \le 1. \end{aligned} \tag{5.25}$$

Evidently, $|\gamma^{(1)}(\tau)| \to 0$ as $\tau \to \infty$ so the beam becomes increasingly incoherent with increased time delay. If $|\tau| \ll \tau_0$ complete coherence is approached. Further, if the average time between collisions is shortened, then the spectrum is broadened and $|\gamma^{(1)}(\tau)| \to 0$ more rapidly with increasing τ. Light beams with such an autocorrelation function are known as *chaotic*, although other light beams with different power spectra, Gaussian, for example, are also termed chaotic.

5.2 Quantum coherence functions

Glauber and others, in a series of papers in the 1960s [2], showed how a quantum theory of coherence can be constructed upon observables in a manner that closely parallels the classical theory. The intensity of a light beam is measured by devices that actually attenuate the beam by absorbing photons. The actual determination of the intensity of a beam is determined by measuring in some way the response of the absorbing system. We shall consider an ideal detector consisting of a single atom of dimension small compared to the wavelength of the light. The absorption of light over a broad band of wavelengths results in the ionization of the atom and the subsequent detection of the photo-electron constitutes the detection of the photon. From the counting of the photo-electrons, the statistical properties of the light field may be determined.

The single-atom detector couples to the quantized field through the dipole interaction (as discussed in Chapter 4),

$$\hat{H}^{(\mathrm{I})} = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}}(\mathbf{r}, t) \tag{5.26}$$

where we are using the Heisenberg picture in which the electric field operator can be written as

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i \sum_{\mathbf{k},s} \left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{1/2} \mathbf{e}_{\mathbf{k}s} \left[\hat{a}_{\mathbf{k}s}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}^\dagger_{\mathbf{k}s}(t) e^{-i\mathbf{k}\cdot\mathbf{r}}\right]. \tag{5.27}$$

Because the wavelength of the light is long compared with the dimensions of the atom, $|\mathbf{k}\cdot\mathbf{r}| \ll 1$ (the dipole approximation) and thus at the atomic detector the

field is

$$\hat{\mathbf{E}}(\mathbf{r},t) \approx i \sum_{\mathbf{k},s} \sqrt{\frac{\hbar\omega_k}{2\varepsilon_0 V}}\, \mathbf{e}_{\mathbf{k}s} \left[\hat{a}_{\mathbf{k}s}(t) - \hat{a}^{\dagger}_{\mathbf{k}s}(t)\right]. \tag{5.28}$$

Furthermore, the component of the field describing absorption is the positive frequency part

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = i \sum_{\mathbf{k},s} \left(\frac{\hbar\omega_k}{2\varepsilon_0 V}\right)^{1/2} \mathbf{e}_{\mathbf{k}s}\, \hat{a}_{\mathbf{k}s}(t) \tag{5.29}$$

as it contains the annihilation operator. We assume that the atom is initially in some state $|g\rangle$ and the field is in some state $|i\rangle$. Upon the absorption of radiation, the atom makes a transition to state $|e\rangle$ and the field to state $|f\rangle$. We assume that $|e\rangle$ is an ionized atomic state and approximate it as a free electron state (a highly idealized model of photo-detection)

$$|e\rangle = \frac{1}{\sqrt{V}}\, e^{i\mathbf{q}\cdot\mathbf{r}}, \tag{5.30}$$

where $\hbar\mathbf{q}$ is the momentum of the ionized electron. Thus with the initial state $|I\rangle = |g\rangle|i\rangle$ and final state $|F\rangle = |e\rangle|f\rangle$ the matrix element for the transition is

$$\langle F|\, \hat{H}^{(\mathrm{I})}\, |I\rangle = -\langle e|\, \hat{\mathbf{d}}\, |g\rangle\, \langle f|\, \hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\, |i\rangle\,. \tag{5.31}$$

The transition probability for the atom–field system is proportional to

$$\left|\langle F|\, \hat{H}^{(\mathrm{I})}\, |I\rangle\right|^2. \tag{5.32}$$

The probability that the *field* undergoes a transition from $|i\rangle$ to $|f\rangle$ is proportional to

$$\left|\langle f|\, \hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\, |i\rangle\right|^2. \tag{5.33}$$

We are really only interested in the final state of the detector, not the field, so we must sum over all the possible final states. For all practical purposes, this set of final states may be regarded as complete (and can always *be* completed with the addition of states for unallowed transactions) so that

$$\begin{aligned}
&\sum_f \left|\langle f|\, \hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\, |i\rangle\right|^2 \\
&\quad = \sum_f \langle i|\hat{\mathbf{E}}^{(-)}(\mathbf{r},t)\, |f\rangle\, \langle f|\, \hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\, |i\rangle \\
&\quad = \langle i|\, \hat{\mathbf{E}}^{(-)}(\mathbf{r},t) \cdot \hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\, |i\rangle\,,
\end{aligned} \tag{5.34}$$

where, of course, $\hat{\mathbf{E}}^{(-)}(\mathbf{r},t) = [\hat{\mathbf{E}}^{(+)}(\mathbf{r},t)]^{\dagger}$.

In the preceding, we have assumed the field to be initially in a pure state. More likely, it will be initially in a mixed state described by a density operator

of the form

$$\hat{\rho}_{\mathrm{F}} = \sum_i P_i \left|i\right\rangle \left\langle i\right| . \tag{5.35}$$

In this case, the expectation value of Eq. (5.34) is replaced by the ensemble average

$$\begin{aligned} &\mathrm{Tr}\left\{\hat{\rho}_{\mathrm{F}} \hat{\mathbf{E}}^{(-)}(\mathbf{r},t)\cdot\hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\right\} \\ &\quad = \sum_i P_i \left\langle i\right| \hat{\mathbf{E}}^{(-)}(\mathbf{r},t)\cdot\hat{\mathbf{E}}^{(+)}(\mathbf{r},t)\left|i\right\rangle . \end{aligned} \tag{5.36}$$

Note the normal ordering of the operators; this is a consequence of our use of an absorbing detector. Henceforth we shall resort to our earlier convention that the fields are of the same polarization, i.e. that we can treat them as scalars. We again use the notation $x = (\mathbf{r}, t)$ and define the function

$$G^{(1)}(x,x) = \mathrm{Tr}\left\{\rho \hat{E}^{(-)}(x)\,\hat{E}^{(+)}(x)\right\} \tag{5.37}$$

which is the intensity of light at the space-time point $x = \mathbf{r}, t$; i.e. $I(\mathbf{r},t) = G^{(1)}(x,x)$ is the quantum analog of the classical expression for the time or ensemble averages of the previous section. For Young's interference experiment as pictured in Fig. 5.1, the positive frequency part of the field at a photo-detector located at position $\mathbf{r}$ at time t is of the superposition of the fields from the two slits:

$$\hat{E}^{(+)}(\mathbf{r},t) = K_1 \hat{E}^{(+)}(\mathbf{r}_1,t_1) + K_2 \hat{E}^{(+)}(\mathbf{r}_2,t_2) \tag{5.38}$$

which is just the positive frequency part of the quantum analog of the classical relation in Eq. (5.1). The intensity of the light on the screen (actually the photo-detector) is

$$\begin{aligned} I(\mathbf{r},t) &= \mathrm{Tr}\left\{\hat{\rho}\hat{E}^{(-)}(\mathbf{r},t)\hat{E}^{(+)}(\mathbf{r},t)\right\} \\ &= |K_1|^2 G^{(1)}(x_1,x_1) + |K_2|^2 G^{(1)}(x_2,x_2) \\ &\quad + 2\mathrm{Re}\left[K_1^* K_2 G^{(1)}(x_1,x_2)\right] \end{aligned} \tag{5.39}$$

where

$$G^{(1)}(x_i,x_i) = \mathrm{Tr}\left\{\hat{\rho}\hat{E}^{(-)}(x_i)\,\hat{E}^{(+)}(x_i)\right\} \qquad i = 1,2,\ldots \tag{5.40}$$

and where

$$G^{(1)}(x_1,x_2) = \mathrm{Tr}\left\{\hat{\rho}\hat{E}^{(-)}(x_1)\,\hat{E}^{(+)}(x_2)\right\} \tag{5.41}$$

for $x_1 = \mathbf{r}_1, t_1$ and $x_2 = \mathbf{r}_2, t_2$. This last quantity is a general first-order correlation function, Eq. (5.40) being special cases. Note that $G^{(1)}(x_i,x_i)$ is just the intensity at the detector of light arriving from x_{i} whereas $G^{(1)}(x_1,x_2)$ is a measure of the correlation of light arriving from both x_1 and x_2, thus a measure of interference.

We may further define, in analogy with the classical coherence function $\gamma^{(1)}(x_1, x_2)$ of Eq. (5.6), the normalized first-order quantum coherence function

$$g^{(1)}(x_1, x_2) = \frac{G^{(1)}(x_1, x_2)}{\left[G^{(1)}(x_1, x_1)\, G^{(1)}(x_2, x_2)\right]^{1/2}} \tag{5.42}$$

for which

$$0 \leq \left|g^{(1)}(x_1, x_2)\right| \leq 1. \tag{5.43}$$

As in the classical case, we may discern three types of coherence:

$$\begin{aligned} &\left|g^{(1)}(x_1, x_2)\right| = 1, && \text{complete coherence} \\ &0 < \left|g^{(1)}(x_1, x_2)\right| < 1, && \text{partial coherence} \\ &\left|g^{(1)}(x_1, x_2)\right| = 0 && \text{incoherent.} \end{aligned} \tag{5.44}$$

From the definition of the correlation function it is possible to show that

$$\begin{aligned} &G^{(1)}(x_1, x_2) = \left[G^{(1)}(x_2, x_1)\right]^* \\ &G^{(1)}(x_i, x_i) \geq 0 \\ &G^{(1)}(x_1, x_1)\, G^{(1)}(x_2, x_2) \geq \left|G^{(1)}(x_1, x_2)\right|^2. \end{aligned} \tag{5.45}$$

Equality in the last expression implies maximum fringe visibility, i.e. $|g^{(1)}(x_1, x_2)| = 1$.

For a single-mode plane-wave quantized field propagating with wave vector $\mathbf{k}$, the positive frequency part of the field operator is

$$\hat{E}^{(+)}(x) = i K \hat{a}\, e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \tag{5.46}$$

where

$$K = \left(\frac{\hbar\omega}{2\varepsilon_0 V}\right)^{1/2}. \tag{5.47}$$

If the field is in a number state $|n\rangle$ then

$$\begin{aligned} G^{(1)}(x, x) &= \langle n|\, \hat{E}^{(-)}(x)\, \hat{E}^{(+)}(x)\, |n\rangle \\ &= K^2 n \end{aligned} \tag{5.48}$$

$$\begin{aligned} G^{(1)}(x_1, x_2) &= \langle n|\, \hat{E}^{(-)}(x_1)\, \hat{E}^{(+)}(x_2)\, |n\rangle \\ &= K^2 n \exp\{i[\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2) - \omega(t_2 - t_1)]\} \end{aligned} \tag{5.49}$$

and thus

$$\left|g^{(1)}(x_1, x_2)\right| = 1. \tag{5.50}$$

For a coherent state $|\alpha\rangle$

$$\begin{aligned} G^{(1)}(x, x) &= \langle \alpha|\, \hat{E}^{(-)}(x)\, \hat{E}^{(+)}(x)\, |\alpha\rangle \\ &= K^2 |\alpha|^2, \end{aligned} \tag{5.51}$$

$$\begin{aligned} G^{(1)}(x_1, x_2) &= \langle \alpha|\, \hat{E}^{(-)}(x_1)\, \hat{E}^{(+)}(x_2)\, |\alpha\rangle \\ &= K^2 |\alpha|^2 \exp\{i[\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2) - \omega(t_2 - t_1)]\}. \end{aligned} \tag{5.52}$$

So we obtain $|g^{(1)}(x_1, x_2)| = 1$ as before. As in the classical case, the key to first-order quantum coherence is the factorization of the expectation value of the correlation function in the numerator of Eq. (5.42), i.e.:

$$G^{(1)}(x_1, x_2) = \left\langle \hat{E}^{(-)}(x_1)\, \hat{E}^{(+)}(x_2) \right\rangle = \left\langle \hat{E}^{(-)}(x_1) \right\rangle \left\langle \hat{E}^{(+)}(x_2) \right\rangle. \tag{5.53}$$

This condition is satisfied for the number and coherent states. The case for a thermal state will be left as an exercise for the reader.

5.3 Young's interference

As a last example of first-order quantum coherence we examine in some detail Young's interference experiment. We follow the presentation of Walls [3]. Again referring to Fig. 5.1, we assume that the source of light is monochromatic and we assume that the pinholes have dimensions of the order of the wavelength of the light. The latter assumption allows us to ignore diffraction effects and allows us to assume that the pinholes are point sources of spherical waves, and that the field at the screen (or at a detector) at position $\mathbf{r}$ at time t is the sum of the spherical waves from each pinhole:

$$\hat{E}^{(+)}(\mathbf{r}, t) = f(r)[\hat{a}_1 e^{iks_1} + \hat{a}_2 e^{iks_2}]\, e^{-i\omega t} \tag{5.54}$$

where

$$f(r) = i \left[\frac{\hbar\omega}{2\varepsilon_0 (4\pi R)} \right]^{1/2} \frac{1}{r}. \tag{5.55}$$

R is the radius of the normalization volume, s_1 and s_2 are the distances from the pinholes to the screen and the $r = |\mathbf{r}|$ appearing in Eq. (5.55) is the result of the approximation $s_1 \approx s_2 = r$. Also, $k = |\mathbf{k}_1| = |\mathbf{k}_2|$ is the magnitude of the wave numbers of the beams from each of the pinholes. The field operators $\hat{a}_1$ and $\hat{a}_2$ are associated with the *radial* modes of the field for photons emitted from pinholes 1 and 2 respectively. The intensity is given by

$$\begin{aligned} I(\mathbf{r}, t) &= \mathrm{Tr}\left\{ \hat{\rho}\, \hat{E}^{(-)}(\mathbf{r}, t)\, \hat{E}^{(+)}(\mathbf{r}, t) \right\} \\ &= |f(r)|^2 \left\{ \mathrm{Tr}\left(\hat{\rho}\, \hat{a}_1^\dagger \hat{a}_1 \right) + \mathrm{Tr}\left(\hat{\rho}\, \hat{a}_2^\dagger \hat{a}_2 \right) \right. \\ &\quad \left. + 2 \left| \mathrm{Tr}\left(\hat{\rho}\, \hat{a}_1^\dagger \hat{a}_2 \right) \right| \cos\Phi \right\} \end{aligned} \tag{5.56}$$

where

$$\mathrm{Tr}\left(\hat{\rho}\, \hat{a}_1^\dagger \hat{a}_2 \right) = \left| \mathrm{Tr}\left(\hat{\rho}\, \hat{a}_1^\dagger \hat{a}_2 \right) \right| e^{i\psi} \tag{5.57}$$

and

$$\Phi = k(s_1 - s_2) + \psi. \tag{5.58}$$

Maximum visibility of the interference fringes occurs for $\Phi = 2\pi m$, where m is an integer, and falls off as $1/r^2$ with increasing distance of the detector from the central fringe.

But we note that the beam falling onto the pinholes can be approximated as a plane-wave mode, and take the corresponding field operator as $\hat{a}$. If the pinholes are of equal size and if we were to put detectors right behind each of them, then incident photons would be equally likely to go through either one. That is, the two pinholes act to split a single beam into two beams. Thus we can write

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{a}_1 + \hat{a}_2) \tag{5.59}$$

where

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij}, \qquad \left[\hat{a}_i, \hat{a}_j^\dagger\right] = 0, \qquad [\hat{a}, \hat{a}^\dagger] = 1. \tag{5.60}$$

Unfortunately, the relation in Eq. (5.59) does not by itself constitute a unitary transformation. There needs to be another mode introduced, what we shall call a "fictitious" mode, to make it unitary. (The reasoning behind this is discussed in detail in Chapter 6.) We shall call this fictitious mode $\hat{b}$, and will let $\hat{b} = (\hat{a}_1 - \hat{a}_2)/\sqrt{2}$, where $[\hat{b}, \hat{b}^\dagger] = 1$. This mode will always be in the vacuum state so it may be ignored in what follows, but we shall keep it for expository purposes.

Assuming the photons incident on the pinholes are in the a mode, in keeping with our description above, an incident state containing n photons is the product state $|n\rangle_a|0\rangle_b$. It is related to a_1 and a_2 mode states according to

$$\begin{aligned} |n\rangle_a\,|0\rangle_b &= \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}\,|0\rangle_a\,|0\rangle_b \\ &= \frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}}\right)^n\left(\hat{a}_1^\dagger + \hat{a}_2^\dagger\right)^n |0\rangle_1\,|0\rangle_2 \end{aligned} \tag{5.61}$$

where $|0\rangle_1|0\rangle_2$ represents the product of the a_1 and a_2 mode's vacuum states. For the case of only one photon we have

$$\begin{aligned} |1\rangle_a\,|0\rangle_b &= \frac{1}{\sqrt{2}}\left(|1\rangle_1\,|0\rangle_2 + |0\rangle_1\,|1\rangle_2\right), \\ &= \frac{1}{\sqrt{2}}\left(|1,0\rangle + |0,1\rangle\right), \end{aligned} \tag{5.62}$$

where in the second line we have introduced compact notation where $|1,0\rangle$ means one photon in mode 1 and zero photons in mode 2 and vice versa for $|0,1\rangle$. Note that $\hat{b} = (\hat{a}_1 - \hat{a}_2)/\sqrt{2}$ operating on the state of Eq. (5.62) yields zero identically, consistent with the idea that photons are not present in the fictitious mode. If the incident state is $|1\rangle_a|0\rangle_b$ then the intensity in the Young's experiment is given by

$$\begin{aligned} I(\mathbf{r}, t) = |f(\mathbf{r})|^2 \Big\{ &\frac{1}{2}\langle 1,0|\,\hat{a}_1^\dagger\hat{a}_1\,|0,1\rangle \\ &+ \frac{1}{2}\langle 0,1|\,\hat{a}_2^\dagger\hat{a}_2\,|0,1\rangle \\ &+ \langle 1,0|\,\hat{a}_1^\dagger\hat{a}_2\,|0,1\rangle\cos\Phi\Big\} \end{aligned} \tag{5.63}$$

which reduces to

$$I(\mathbf{r}, t) = |f(\mathbf{r})|^2 [1 + \cos\Phi]. \tag{5.64}$$

In the case of two photons, we have from Eq. (5.61),

$$|2\rangle_a |0\rangle_b = \frac{1}{2}[|2, 0\rangle + \sqrt{2}\,|1, 1\rangle + |0, 2\rangle] \tag{5.65}$$

where the meanings of the state labels on the right-hand side should be clear. (Note that the application of $\hat{b} = (\hat{a}_1 - \tilde{a}_2)/\sqrt{2}$ to Eq. (5.65) again yields zero.) This state results in the intensity

$$I(\mathbf{r}, t) = 2\,|f(\mathbf{r})|^2 [1 + \cos\Phi]. \tag{5.66}$$

The general n-photon state gives

$$I(\mathbf{r}, t) = n\,|f(\mathbf{r})|^2 [1 + \cos\Phi]. \tag{5.67}$$

For a coherent state incident on the pinholes, we have, by using the displacement operator and Eq. (5.59),

$$\begin{aligned} |\alpha\rangle_a |0\rangle_b &= \hat{D}_a(\alpha) |0\rangle_a |0\rangle_b = \hat{D}_1\left(\frac{\alpha}{\sqrt{2}}\right) \hat{D}_2\left(\frac{\alpha}{\sqrt{2}}\right) |0, 0\rangle \\ &= \left|\frac{\alpha}{\sqrt{2}}\right\rangle_1 \left|\frac{\alpha}{\sqrt{2}}\right\rangle_2 \end{aligned} \tag{5.68}$$

where

$$\hat{D}_i(\alpha/\sqrt{2}) = \exp\left(\frac{\alpha}{\sqrt{2}}\hat{a}^\dagger - \frac{\alpha^*}{\sqrt{2}}\hat{a}\right) \qquad i = 1, 2. \tag{5.69}$$

The intensity in this case is

$$I(\mathbf{r}, t) = |\alpha|^2\,|f(\mathbf{r})|^2 [1 + \cos\Phi]. \tag{5.70}$$

In all of these cases, the two interfering modes are derived from the same single-mode beam incident on the two holes. They are all first-order coherent, giving rise to the same interference pattern; only the overall intensity is affected by the precise number, or average number, of photons. Note that throughout, as promised, the fictitious b mode remains in the vacuum and as such we could ignore it, as in fact, did Walls [3], but we think it is important to recognize that unitarity must be preserved in all transformations.

Equation (5.63) seems to confirm Dirac's [4] famous remark that "each photon interferes only with itself, interference between different photons does not occur". Unfortunately, this remark, probably intended as a metaphor to delineate the sharp contrast between the classical and quantum pictures of the interference of light, has been taken out of context and interpreted by some to imply that photons from independent light beams cannot interfere. If the modes a_1 and a_2 are truly independent then the product number state $|n_1\rangle|n_2\rangle$, in fact, does not give rise to interference fringes, as is easy to see. But if the modes are in coherent states

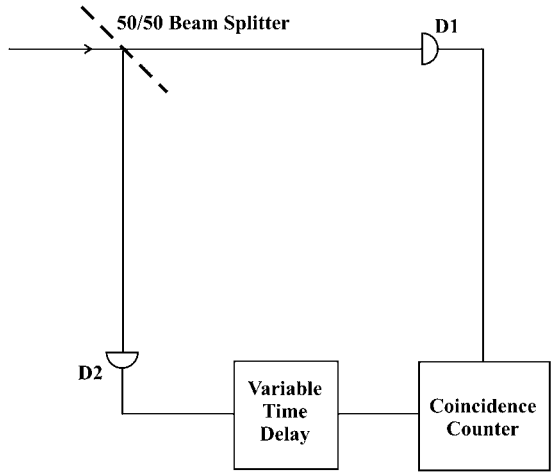


Fig. 5.3. Sketch of the setup for the Hanbury-Brown and Twiss experiment.

$|\alpha_1\rangle|\alpha_2\rangle$ arising, say, from *independent* lasers, then

$$I(\mathbf{r}, t) = |f(\mathbf{r})|^2\{|\alpha_1|^2 + |\alpha_2|^2 + 2|\alpha_1^*\alpha_2| \cos \Phi\} \tag{5.71}$$

which clearly exhibits interference fringes. Interference in light emitted by independent lasers was demonstrated experimentally by Magyar and Mandel [5] just a few years after the invention of the laser.

5.4 Higher-order coherence functions

First-order coherence in Young's experiment may be understood mathematically as the result of the factorization of the expectation values in the correlation function of the fields in both the classical and quantum cases. Such an experiment is able to determine the degree to which a light source is monochromatic, or to determine the coherence length of the light, but it says nothing about the statistical properties of the light. That is, first-order coherence experiments are unable to distinguish between states of light with identical spectral distributions but with quite different photon number distributions. We have seen that a single-mode field in either a number state or a coherent state is first-order quantum coherent yet the photon distributions of these states are strikingly different.

In the 1950s, Hanbury Brown and Twiss in Manchester [6] developed a new kind of correlation experiment that involved the correlation of intensities rather than of fields. A sketch of the experiment is shown in Fig. 5.3. Detectors $D1$ and $D2$ are the same distance from the beam splitter. This setup measures a delayed coincidence rate where one of the detectors registers a count at time t and the other a count at $t + \tau$. If τ, the time delay, is smaller than the coherence time τ_0, information on the statistics of the light beam striking the beam splitter can be determined.

The rate of coincident counts is proportional to the time, or ensemble, average

$$C(t, t+\tau) = \langle I(t)\, I(t+\tau)\rangle \tag{5.72}$$

where $I(t)$ and $I(t+\tau)$ are the instantaneous intensities at the two detectors (these are classical quantities here). If we assume that the fields are stationary, the average is a function only of t. If the average of the intensity at each detector is $\langle I(t)\rangle$, then the probability of obtaining a coincidence count with time delay τ is

$$\begin{aligned}\gamma^{(2)}(\tau) &= \frac{\langle I(t)\, I(t+\tau)\rangle}{\langle I(t)\rangle^2} \\ &= \frac{\langle E^*(t)\, E^*(t+\tau)\, E(t+\tau)\, E(t)\rangle}{\langle E^*(t)\, E(t)\rangle^2}.\end{aligned} \tag{5.73}$$

This is the classical second-order coherence function. If the detectors are at different distances from the beam splitter, the second-order coherence function can be generalized to

$$\begin{aligned}\gamma^{(2)}(x_1, x_2; x_2, x_1) &= \frac{\langle I(x_1)\, I(x_2)\rangle}{\langle I(x_1)\rangle\ \langle I(x_2)\rangle} \\ &= \frac{\langle E^*(x_1)\, E^*(x_2)\, E(x_2)\, E(x_1)\rangle}{\langle |E(x_1)|^2\rangle\ \langle |E(x_2)|^2\rangle}.\end{aligned} \tag{5.74}$$

By analogy to first-order coherence, there is said to be classical coherence to second order if $|\gamma^{(1)}(x_1, x_2)| = 1$ and $\gamma^{(2)}(x_1, x_2; x_2, x_1) = 1$. The second condition requires the factorization

$$\langle E^*(x_1) E^*(x_2) E(x_2) E(x_1)\rangle = \langle |E(x_1)|^2\rangle\ \langle |E(x_2)|^2\rangle. \tag{5.75}$$

For a plane wave propagating in the z-direction, as given by Eq. (5.16), it is easy to show that

$$\langle E^*(t)\, E^*(t+\tau)\, E(t+\tau)\, E(t)\rangle = E_0^4 \tag{5.76}$$

and thus $\gamma^{(2)}(\tau) = 1$. For any light beam of constant, non-fluctuating, intensity, we have $I(t) = I(t+\tau) = I_0$, $\gamma^{(2)}(\tau) = 1$.

However, unlike the first-order coherence function, the second-order coherence function is not restricted to be unity or less. To see this, we first consider the zero time-delay coherence function

$$\gamma^{(2)}(0) = \frac{\langle I(t)^2\rangle}{\langle I(t)\rangle^2}. \tag{5.77}$$

For a sequence of N measurements taken at times $t_1, t_2, \ldots, t_N$, the required averages are given by

$$\begin{aligned}\langle I(t)\rangle &= \frac{I(t_1) + I(t_2) + \cdots + I(t_N)}{N} \\ \langle I(t)^2\rangle &= \frac{I(t_1)^2 + I(t_2)^2 + \cdots + I(t_N)^2}{N}.\end{aligned} \tag{5.78}$$

Now according to Cauchy's inequality applied to a pair of measurements at times t_1 and t_2, we have

$$2I(t_1)\,I(t_2) \le I(t_1)^2\,I(t_2)^2\,. \tag{5.79}$$

Applying this to all the cross terms in $\langle I(t)\rangle^2$ it follows that

$$\langle I(t)^2\rangle \ge \langle I(t)\rangle^2 \tag{5.80}$$

and thus

$$1 \le \gamma^{(2)}(0) < \infty, \tag{5.81}$$

there being no way to establish an upper limit.

For nonzero time delays, the positivity of the intensity ensures that $0 \le \gamma^{(2)}(\tau) < \infty$, $(\tau \ne 0)$. But from the inequality of Eq. (5.81) it can be established that

$$\begin{aligned}&[I(t_1)I(t_1+\tau)+\cdots+I(t_N)I(t_N+\tau)]^2\\ &\quad\le [I(t_1)^2+\cdots+I(t_N)^2][I(t_1+\tau)^2+\cdots+I(t_N+\tau)^2].\end{aligned} \tag{5.82}$$

For a long series of many measurements, the two series on the right-hand side are equivalent so that

$$\langle I\,(t)\,I\,(t+\tau)\rangle \le \langle I\,(t)\rangle^2 \tag{5.83}$$

and thus we arrive at

$$\gamma^{(2)}\,(\tau) \le \gamma^{(2)}\,(0)\,. \tag{5.84}$$

The results reported in Eqs. (5.81) and (5.84) establish limits for classical light fields. Later we shall show that some quantum states of light violate the quantum-mechanical version of the inequality of Eq. (5.84).

For a light source containing a large number of independently radiating atoms undergoing collisional broadening, it can be shown [7] that the first- and second-order coherence functions are related according to

$$\gamma^{(2)}\,(\tau) = 1 + \left|\gamma^{(1)}\,(\tau)\right|^2\,, \tag{5.85}$$

a relation that holds for all kinds of chaotic light. Evidently, since $0 \le |\gamma^{(1)}(\tau)| \le 1$ it follows that $1 \le |\gamma^{(2)}(\tau)| \le 2$. From the result in Eq. (5.25) we have, for sources with Lorentzian spectra,

$$\gamma^{(2)}\,(\tau) = 1 + e^{-2|\tau|/\tau_0}. \tag{5.86}$$

Although for $\tau \to \infty$, $\gamma^{(2)}(\tau) \to 1$, it is evident that for zero time delay, $\tau \to 0$, $\gamma^{(2)}(0) = 2$. In fact, for any kind of chaotic light, $\gamma^{(2)}(0) = 2$. The implication of this result is as follows. If light incident on one of the detectors is independent of the light incident on the other, there should be a uniform coincidence rate independent of t. This is what Hanbury Brown and Twiss [6] expected. By using

an elementary (but wrong!) picture in which the photons are emitted independently by the source, and assuming that the beam splitter did not split photons but merely reflected or transmitted them, Hanbury Brown and Twiss expected to be able to demonstrate the existence of photons. They found, for zero time delay, twice the detection rate compared with the rate at long time delays. If photons exist, they evidently arrive in pairs at zero time delay but independently at long time delays. That photons arrive in "bunched" pairs is now known as the *photon bunching effect* (also known as the Hanbury Brown and Twiss effect). Note that by measuring the coincidence counts at increasing delay times, it is possible to measure the coherence time τ_0 of the source.

In the previous paragraph, we spoke of the bunching of photons, even though the important result of Eq. (5.86) was not derived on the basis of the quantum theory of light. We now introduce the second-order quantum coherence function and show that light in a coherent state, as obtained (to a reasonable approximation) from a well-stabilized laser, is coherent to second order and that thermal light sources exhibit the photon bunching effect. In these cases, the quantum and classical pictures agree, but it will become clear that there are instances where the quantum theory predicts situations for which there is no classical counterpart.

We extend the argument used in the first-order case to the detection of two photons by absorption. The transition probability for the absorption of two photons is proportional to

$$\left| \langle f | \, \hat{E}^{(+)}(\mathbf{r}_2, t_2) \, \hat{E}^{(+)}(\mathbf{r}_1, t_1) \, |i\rangle \right|^2 \tag{5.87}$$

which, after summation over all final states becomes

$$\langle i | \, \hat{E}^{(-)}(\mathbf{r}_1, t_1) \, \hat{E}^{(-)}(\mathbf{r}_2, t_2) \, \hat{E}^{(+)}(\mathbf{r}_2, t_2) \, \hat{E}^{(+)}(\mathbf{r}_1, t_1) \, |i\rangle \, . \tag{5.88}$$

Generalizing to cases of nonpure field states we introduce the second-order quantum correlation function

$$G^{(2)}(x_1, x_2; x_2, x_1) = \mathrm{Tr} \left\{ \hat{\rho} \hat{E}^{(-)}(x_1) \, \hat{E}^{(-)}(x_2) \, \hat{E}^{(+)}(x_2) \, \hat{E}^{(+)}(x_1) \right\} \tag{5.89}$$

which is to be interpreted as the ensemble average of $I(x_1)I(x_2)$. As in the first-order case, the normal ordering of the field operators for absorptive detection is important and must be preserved. We define the second-order quantum coherence function as

$$g^{(2)}(x_1, x_2; x_2, x_1) = \frac{G^{(2)}(x_1, x_2; x_2, x_1)}{G^{(1)}(x_1, x_1) \, G^{(1)}(x_2 x_2)} \tag{5.90}$$

where $g^{(2)}(x_1, x_2; x_2, x_1)$ is the joint probability of detecting one photon at $\mathbf{r}_1$ at time t_1 and a second at $\mathbf{r}_2$ at time t_2. A quantum field is said to be second-order coherent if $|g^{(1)}(x_1, x_2)| = 1$ and $g^{(2)}(x_1, x_2; x_2, x_1) = 1$. This requires that $G^{(2)}(x_1, x_2; x_2, x_1)$ factorize according to

$$G^{(2)}(x_1, x_2; x_2, x_1) = G^{(1)}(x_1, x_1) \, G^{(1)}(x_2, x_2) \, . \tag{5.91}$$

At a fixed position, $g^{(2)}$ depends only on the time difference $\tau = t_2 - t_1$:

$$g^{(2)}(\tau) = \frac{\left\langle \hat{E}^{(-)}(t)\,\hat{E}^{(-)}(t+\tau)\,\hat{E}^{(+)}(t+\tau)\,\hat{E}^{(+)}(t)\right\rangle}{\left\langle \hat{E}^{(-)}(t)\,\hat{E}^{(+)}(t)\right\rangle \left\langle \hat{E}^{(-)}(t+\tau)\,\hat{E}^{(+)}(t+\tau)\right\rangle} \tag{5.92}$$

which is interpreted as the conditional probability that if a photon is detected at t one is also detected at $t + \tau$.

For a single-mode field of the form given by Eq. (5.46), $g^{(2)}(\tau)$ reduces to

$$\begin{aligned} g^{(2)}(\tau) &= \frac{\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a}\rangle}{\langle \hat{a}^\dagger \hat{a}\rangle^2} = \frac{\langle \hat{n}(\hat{n}-1)\rangle}{\langle \hat{n}\rangle^2} \\ &= 1 + \frac{\langle (\Delta \hat{n})^2\rangle - \langle \hat{n}\rangle}{\langle \hat{n}\rangle^2}, \end{aligned} \tag{5.93}$$

which is independent of τ.

For the field in a coherent state $|\alpha\rangle$ it follows that

$$g^{(2)}(\tau) = 1 \tag{5.94}$$

meaning that the probability of a delayed coincidence is independent of time. This state is second-order coherent. For a field in a single-mode thermal state (all other modes filtered out) given by Eq. (2.138) it can be shown that

$$g^{(2)}(\tau) = 2 \tag{5.95}$$

indicating a higher probability of detecting coincident photons. For a multimode (unfiltered) thermal state it can be shown [7] that, just as in the classical case,

$$g^{(2)}(\tau) = 1 + \left|g^{(1)}(\tau)\right|^2 \tag{5.96}$$

which lies in the range $1 \leq g^{(2)}(\tau) \leq 2$. For collision broadened light with a Lorentzian spectrum and a first-order coherence function

$$g^{(1)}(\tau) = e^{-i\omega_0 \tau - |\tau|/\tau_0} \tag{5.97}$$

(see Appendix B) we have

$$g^{(2)}(\tau) = 1 + e^{-2|\tau|/\tau_0} \tag{5.98}$$

which is just as in the classical case. But here it is legitimate to interpret the result in terms of the arrival of photons. For $|\tau| \ll \tau_0$, the probability of getting two photon counts within the time $|\tau|$ is large compared with the random case. For zero time delay, $g^{(2)}(0) = 2$, and $g^{(2)}(\tau) < g^{(2)}(0)$. This inequality characterizes photon bunching. For a multimode coherent state, using the definition of Eq. (5.93), it can be shown that

$$g^{(2)}(\tau) = 1 \tag{5.99}$$

and thus the photons arrive randomly as per the Poisson distribution, $g^{(2)}(\tau)$ being independent of the delay time. But there is another possibility, the case where $g^{(2)}(0) < g^{(2)}(\tau)$. This is the opposite of photon bunching, *photon antibunching*

[8]. For this, the photons tend to arrive evenly spaced in time, the probability of obtaining coincident photons in a time interval τ is less than for a coherent state (the random case). As we will show in Chapter 7, this situation is quite nonclassical in the sense that apparent negative probabilities are involved, meaningless for classical fields. But for now let us consider the single-mode field in a number state $|n\rangle$ from which it follows that

$$g^{(2)}(\tau) = g^{(2)}(0) = \begin{cases} 0 & n = 0, 1 \\ 1 - \dfrac{1}{n} & n \geq 2 \end{cases}. \tag{5.100}$$

Evidently $g^{(2)}(0) < 1$ and this is outside the allowed range for its classical counterpart $\gamma^{(2)}(0)$ as discussed earlier. The fact that $g^{(2)}(0)$ takes on classically forbidden values may be interpreted as a quantum-mechanical violation of the Cauchy inequality. Note that $g^{(2)}(0)$ will be less than unity whenever $\langle(\Delta\hat{n})^2\rangle < \langle\hat{n}\rangle$, according to Eq. (5.94). (For a number state $\langle(\Delta\hat{n})^2\rangle = 0$.) States for which this condition holds are *sub-Poissonian*. (States that possess sub-Poissonian statistics are also nonclassical, as we shall discuss in Chapter 7.) Since $g^{(2)}(\tau)$ is constant for the single-mode field, photon antibunching does not occur, the requirement *for* it to occur being $g^{(2)}(0) < g^{(2)}(\tau)$. The point is that photon antibunching and sub-Poissonian statistics are different effects although they have often been confused as being essentially the same thing. They are not [9].

The extension of the concept of quantum coherence to the nth order is straightforward. The nth-order quantum correlation function is given by

$$G^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1) = \mathrm{Tr}\left\{\hat{\rho}\hat{E}^{(-)}(x_1)\ldots\hat{E}^{(-)}(x_n)\hat{E}^{(+)}(x_n)\ldots\hat{E}^{(+)}(x_1)\right\} \tag{5.101}$$

and the nth-order coherence function is then defined as

$$g^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1) = \frac{G^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1)}{G^{(1)}(x_1, x_1)\ldots G^{(1)}(x_n, x_n)}. \tag{5.102}$$

Because $G^{(n)}$ contains counting rates (intensities and coincidence rates) which are always positive, it follows that

$$G^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1) \geq 0. \tag{5.103}$$

Generalizing on our definition of second-order coherence, a field is said to be nth-order coherent if

$$\left|g^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1)\right| = 1 \tag{5.104}$$

for all $n \geq 1$. If Eq. (5.103) holds for $n \to \infty$, the state is said to be fully coherent. The necessary and sufficient condition for Eq. (5.103) to hold is that the correlation function be factorable, i.e.

$$G^{(n)}(x_1, \ldots x_n; x_n, \ldots x_1) = G^{(1)}(x_1, x_1)\ldots G^{(1)}(x_n, x_n), \tag{5.105}$$

a condition that automatically holds for coherent states.

Problems

1. Derive an expression for the interference pattern in a Young's double-slit experiment for an incident field n-photon state; i.e. verify Eq. (5.66).
2. Derive an expression for the interference pattern in a Young's double-slit experiment for an incident thermal light beam.
3. Show that thermal light is first-order coherent but not second- or higher-order coherent.
4. Consider the superposition state of the vacuum and one-photon states,

$$|\psi\rangle = C_0 |0\rangle + C_1 |1\rangle ,$$

where $|C_0|^2 + |C_1|^2 = 1$, and investigate its coherence properties. Note that quantum mechanically it is "coherent" because it is a pure state. Compare your result with that obtained for the mixture

$$\rho = |C_0|^2 |0\rangle \langle 0| + |C_1|^2 |1\rangle \langle 1| .$$

5. Assuming $|\alpha|$ large, discuss the coherence properties of the superposition of two coherent states (sometimes called the Schrödinger cat states)

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\alpha\rangle + |-\alpha\rangle)$$

and compare with those of the mixture

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

References

[1] G. R. Fowles, *Introduction to Modern Optics*, 2nd edition (Mineola: Dover, 1989), p. 58; F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*, 2nd edition (Englewood Cliffs: Prentice Hall, 1993), p. 247.

[2] R. J. Glauber, *Phys. Rev.*, **130** (1963), 2529; *Phys. Rev.*, **131** (1963), 2766; U. M. Titulaer and R. J. Glauber, *Phys. Rev.*, **140** (165), B676. See also C. L. Mehta and E. C. G. Sudarshan, *Phys. Rev*., **138** (1965), B274.

[3] D. F. Walls, *Am. J. Phys*., **45** (1977), 952.

[4] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th edition (Oxford: Oxford University Press, 1958), p. 9.

[5] G. Magyar and L. Mandel, *Nature*, **198** (1963), 255.

[6] R. Hanbury Brown and R. Q. Twiss, *Nature*, **177** (1956), 27. For a complete and historical account of intensity interferometry, see R. Hanbury Brown, *The Intensity Interferometer* (London: Taylor and Francis, 1974).

[7] See R. Loudon, *The Quantum Theory of Light*, 3rd edition (Oxford: Oxford University Press, 2000), chapter 3.

[8] See R. Loudon, *Phys. Bull*., **27** (1976), 21, reprinted in L. Allen and P. L. Knight, *Concepts of Quantum Optics* (Oxford: Pergamon, 1983), p. 174; and the review article, D. F. Walls, *Nature*, **280** (1979), 451.

[9] X. T. Zou and L. Mandel, *Phys. Rev. A*, **41** (1990), 475.

Bibliography

A standard source of information on classical coherence theory, and classical optics in general is:

M. Born and E. Wolf, *Principles of Optics*, 7th edition (Cambridge: Cambridge University Press, 1999).

Another useful book with a good chapter on classical coherence theory is:

S. G. Lipson, H. Lipson, and D. S. Tannhauser, *Optical Physics* 3rd edition (Cambridge: Cambridge University Press, 1995).

A book discussing both classical and quantum coherence is the tome:

L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press, 1995).

Some other books with good discussion of quantum optical coherence theory are the following.

M. Weissbluth, *Photon–Atom Interactions* (New York: Academic Press, 1989). Our discussion of quantum coherence is fairly close to the one presented in this book.

R. Loudon, *The Quantum Theory of Light*, 3rd edition (Oxford: Oxford University Press, 2000).

D. F. Walls and G. J. Milburn, *Quantum Optics*, 2nd edition (Berlin: Springer, 1994).

Though dated, the following classic review article is recommended.

L. Mandel and E. Wolf, Coherence Properties of Optical Fields, *Rev. Mod. Phys.*, **37** (1965), 231.

Chapter 6
Beam splitters and interferometers

6.1 Experiments with single photons

Central to the entire discipline of quantum optics, as should be evident from the preceding chapters, is the concept of the photon. Yet it is perhaps worthwhile to pause and ask: what is the evidence for the existence of photons? Most of us first encounter the photon concept in the context of the photo-electric effect. As we showed in Chapter 5, the photo-electric effect is, in fact, used to indirectly detect the presence of photons – the photo-electrons being the entities counted. But it turns out that some aspects of the photo-electric effect can be explained without introducing the concept of the photon. In fact, one can go quite far with a semiclassical theory in which only the atoms are quantized with the field treated classically. But we hasten to say that, for a satisfactory explanation of *all* aspects of the photo-electric effect, the field *must* be quantized. As it happens, the other venerable "proof" of the existence of photons, the Compton effect, can also be explained without quantized fields.

In an attempt to obtain quantum effects with light, Taylor, in 1909 [1], obtained interference fringes in an experiment with an extremely weak source of light. His source was a gas flame and the emitted light was attenuated by means of screens made of smoked glass. The "double slit" in the experiment was, in fact, a needle whose shadow on a screen exhibited the fringes of a diffraction pattern when exposed to direct light from the source. But Taylor found that the fringes persisted upon attenuation of the source, even down to the lowest intensities where, one could naively conclude, on the basis of simple energy considerations, there was at most only one photon at a time between the source and the screen. Apparently, photons passing by the needle one at a time give rise to interference. Presumably, this is the origin of Dirac's famous remark [2] that "each photon interferes only with itself, interference between two photons does not occur". But we now know, as discussed in Chapter 5, that a thermal source, such as the gas flame used by Taylor, does not produce photons one-at-a-time but rather produces them in bunches. Hence it is naïve and wrong to use energy considerations alone to determine the number of photons between the source and the screen at any given time; there is a strong likelihood that there are two photons present, the

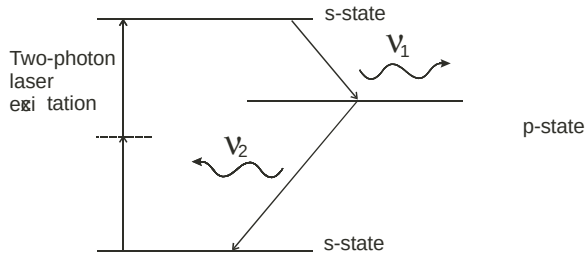


Fig. 6.1. Energy-level diagram of the single photon source for the experiments of Grangier *et al.* A calcium atom is irradiated by a laser, which excites the atom by two-photon absorption to a high-lying s-state. The atom then undergoes a cascade decay first to a p-state, emitting a photon of frequency ν_1, used as the trigger photon, and then to the original s-state emitting a photon of frequency ν_2.

photon bunching effect. A laser source produces photons randomly, so even when attenuated there is at least some chance that there may be more than one photon between the source and the screen. To get as close as possible to having a single photon between source and screen requires a source of antibunched photons. This in turn requires a source consisting of only a very few atoms, ideally a single atom.

Such a source was developed only relatively recently by Grangier *et al.* [3] originally for the purpose of a fundamental test of quantum mechanics, namely a search for violations of Bell's inequality, and then used to demonstrate the indivisibility of photons. This source consists of a beam of calcium atoms irradiated by laser light exciting the atoms to a high-lying s-state. The s-state undergoes a rapid decay to a p-state emitting a photon of frequency ν_1. Subsequently the atom rapidly undergoes another rapid decay, this time to the ground s-state, by emitting a second photon, this one having a frequency ν_2 (see Fig. 6.1). The photons, to conserve momentum, are emitted in opposite directions. In the experiment described in Reference [3], the first photon, detected by D_{trig}, was used as a "trigger" to alert a set of photo-detectors placed at the outputs of a 50:50 beam splitter upon which the second photon falls as illustrated in Fig. 6.2. The trigger tells the photo-detectors to expect a photon to emerge from the beam splitter by "gating" the detection electronics for a brief time interval. This eliminates spurious counts due to photons entering the detectors from irrelevant sources. The experimental setup, as pictured in Fig. 6.2, is such that only the particle nature of the photons will be manifested. That is, a single photon falling on the beam splitter would be either reflected into detector D_{ref} or transmitted into detector D_{tran}, i.e. it is a "which path" experiment and no interference effects are expected. There should be no simultaneous counts (the counts should anti-correlated) of reflected and transmitted photons and, because the beam splitter is 50:50, repeated runs of the experiment should result in each of the two detectors firing approximately 50% of the time. These expectations were confirmed by the investigators.

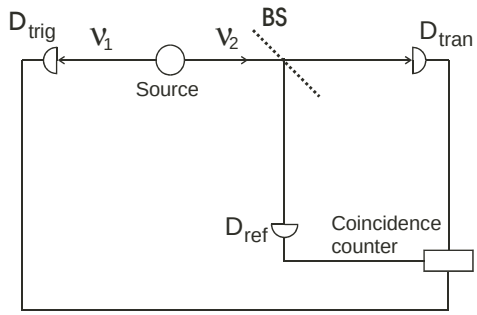


Fig. 6.2. Anti-correlation experiment of Grangier *et al.* Detection of the trigger photon alerts the coincidence counters to expect a photon to pass through the beam splitter. The beam splitter is 50:50.

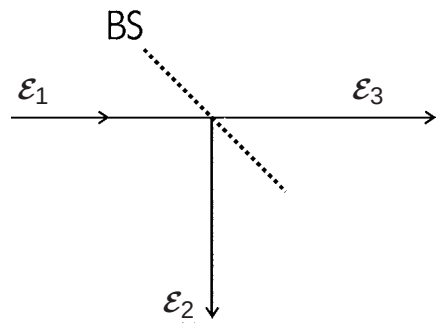


Fig. 6.3. Classical beam splitting. A classical field of amplitude $\mathcal{E}_1$ is split into fields of amplitudes $\mathcal{E}_2$ and $\mathcal{E}_3$.

6.2 Quantum mechanics of beam splitters

At this point we must pause and consider the beam splitter in a fully quantum-mechanical context. In the previous chapter, we used the notion of a beam splitter in a somewhat cavalier manner. We were able to get away with it because, for "classical"-like light beams, coherent and thermal beams, the quantum and classical treatments of beam splitters agree. But at the level of a single or few photons, the classical approach to beam splitting produces erroneous and quite misleading results.

To see how classical reasoning over beam splitting goes wrong, let us consider first a classical light field of complex amplitude $\mathcal{E}_1$ incident upon a lossless beam splitter as indicated in Fig. 6.3. $\mathcal{E}_2$ and $\mathcal{E}_3$ are the amplitudes of the reflected and transmitted beams respectively. If r and t are the (complex) reflectance and transmittance respectively of the beam splitter, then it follows that

$$\mathcal{E}_2 = r\mathcal{E}_1 \quad \text{and} \quad \mathcal{E}_3 = t\mathcal{E}_1. \tag{6.1}$$

For a 50:50 beam splitter we would have $|r| = |t| = 1/\sqrt{2}$. However, for the sake of generality, we do not impose this condition here. Since the beam splitter is assumed lossless, the intensity of the input beam should equal the sum of the intensities of the two output beams:

$$|\mathcal{E}_1|^2 = |\mathcal{E}_2|^2 + |\mathcal{E}_3|^2 \tag{6.2}$$

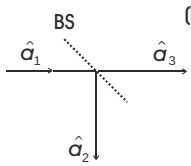


Fig. 6.4. Naïve, and incorrect, quantum mechanical depiction of a beam splitter.

which requires that

$$|r|^2 + |t|^2 = 1. \tag{6.3}$$

To treat the beam splitter quantum mechanically we might try replacing the classical complex field amplitudes $\mathcal{E}_i$ by a set of annihilation operators $\hat{a}_i$ ($i = 1, 2, 3$) as indicated in Fig. 6.4. In analogy with the classical case we might try setting

$$\hat{a}_2 = r\hat{a}_1 \quad \text{and} \quad \hat{a}_3 = t\hat{a}_1. \tag{6.4}$$

However, the operators of each of the fields are supposed to satisfy the commutation relations

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij}, \quad [\hat{a}_i, \hat{a}_j] = 0 = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] \quad (i, j = 1, 2, 3), \tag{6.5}$$

but it is easy to see that for the operators of Eq. (6.4) we obtain

$$\begin{aligned} \left[\hat{a}_2, \hat{a}_2^\dagger\right] &= |r|^2 \left[\hat{a}_1, \hat{a}_1^\dagger\right] = |r|^2, \\ \left[\hat{a}_3, \hat{a}_3^\dagger\right] &= |t|^2 \left[\hat{a}_1, \hat{a}_1^\dagger\right] = |t|^2, \\ \left[\hat{a}_2, \hat{a}_3^\dagger\right] &= rt^* \neq 0, \text{ etc.} \end{aligned} \tag{6.6}$$

Thus the transformations in Eq. (6.4) do not preserve the commutation relations and therefore cannot provide the correct quantum description of a beam splitter. This conundrum is resolved as follows: in the classical picture of the beam splitter there is an unused "port" which, being empty of an input field, has no effect on the output beams. However, in the quantum-mechanical picture, the "unused" port still contains a quantized field mode albeit in the vacuum state and, as we have repeatedly seen, the fluctuations of the vacuum lead to important physical effects. The situation with the beam splitter is no exception. In Fig. 6.5 we indicate all the modes required for a proper quantum description of the beam splitter, $\hat{a}_0$ representing the field operator of the classically vacant input mode. Also indicated are two sets of transmittances and reflectances, allowing for the possibility of an asymmetric beam splitter. We now write the beam-splitter transformations for the field operators as

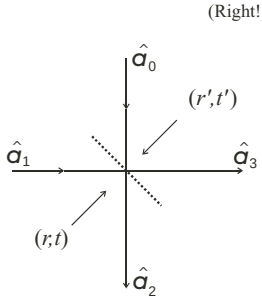


Fig. 6.5. The correct quantum-mechanical depiction of a beam splitter.

$$\hat{a}_2 = r\hat{a}_1 + t'\hat{a}_0, \quad \hat{a}_3 = t\hat{a}_1 + r'\hat{a}_0 \tag{6.7}$$

or collectively as

$$\begin{pmatrix} \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} t' & r \\ r' & t \end{pmatrix} \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \end{pmatrix}. \tag{6.8}$$

It is easily seen that the commutation relations of Eq. (6.5) are satisfied as long as the following relations hold:

$$\left|r'\right| = |r|, |t| = \left|t'\right|, |r|^2 + |t|^2 = 1, \; r^*t' + r't'^* = 0, \text{ and } r^*t + r't'^* = 0. \tag{6.9}$$

These relations are known as the reciprocity relations and can also be derived on the basis of energy conservation.

Let us examine a couple of relevant examples. The phase shifts of the reflected and transmitted beams depend on the construction of the beam splitter [4]. If the beam splitter is constructed as a single dielectric layer, the reflected and transmitted beams will differ in phase by a factor of $\exp(\pm\, i\pi/2) = \pm\, i$. For a 50:50 beam splitter, assuming the reflected beam suffers a $\pi/2$ phase shift, the input and output modes are related according to

$$\hat{a}_2 = \frac{1}{\sqrt{2}}\,(\hat{a}_0 + i\hat{a}_1)\,, \quad \hat{a}_3 = \frac{1}{\sqrt{2}}\,(i\hat{a}_0 + \hat{a}_1)\,. \tag{6.10}$$

Since the transformation between input and output modes must be unitary, we may write Eq. (6.8) as

$$\begin{pmatrix} \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \hat{U}^\dagger \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \end{pmatrix} \hat{U} \tag{6.11}$$

where $\hat{U}$ is a unitary operator. This transformation constitutes a Heisenberg picture formulation of a beam splitter. For the specific transformation represented by Eq. (6.10), the operator $\hat{U}$ has the form

$$\hat{U} = \exp\left[i\,\frac{\pi}{4}\left(\hat{a}_0^\dagger \hat{a}_1 + \hat{a}_0 \hat{a}_1^\dagger\right)\right], \tag{6.12}$$

easily checked using the Baker–Hausdorf lemma in Eq. (6.11).

On the other hand, we may adopt the Schrödinger picture and ask the following question. For a given input state to the beam splitter, what is the output state? Remembering that all photon number states $|n\rangle$, hence any superposition or any statistical mixture of such states, may be constructed by the action of n powers of the creation operator on the vacuum, we may use Eqs. (6.7) and (6.8) to construct the output states from the action of the transformed creation operators on the vacuum states of the output modes, it being obvious that an input vacuum transforms to an output vacuum: $|0\rangle_0|0\rangle_1 \to |0\rangle_2|0\rangle_3$.

As an example, consider the single photon input state $|0\rangle_0|1\rangle_1$ which we may write as $\hat{a}_1^\dagger|0\rangle_0|0\rangle_1$. For the beam splitter described by Eq. (6.10) we find that $\hat{a}_1^\dagger = (i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger)/\sqrt{2}$. Thus we may write, using that $|0\rangle_0|0\rangle_1 \xrightarrow{\text{BS}} |0\rangle_2|0\rangle_3$,

$$\begin{aligned} |0\rangle_0\,|1\rangle_1 &\xrightarrow{\text{BS}} \frac{1}{\sqrt{2}}\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{1}{\sqrt{2}}\,(i\,|1\rangle_2\,|0\rangle_3 + |0\rangle_2\,|1\rangle_3)\,. \end{aligned} \tag{6.13}$$

This is an important result. It says that a single-photon incident at one of the input ports of the beam splitter, the other port containing only the vacuum, will be either transmitted or reflected with equal probability. Of course, this is precisely as we earlier claimed and explains why no coincident counts are to be expected with photon counters placed at the outputs of the beam splitter, as confirmed by the experiment of Grangier *et al.* [3]. Actually, because the beam splitter is well understood, the lack of coincident counts in the above experiment may be taken as an indication that the source is truly producing single-photon states. (Obviously,

the beam splitter is a *passive* device that neither creates nor destroys photons. There are, of course, *active* devices that convert one photon into two, for example, and we shall encounter these in the next chapter.)

One other point needs to be made about the output state of Eq. (6.13). It is an *entangled* state: it cannot be written as a simple product of states of the individual modes 2 and 3. The density operator (see Appendix A) for the (pure) state of Eq. (6.13) is

$$\begin{aligned}\hat{\rho}_{23} = \frac{1}{2}\{&|1\rangle_2|0\rangle_{32}\langle 1|_3\langle 0| + |0\rangle_2|1\rangle_{32}\langle 0|_3\langle 1| \\ &+ i\,|1\rangle_2|0\rangle_{32}\langle 0|_3\langle 1| - i\,|0\rangle_2|1\rangle_{32}\langle 1|_3\langle 0|\}\,. \end{aligned} \tag{6.14}$$

In placing detectors in the two output beams, we are measuring the full "coherence" as described by the state vector of Eq. (6.13) or equivalently the density operator of Eq. (6.14). Suppose, on the other hand, we make no measurement of, say, mode 3. Mode 2 is then described by the reduced density operator obtained by tracing over the states of the unmeasured mode (see Appendix A):

$$\begin{aligned}\hat{\rho}_2 = \mathrm{Tr}_3\hat{\rho}_{23} &= \sum_{n=0}^{\infty}{}_3\langle n|\hat{\rho}_{23}|n\rangle_3 \\ &= \frac{1}{2}(|0\rangle_{22}\langle 0| + |1\rangle_{22}\langle 1|)\,. \end{aligned} \tag{6.15}$$

This represents merely a statistical mixture, there being no "off-diagonal" coherence terms of the form $|0\rangle\langle 1|$ or $|1\rangle\langle 0|$. Thus placing a detector in only one of the output beams yields random results, 0 or 1, each 50% of the time, just as we would expect.

Before moving on to single-photon interference, let us consider two more examples of beam splitting. First we consider a coherent state, a classical-like state, rather the opposite of the highly nonclassical single-photon state, incident on the beam splitter with, again, only the vacuum in the other input port. That is, the initial state is $|0\rangle_0|\alpha\rangle_1 = \hat{D}_1(\alpha)|0\rangle_0|0\rangle_1$ where $\hat{D}_1(\alpha) = \exp(\alpha\hat{a}_1^\dagger - \alpha^*\hat{a}_1)$ is the displacement operator for mode 1. We may then, following the procedure above, obtain the output state according to

$$\begin{aligned}|0\rangle_0|\alpha\rangle_1 &\xrightarrow{\text{BS}} \exp\left[\frac{\alpha}{\sqrt{2}}\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right) - \frac{\alpha^*}{\sqrt{2}}(-i\,\hat{a}_2 + \hat{a}_3)\right]|0\rangle_2|0\rangle_3 \\ &= \exp\left[\left(\frac{i\,\alpha}{\sqrt{2}}\right)\hat{a}_2^\dagger - \left(\frac{-i\,\alpha^*}{\sqrt{2}}\right)\hat{a}_2\right]\exp\left[\left(\frac{\alpha}{\sqrt{2}}\right)\hat{a}_3^\dagger - \left(\frac{\alpha^*}{\sqrt{2}}\right)\hat{a}_3\right]|0\rangle_2|0\rangle_3 \\ &= \left|\frac{i\,\alpha}{\sqrt{2}}\right\rangle_2\left|\frac{\alpha}{\sqrt{2}}\right\rangle_3. \end{aligned} \tag{6.16}$$

Evidently we obtain the result expected for a classical light wave where the incident intensity is evenly divided between the two output beams, e.g. half the incident average photon number, $|\alpha|^2/2$, emerges in each beam. We also naturally obtain the phase shift $i = e^{i\pi/2}$ for the reflected wave, as expected. Finally, note that the output is not entangled.

Essentially everything about beam splitting with a coherent state is classical and in that regard it is worth inserting here a note of caution. The case of a single photon incident on the beam splitter is *not* obtainable as the limiting case of an incident coherent state, i.e. for $|\alpha|$ small. It is easy to check that the single-photon result of Eq. (6.13) is not obtainable in any way as a limiting case of Eq. (6.16). This is quite obvious without doing any calculations as the former is entangled whereas the latter is not. Entanglement cannot arise as a limiting case of a product state. The point is, as mentioned earlier in the chapter, that attempts at extrapolating low-field behavior from classical results are misleading and quite wrong.

As a last example with operator transformations, we return to the strictly quantum domain and consider the situation where single photons are simultaneously injected into the two input ports of our 50:50 beam splitter, the incident state being $|1\rangle_0|1\rangle_1 = \hat{a}_0^\dagger \hat{a}_1^\dagger |0\rangle_0|0\rangle_1$. Again, following the previous procedure with $\hat{a}_0^\dagger = (\hat{a}_2^\dagger + i\hat{a}_3^\dagger)/\sqrt{2}$ and $\hat{a}_1^\dagger = (i\hat{a}_2^\dagger + \hat{a}_3^\dagger)/\sqrt{2}$ we have

$$\begin{aligned} |1\rangle_0\,|1\rangle_1 &\xrightarrow{\text{BS}} \frac{1}{2}\left(\hat{a}_2^\dagger + i\,\hat{a}_3^\dagger\right)\left(i\,\hat{a}_2^\dagger + \hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{i}{2}\left(\hat{a}_2^\dagger\hat{a}_2^\dagger + \hat{a}_3^\dagger\hat{a}_3^\dagger\right)|0\rangle_2\,|0\rangle_3 \\ &= \frac{i}{\sqrt{2}}\left(|2\rangle_2\,|0\rangle_3 + |0\rangle_2\,|2\rangle_3\right). \end{aligned} \tag{6.17}$$

Apparently, the two photons emerge together such that photo-detectors placed in the output beams should not register simultaneous counts. But unlike the case of a single incident photon, the physical basis for obtaining no simultaneous counts is not a result of the particle-like nature of photons. Rather, it is caused by interference (a wave-like effect) between two possible ways of obtaining the (absent) output state $|1\rangle_2|1\rangle_3$: the process where both photons are transmitted (Fig. 6.6(a)) and the process where they are both reflected (Fig. 6.6(b)). Note the indistinguishability of the two processes for the output state $|1\rangle_2|1\rangle_3$. There is a simple and rather intuitive way of understanding this result. Recall Feynman's rule [5] for obtaining the probability for an outcome that can occur by several indistinguishable processes: one simply adds the probability *amplitudes* of all the processes and then calculates the square of the modulus. Assuming that our beam splitter is described by Eq. (6.10), the reflected photons each acquire an $e^{i\pi/2} = i$ phase shift. The amplitude for transmission for each photon is $A_\mathrm{T} = 1/\sqrt{2}$ and the amplitude for reflection for each is $A_\mathrm{R} = i/\sqrt{2}$. The amplitude that *both* photons are transmitted is $A_\mathrm{T} \cdot A_\mathrm{T}$ and that both are reflected is $A_\mathrm{R} \cdot A_\mathrm{R}$. Thus the probability of the photons emerging in both output beams is

$$P_{11} = |A_\mathrm{T} \cdot A_\mathrm{T} + A_\mathrm{R} \cdot A_\mathrm{R}|^2 = \left|\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\frac{i}{\sqrt{2}}\right|^2 = 0. \tag{6.18}$$

An experimental demonstration of this effect was first performed by Hong, Ou, and Mandel [6] and is discussed in Chapter 9.

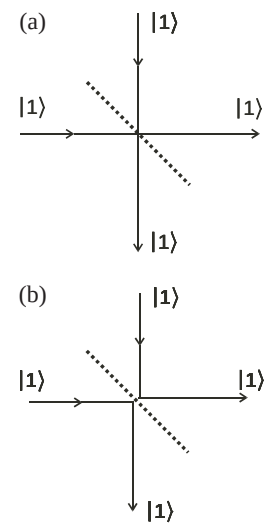


Fig. 6.6. Two indistinguishable processes with simultaneous single photon inputs. (a) Both photons are transmitted. (b) Both photons are reflected. These processes interfere destructively with each other.

It may be attempting to interpret the result in Eq. (6.17) as owing to the bosonic nature of photons, a kind of clustering in the sense of Bose–Einstein condensation (BEC). Indeed, in the case of fermions, such as in neutron interferometry, the output of a beam splitter for the corresponding input would find the fermions always in different beams in accordance with the Pauli exclusion principle. Of course, this behavior and that of the photons are linked to the statistical properties of the particles and hence to the nature of the operators that describe them. But for photons it would be misleading to push the clustering analogy à la BEC too far. For example, if n photons are injected into each port simultaneously, i.e. the input state being $|n\rangle_0|n\rangle_1$, the output state is *not* $\sim |2n\rangle_2|0\rangle_3 + e^{i\phi}|0\rangle_2|2n\rangle_3$, except for $n = 1$, as the reader may easily check.

We now reconsider the case of the vacuum and coherent state incident to a beam splitter as given in Eq. (6.16). We expand the product state $|0\rangle_0|\alpha\rangle_1$ as

$$|\text{in}\rangle = |0\rangle_0\,|\alpha\rangle_1 = e^{-|\alpha|^2/2}\sum_{N=0}^{\infty}\frac{\alpha^N}{\sqrt{N!}}\,|0\rangle_0\,|N\rangle_1 \tag{6.19}$$

where it is evident that we have a superposition of the product states $|0\rangle_0|N\rangle_1$. Because the beam splitter conserves the total number of photons, we should be able to take the expansion for the product state in the last line of Eq. (6.16) and by partitioning it into sectors of total photon numbers N, deduce the beam-splitter transformation for the input states $|0\rangle_0|N\rangle_1$. We write

$$|\text{out}\rangle = \left|\frac{i\,\alpha}{\sqrt{2}}\right\rangle_2\left|\frac{\alpha}{\sqrt{2}}\right\rangle_3 = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(i\,\alpha/\sqrt{2})^n(\alpha/\sqrt{2})^m}{\sqrt{n!m!}}\,|n\rangle_2\,|m\rangle_3, \tag{6.20}$$

and setting $m = N - n$ we have, after some rearrangement,

$$|\text{out}\rangle = e^{-|\alpha|^2/2}\sum_{N=0}^{\infty}\frac{\alpha^N}{\sqrt{N!}}\,|\psi_N\rangle_{2,3}, \tag{6.21}$$

where

$$|\psi_N\rangle_{2,3} = \frac{1}{2^{N/2}}\sum_{n=0}^{N} i^n\left[\frac{N!}{n!\,(N-n)!}\right]^{1/2}|n\rangle_2\,|N-n\rangle_3. \tag{6.22}$$

Thus it follows that the beam splitter causes the transformation $|0\rangle_0|N\rangle_1 \to |\psi_N\rangle_{2,3}$. The state $|\psi_N\rangle_{2,3}$ is a binomial state in that N photons are binomially distributed over the two modes.

There is one other way to obtain this result and that is by using projection operators (see Appendix B). (We stress that only the mathematics of the projection operators is used here; no state reductive measurements are implied.) If we let $\hat{P}_N^{(0,1)} = |0\rangle_0|N\rangle_1\,{}_1\langle 0|_1\langle N|$ be the projection operator onto the N photon sector of the 0–1 modes, the corresponding projection operator onto the 2–3 modes is

$$\hat{P}_N^{(2,3)} = \sum_{k=0}^{N}|k\rangle_2\,|N-k\rangle_3\,{}_2\langle k|_3\,\langle N-k|. \tag{6.23}$$

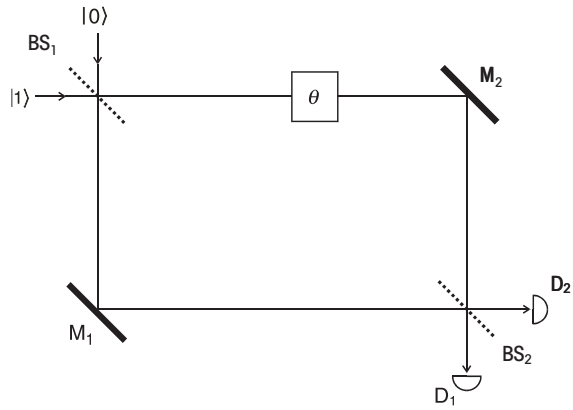


Fig. 6.7. A Mach–Zehnder interferometer with a single-photon input. The beam splitters BS_1 and BS_2 are 50:50, M_1 and M_2 are mirrors, and the box labeled θ represents the relative phase shift between the two arms.

Evidently, we have

$$|0\rangle_0\,|N\rangle_1 = \frac{\hat{P}_N^{(0,1)}\,|\text{in}\rangle}{\langle\text{in}|\,\hat{P}_N^{(0,1)}\,|\text{in}\rangle^{1/2}}, \tag{6.24}$$

and so it is that

$$|\psi_N\rangle_{2,3} = \frac{\hat{P}_N^{(2,3)}\,|\text{out}\rangle}{\langle\text{out}|\,\hat{P}_N^{(2,3)}\,|\text{out}\rangle^{1/2}}. \tag{6.25}$$

Similar considerations will allow us to obtain the output corresponding to the input state $|N\rangle_0|N\rangle_1$ starting with $|\text{in}\rangle = |\alpha\rangle_0|\alpha\rangle_1$. The demonstration of this is be left as an exercise (see Problem 8).

6.3 Interferometry with a single photon

We return now to the single-photon experiment of Grangier *et al*. [3]. As we have said, these investigators used an atomic source to produce single photons as verified by the lack of coincident photon counts with detectors placed at the output ports of a beam splitter. Of course, such a setup addresses only the particle nature of the photons and no interference is exhibited. To obtain interference effects with single photons we must scramble the pathways to the detectors as interference occurs in the lack of "which path" information. A convenient way to accomplish this is to construct a Mach–Zehnder interferometer (MZI), which consists of two beam splitters and a set of mirrors as illustrated in Fig. 6.7. Interference occurs because the detectors D_1 and D_2 cannot distinguish between the photon taking the clockwise or the counter-clockwise paths. In the counter-clockwise path is placed a phase shifter, which could simply be a length of optical fiber, so that the lengths of the paths about the MZI are different and adjustable. This phase shift may be represented by the unitary operator $\exp(i\theta\hat{n})$ where $\hat{n}$ is to be understood as the number operator of the field in that section of the interferometer. The angle θ in reality represents a relative phase shift between the two paths.

To see how single-photon interference arises in the MZI, we start with the input state $|0\rangle|1\rangle$ where we shall adopt the convention that the states propagating along the counter-clockwise path precede those along the clockwise, as indicated in Fig. 6.7. Assuming that the beam splitters are described by Eqs. (6.10), the first beam splitter (BS_1) transforms the input as

$$|0\rangle\,|1\rangle \xrightarrow{\mathrm{BS}_1} \frac{1}{\sqrt{2}}\left(|0\rangle\,|1\rangle + i\,|1\rangle\,|0\rangle\right). \tag{6.26}$$

The mirrors contribute a factor of $e^{i\,\pi/2}$ to each term, amounting to an irrelevant phase, which we omit. The phase shifter in the clockwise path causes a phase change in the first component;

$$\frac{1}{\sqrt{2}}(|0\rangle|1\rangle + i|1\rangle|0\rangle) \xrightarrow{\theta} \frac{1}{\sqrt{2}}(e^{i\,\theta}|0\rangle|1\rangle + i|1\rangle|0\rangle). \tag{6.27}$$

At the second beam splitter (BS_2) we have the following transformations

$$\begin{aligned} |0\rangle|1\rangle &\xrightarrow{\mathrm{BS}_2} \frac{1}{\sqrt{2}}(|0\rangle|1\rangle + i|1\rangle|0\rangle), \\ |1\rangle|0\rangle &\xrightarrow{\mathrm{BS}_2} \frac{1}{\sqrt{2}}(|1\rangle|0\rangle + i|0\rangle|1\rangle), \end{aligned} \tag{6.28}$$

where (mindful of the propagation direction convention) on the right-hand side the first state in each of the product states refers to the beam directed toward detector D_1 while the second is toward D_2. Thus the transformation on the total state due to the second beam splitter is

$$\frac{1}{\sqrt{2}}(e^{i\,\theta}|0\rangle|1\rangle + i|1\rangle|0\rangle) \xrightarrow{\mathrm{BS}_2} \frac{1}{2}[(e^{i\,\theta}-1)|0\rangle|1\rangle + i(e^{i\,\theta}+1)|1\rangle|0\rangle]. \tag{6.29}$$

The probability that the state $|0\rangle|1\rangle$ is detected (that only D_2 fires) is

$$P_{01} = \frac{1}{2}\,(1-\cos\theta) \tag{6.30}$$

whereas the probability that the state $|1\rangle|0\rangle$ is detected (that only D_1 fires) is

$$P_{10} = \frac{1}{2}\,(1+\cos\theta)\,. \tag{6.31}$$

Obviously, as the path length, hence θ, is changed, these probabilities exhibit oscillations, which, of course, are the interference fringes indicative of single-photon interference. It is precisely these fringes that have been observed in the experiments of Grangier *et al*. [3].

6.4 Interaction-free measurement

We now wish to take advantage of the machinery developed so far to expose what surely must be one of the most peculiar features of quantum mechanics, namely the capability of detecting the presence of an object without scattering any quanta (in this case photons) off it. Conventional wisdom would have it that in order to detect the presence of an object one would minimally be required to detect one scattered photon from it. But recently, Elitzur and Vaidman [7] showed

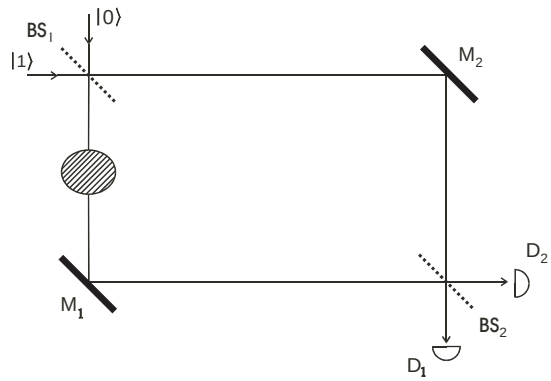


Fig. 6.8. Experimental set for interaction free-measurement. A Mach–Zehnder interferometer contains an object, the shaded ellipse, in one arm.

theoretically, and Kwiat *et al*. [8] showed experimentally, that the conventional wisdom is not quite right.

To set the stage for our presentation of this counter-intuitive effect, let us reconsider the MZI of Fig. 6.7 and choose the path lengths in both arms to be the same; that is, we choose $\theta = 0$. Under this condition only detector D_1 fires. D_2 does not fire due to destructive interference between the two paths of the interferometer. Suppose now an object is placed in one of the arms of the MZI as shown in Fig. 6.8. The presence of the object, in this case in the counter-clockwise arm, changes everything. Imagine now that we place a detector, or for that matter a number of detectors, in the vicinity of the object such that we would be able to detect a photon scattered from it. Upon detection of such a photon we know two things: that the photon took the counter-clockwise path after the first beam splitter (that is, we gain "which path" information) and that there is an object in that path. But it turns out that it is not even necessary to detect a scattered photon to determine the presence of the object. Suppose there are no extra detectors near the object and we confine our attention to the detectors at the outputs of the MZI. If, for any given run of the experiment, neither of these detectors fires then we again know the path of the photon and that it had to be scattered by the object. Recall that after the first beam splitter there is a probability of $1/2$ that the photon is in the counter-clockwise path and thus a probability of $1/2$ that the photon scatters from the object, hence in half of the runs of the experiment neither detector fires. But there is a probability of $1/2$ that the photon takes the clockwise path. Because there is no longer an open path to cause interference, at the second beam splitter the photon has a 50:50 chance of being transmitted to detector D_1 or reflected to detector D_2. So in *these* cases where detectors do fire, D_1 fires half the time and D_2 the other half. So then, for the interferometer as a *whole*, there is a probability of $1/2$ that neither detector fires, a probability of $1/4$ each that D_1 or D_2 fire. Remember, the MZI was initially configured such that with no object in either path, *only* D_1 would fire, 100% of the time. So whenever D_2 fires we *know* that there must be an object present in one arm of the MZI. The fact that D_2 fires at all (and it will fire 25% of the time) may be taken as a

detection of the object in one of the arms of the interferometer even though no photon has been scattered by it.

This bizarre result is ultimately a manifestation of a rather general feature of quantum mechanics known as nonlocality – the apparently instantaneous effects of certain kinds of influences due to interactions and measurements. Another example of nonlocal effects is the Aharonov–Bohm effect, in which an electron wave function accumulates phase due to a distant and confined magnetic flux that never directly interacts with the electron [9], and the violation of various forms of Bell's inequalities. Bell's inequalities will be discussed in Chapter 9.

6.5 Interferometry with coherent states of light

Having examined interferometry at the level of a single photon where quantum effects are expected to be pre-eminent, we now examine interferometry at the other extreme, where we assume an input coherent state. The result should be essentially the classical one.

To this end we envision the MZI of Fig. 6.7 but where the mode previously containing a single photon now contains a coherent state $|\alpha\rangle$. Following the previous convention, and using Eq. (6.16), the first beam splitter performs the transformation

$$|0\rangle\,|\alpha\rangle \xrightarrow{\mathrm{BS}_1} \left|\frac{i\alpha}{\sqrt{2}}\right\rangle \left|\frac{\alpha}{\sqrt{2}}\right\rangle. \tag{6.32}$$

Ignoring the common phase shifts from the mirrors, the effect of the phase shifter in the clockwise arm is

$$\left|\frac{i\,\alpha}{\sqrt{2}}\right\rangle \left|\frac{\alpha}{\sqrt{2}}\right\rangle \xrightarrow{\theta} \left|\frac{i e^{i\,\theta}\alpha}{\sqrt{2}}\right\rangle \left|\frac{\alpha}{\sqrt{2}}\right\rangle. \tag{6.33}$$

Lastly BS_2 performs the transformation

$$\left|\frac{i e^{i\,\theta}\alpha}{\sqrt{2}}\right\rangle \left|\frac{\alpha}{\sqrt{2}}\right\rangle \xrightarrow{\mathrm{BS}_2} \left|\frac{i(e^{i\,\theta}+1)\alpha}{\sqrt{2}}\right\rangle \left|\frac{(e^{i\,\theta}-1)\alpha}{\sqrt{2}}\right\rangle. \tag{6.34}$$

In a typical phase-shift measurement with an MZI, the shift θ is determined by subtracting the intensities obtained from the detectors D_1 and D_2. This difference is proportional to the number difference operator of the two modes. Let the operators $(\hat{a},\ \hat{a}^\dagger)$ represent the beam directed to detector D_1 and $(\hat{b},\ \hat{b}^\dagger)$ of that toward D_2. The number difference operator we then designate as $\hat{O} = \hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}$. From the right-hand side of Eq. (6.34) we obtain $\langle\hat{O}\rangle = |\alpha|^2\cos\theta$. This is essentially the result obtained by using coherent classical light waves. But it is not the end of the story.

We must never forget that coherent states, even though they have many classical-like properties, are fundamentally quantum mechanical and thus carry with them quantum fluctuations, as we have seen earlier. These fluctuations impose limits on the accuracy of the phase measurements. By using the

calculus of error propagation, the uncertainty of the phase measurements is given by

$$\Delta\theta = \Delta O \Big/ \left| \frac{\partial \langle \hat{O} \rangle}{\partial \theta} \right| \tag{6.35}$$

where $\Delta O = \sqrt{\langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2}$. With the coherent state as input the phase uncertainty is found to be

$$\Delta\theta = \frac{1}{\sqrt{\bar{n}} |\sin\theta|} \tag{6.36}$$

where $\bar{n} = |\alpha|^2$. Clearly the phase uncertainty depends on the relative phase θ and evidently we obtain the optimal, or minimum, phase uncertainty $\Delta\theta_{\min} = 1/\sqrt{\bar{n}}$ for θ equal to odd multiples of $\pi/2$. This is the best we can do with classical-like states.

However, on fundamental grounds it should be possible to exceed this limitation, often known as the standard quantum limit, by the use of manifestly nonclassical states of the field to do interferometry. In fact, the fundamental limit on the phase uncertainty, known as the Heisenberg limit, appears to be $\Delta\theta_{\rm H} = 1/\bar{n}$. Reaching this limit has been an important goal in quantum-optics research over the past two decades, and continues to the present, particularly in connection with attempts to detect weak signals such as those expected from gravitational waves passing through the Earth. In fact, these signals are expected to be so weak that they would ordinarily be masked by the vacuum noise of the detector. The current generation of gravity-wave detectors, such as those associated with the LIGO and VIRGO projects which involve large-scale interferometers, could benefit from the use of nonclassical states to enhance the sensitivity of phase measurements. Nonclassical states will be the subject of the next chapter. Improved sensitivity using *entangled* states in interferometers is discussed in Chapter 11.

Problems

1. Show that the beam splitter transformation equations of Eq. (6.10) do indeed follow from Eq. (6.11).
2. Consider the operator $\hat{U}(\theta) = \exp[i\,\theta \hat{J}_1]$ where $\hat{J}_1$ is defined as the operator $\hat{J}_1 \equiv (\hat{a}_0^\dagger \hat{a}_1 + \hat{a}_1^\dagger \hat{a}_0)/2$. Obviously for $\theta = \pi/2$ we obtain the transformation for the 50:50 beam splitter of Eqs. (6.10) and (6.11). For the more general case where the beam splitter is *not* 50:50, obtain the corresponding transformation of the mode operators. Relate the angle θ to the parameters r, t, r', t'.
3. Consider a beam splitter constructed in such a way that its operator is described by the transformation $\hat{U}(\theta) = \exp[i\,\theta \hat{J}_2]$, where $\hat{J}_2 \equiv (\hat{a}_0^\dagger \hat{a}_1 - \hat{a}_1^\dagger \hat{a}_0)/2i$. Obtain the transformation of the mode operators for this beam splitter for the general case and for the case of a 50:50 beam splitter.

4. Defining the operator $\hat{J}_3 \equiv (\hat{a}_0^\dagger \hat{a}_0 - \hat{a}_1^\dagger \hat{a}_1)/2$, show that the operators $\hat{J}_1$, $\hat{J}_2$, and $\hat{J}_3$ satisfy the angular momentum commutation relations $[\hat{J}_i, \hat{J}_j] = i\varepsilon_{ijk}\hat{J}_k$, where ε_{ijk} is the usual completely anti-symmetric tensor. Further, show that the operator $\hat{J}_0 \equiv (\hat{a}_0^\dagger \hat{a}_0 + \hat{a}_1^\dagger \hat{a}_1)/2$, commutes with all the operators $\hat{J}_1$, $\hat{J}_2$, and $\hat{J}_3$ and with the square of the "angular momentum" $\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2 = \hat{J}_0(\hat{J}_0 + 1)$. Comment on the association of beam splitter transformations and rotations in fictitious three dimensional space.
5. Suppose that the input to a beam splitter of the type described by the unitary operator $\hat{U}(\theta) = \exp[i\,\theta \hat{J}_1]$ is the state $|\text{in}\rangle = |0\rangle_0|N\rangle_1$. Show that the output state is given by
$$|\text{out}\rangle = [1 + \tan^2(\theta/2)]^{-N/2} \sum_{k=0}^{N} \binom{N}{k}^{1/2} [i\tan(\theta/2)]^k |k\rangle_2 |N-k\rangle_3.$$
Examine the joint probability $P_{n,m}$ for finding n photons in mode 2 and m in mode 3 and show that the nonzero elements form a binomial distribution. Is the state $|\text{out}\rangle$ an entangled state? If so, how does the entanglement change with the angle θ?
6. Suppose the input state to a 50:50 $\hat{J}_1$-type beam splitter is $|\text{in}\rangle = |\alpha\rangle_0|\beta\rangle_1$. Obtain the output state. Is it entangled?
7. Consider the input state $|\text{in}\rangle = |N\rangle_0|N\rangle_1$ to a $\hat{J}_1$-type beam splitter. Argue, without actually performing the transformations, that there will be no odd-numbered photon states in either of the output modes. (Hint: consider the number of ways that each set of odd-numbered states can occur and add up the amplitudes.) Show that the projection operator associated with modes 2 and 3 for the above input state containing a total of $2N$ photons is $\hat{P}_{2N}^{(2,3)} = \sum_{k=0}^{N} |2k\rangle_2 |2N-2k\rangle_3\, {}_2\langle 2k|_3\langle 2N-2k|$.
8. For a 50:50 beam splitter of the type described by the operator $\hat{J}_1$, show that with the input state $|\text{in}\rangle = |N\rangle_0|N\rangle_1$, the output state is given by
$$|\text{out}\rangle = \sum_{k=0}^{N} \left[\binom{2k}{k} \binom{2N-2k}{N-k} \left(\frac{1}{2}\right)^{2N} \right]^{1/2} |2k\rangle_2\, |2N-2k\rangle_3.$$
Examine the joint photon probability $P_{n,m}$ for finding n photons in mode 2 and m photons in mode 3. In this case, the nonzero elements form a distribution known as the arcsine distribution [10].
9. For the beam splitter of the previous question, suppose that $|\text{in}\rangle = |0\rangle_0(|\alpha\rangle_1 + |-\alpha\rangle_1)/\sqrt{2}$, where $|\alpha|$ is assumed to be large enough so that $\langle -\alpha \mid \alpha\rangle \simeq 0$. Obtain the output state and examine the question of its entanglement.
10. Consider the Mach–Zehnder interferometer of Fig. 6.7 but now with the input state at the first beam splitter identical to that of the previous problem. Obtain the expectation value of the number difference operator $\hat{O} = \hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}$ and examine the uncertainty in the measurement of the phase shift θ for this case of input state. Compare the result obtained using the coherent state in one of the input modes.
11. The description of interaction-free measurements given in Section 6.4 assumed that the beam splitters involved are 50:50. The presence of an object in one arm is determined when one of the detectors, the one initially not firing because of interference, with no object in one of the arms, begins to fire. It will fire in one-fourth of the runs. Examine

the question: can the presence of an object in one arm be detected more efficiently (i.e. in a greater number of experimental runs) by using beam splitters that are not 50:50?

References

[1] G. I. Taylor, *Proc. Camb. Philos. Soc*., **15** (1909), 114.
[2] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th edition (Oxford: Oxford University Press, 1958), p. 9.
[3] P. Grangier, G. Roger and A. Aspect, *Europhys. Lett*., **1** (1986), 173.
[4] See M. W. Hamilton, *Am. J. Phys*., **68** (2000), 186.
[5] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics* (Reading: Addison-Wesley, 1965), Vol. III, chapter 3.
[6] C. K. Hong, Z. Y. Ou and L. Mandel, *Phys. Rev. Lett*., **59** (1987), 2044.
[7] A. C. Elitzur and L. Vaidman, *Found. Phys*., **23** (1993), 987. See also, R. H. Dicke, *Am. J. Phys*., **49** (1981), 925.
[8] P. Kwiat, H. Weinfurter, T. Herzog and A. Zielinger, *Phys. Rev. Lett*., **74** (1995), 4763.
[9] Y. Aharanov and D. Bohm, *Phys. Rev*., **115** (1959), 484.
[10] See W. Feller, *An Introduction to the Probability Theory and its Applications*, 3rd edition (New York: Wiley, 1968), Vol. I, p. 79.

Bibliography

The following are papers on the quantum description of beam splitters.

A. Zeilinger, *Am. J. Phys*., **49** (1981), 882.
S. Prasad, M. O. Scully and W. Martienssen, *Opt. Commun*., **62** (1987), 139.
Z. Y. Ou, C. K. Hong and L. Mandel, *Opt. Commun*., **63** (1987), 118.
Z. Y. Ou and L. Mandel, *Am. J. Phys*., **57** (1989), 66.

Though not explicitly discussed here in the text (but see the Problems above), there is a very useful formulation of the description of beam splitters in terms of the angular-momentum operators, or in the language of group theory, the unitary group SU(2). *The following papers discuss this formulation*.

B. Yurke, S. L. McCall and J. R. Klauder, *Phys. Rev. A*, **33** (1986), 4033.
R. A. Campos and B. A. E. Saleh, *Phys. Rev. A*, **40** (1989), 1371.
U. Leonhardt, *Phys. Rev. A*, **48** (1993), 3265.

Chapter 7
Nonclassical light

"The word 'classical' means only one thing in science: it's wrong!" [1]

We have previously emphasized the fact that *all* states of light are quantum mechanical and are thus nonclassical, deriving some quantum features from the discreteness of the photons. Of course, in practice, the nonclassical features of light are difficult to observe. (We shall use "quantum mechanical" and "nonclassical" more or less interchangeably here.) Already we have discussed what must certainly be the most nonclassical of all nonclassical states of light – the single-photon state. Yet, as we shall see, it is possible to have nonclassical states involving a very large number of photons. But we need a criterion for nonclassicality. Recall that in Chapter 5 we discussed such a criterion in terms of the quasi-probability distribution known as the P function, $P(\alpha)$. States for which $P(\alpha)$ is positive everywhere or no more singular than a delta function, are classical whereas those for which $P(\alpha)$ is negative or more singular than a delta function are nonclassical. We have shown, in fact, that $P(\alpha)$ for a coherent state *is* a delta function, and Hillery [2] has shown that *all* other pure states of the field will have functions $P(\alpha)$ that are negative in some regions of phase space and are more singular than a delta. It is evident that the variety of possible nonclassical states of the field is quite large.

In this chapter, we shall discuss some of the most important examples of nonclassical states. We begin with the squeezed states, i.e. the quadrature and the number squeezed states (the latter also known as states with sub-Poissonian statistics), then discuss again the nonclassical nature of photon antibunching, introduce the Schrödinger cat states, the states of correlated double beams, higher-order squeezing, and finally, broadband squeezed light.

7.1 Quadrature squeezing

If two operators $\hat{A}$ and $\hat{B}$ satisfy the commutation relation $[\hat{A},\ \hat{B}] = i\hat{C}$, it follows that

$$\left\langle (\Delta\hat{A})^2 \right\rangle \left\langle (\Delta\hat{B})^2 \right\rangle \geq \frac{1}{4}|\langle \hat{C} \rangle|^2. \tag{7.1}$$

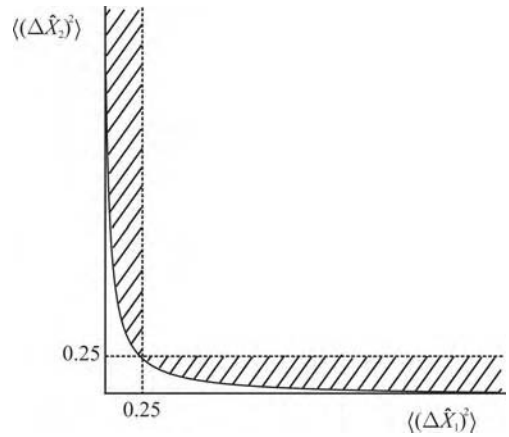


Fig. 7.1. A graphical characterization of quadature squeezing. Squeezing exists for the points inside the shaded areas. The solid-line border is a hyperbola determined from the equalized uncertainty relation $\left\langle(\Delta\hat{X}_1)^2\right\rangle \left\langle(\Delta\hat{X}_2)^2\right\rangle = 1/16$.

A state of the system is said to be squeezed if either

$$\left\langle(\Delta\hat{A})^2\right\rangle < \frac{1}{2}|\langle\hat{C}\rangle| \quad \text{or} \quad \left\langle(\Delta\hat{B})^2\right\rangle < \frac{1}{2}|\langle\hat{C}\rangle|. \tag{7.2}$$

(Because of Eq.(7.1), we obviously cannot have both variances less than $|\langle\hat{C}\rangle|/2$ simultaneously.) Squeezed states for which the $[\hat{X}_1, \hat{X}_2] = i/2$ equality holds in Eq. (7.1) are sometimes known as ideal squeezed states and are an example of the "intelligent" states we discussed in Chapter 3.

In the case of quadrature squeezing, we take $\hat{A} = \hat{X}_1$ and $\hat{B} = \hat{X}_2$, with$\hat{X}_1$ and $\hat{X}_2$ being the quadrature operators of Eqs. (2.52) and (2.53) satisfying Eq. 2.55 and thus $\hat{C} = 1/2$. From Eq. (2.56) it follows that quadrature squeezing exists whenever [3]

$$\left\langle(\Delta\hat{X}_1)^2\right\rangle < \frac{1}{4} \quad \text{or} \quad \left\langle(\Delta\hat{X}_2)^2\right\rangle < \frac{1}{4}. \tag{7.3}$$

We have already established that, for a coherent state $|\alpha\rangle$, the equality in Eq. (2.56) holds and that the variances of two quadratures are equal: $\langle(\Delta\hat{X}_1)^2\rangle = \langle(\Delta\hat{X}_2)^2\rangle = 1/4$. Not only that, but the result for the coherent state is exactly the same as for the vacuum (see Eq. (2.59)). States for which one of the conditions in Eq. (7.3) holds will have less "noise" in one of the quadratures than for a coherent state or a vacuum state – the fluctuations in that quadrature are squeezed. Of course, the fluctuations in the other quadrature must be enhanced so as to not violate the uncertainty relation. There are squeezed states for which the uncertainty relation is equalized, but this need not be the case in general. See Fig. 7.1 for a graphical representation of the range of squeezing.

Before presenting specific examples of squeezed states, we wish to demonstrate why quadrature squeezing must be considered a nonclassical effect. To this end, we express the relevant expectation values in terms of the P function. Using

Eq. (3.108) it is easy to show that

$$\left\langle (\Delta \hat{X}_1)^2 \right\rangle = \frac{1}{4}\left\{1 + \int P(\alpha)[(\alpha + \alpha^*) - (\langle \hat{a} \rangle + \langle \hat{a}^\dagger \rangle)]^2 d^2\alpha \right\} \tag{7.4}$$

and

$$\left\langle (\Delta \hat{X}_2)^2 \right\rangle = \frac{1}{4}\left\{1 + \int P(\alpha)[(\alpha - \alpha^*)/i - (\langle \hat{a} \rangle - \langle \hat{a}^\dagger \rangle)/i]^2 d^2\alpha \right\} \tag{7.5}$$

where

$$\langle \hat{a} \rangle = \int P(\alpha)\alpha \, d^2\alpha \quad \text{and} \quad \langle \hat{a}^\dagger \rangle = \int P(\alpha)\alpha^* d^2\alpha. \tag{7.6}$$

Because the term inside the square brackets is, of course, always positive, it is evident that the condition $\langle (\Delta \hat{X}_{1,2})^2 \rangle < 1/4$ requires that $P(\alpha)$ be nonpositive at least in some regions of phase space.

It is sometimes convenient to introduce a generic quadrature operator

$$\hat{X}(\vartheta) = \frac{1}{2}(\hat{a}\, e^{-i\,\vartheta} + \hat{a}^\dagger e^{i\,\vartheta}) \tag{7.7}$$

where obviously $\hat{X}(0) = \hat{X}_1$ and $\hat{X}(\pi/2) = \hat{X}_2$. To characterize squeezing, we introduce the parameter

$$s(\vartheta) = \frac{\left\langle (\Delta \hat{X}(\vartheta))^2 \right\rangle - 1/4}{1/4} = 4\left\langle (\Delta \hat{X}(\vartheta))^2 \right\rangle - 1. \tag{7.8}$$

Squeezing exists for some value of the angle ϑ whenever $-1 \leq s(\vartheta) < 0$. Alternatively, we can introduce the normally ordered variance $\langle : (\Delta \hat{X}(\vartheta))^2 : \rangle$ such that squeezing exists whenever $-\frac{1}{4} \leq \langle : (\Delta \hat{X}(\vartheta))^2 : \rangle < 0$. In terms of the normally ordered variance we have

$$s(\vartheta) = 4\left\langle : (\Delta \hat{X}(\vartheta))^2 : \right\rangle. \tag{7.9}$$

So, how can squeezed states be generated? One way of generating a squeezed state mathematically, in a manner corresponding to a physical process that we discuss below, is through the action of a "squeeze" operator defined as

$$\hat{S}(\xi) = \exp\left[\frac{1}{2}(\xi^* a^2 - \xi a^{\dagger 2})\right], \tag{7.10}$$

where $\xi = r\, e^{i\,\theta}$, and where r is known as the squeeze parameter and $0 \leq r < \infty$ and $0 \leq \theta \leq 2\pi$. This operator $\hat{S}(\xi)$ is a kind of two-photon generalization of the displacement operator used to define the usual coherent states of a single-mode field, as discussed in Chapter 3. Evidently, the operator $\hat{S}(\xi)$ acting on the vacuum would create some sort of "two-photon coherent state" as it is clear that photons will be created or destroyed in pairs by the action of this operator. To get some

sense of what happens as the result of acting with this operator, let us consider the state

$$|\psi_s\rangle = \hat{S}(\xi)\,|\psi\rangle \tag{7.11}$$

where $|\psi\rangle$ is for the moment arbitrary and $|\psi_s\rangle$ denotes the state generated by the action of $\hat{S}(\xi)$ on $|\psi\rangle$. To obtain the variances of $\hat{X}_1$ and $\hat{X}_2$ we need the expectation values of $\hat{a}$, $\hat{a}^2$ etc. To this end we need the following results obtained from application of the Baker–Hausdorf lemma:

$$\begin{aligned}\hat{S}^\dagger(\xi)\,\hat{a}\,\hat{S}(\xi) &= \hat{a}\cosh r - \hat{a}^\dagger e^{i\theta}\sinh r\\ \hat{S}^\dagger(\xi)\,\hat{a}^\dagger\,\hat{S}(\xi) &= \hat{a}^\dagger\cosh r - \hat{a}e^{-i\theta}\sinh r\end{aligned} \tag{7.12}$$

where $\hat{S}^\dagger(\xi) = \hat{S}(-\xi)$. Thus we have

$$\langle\psi_s|\,\hat{a}\,|\psi_s\rangle = \langle\psi|\,\hat{S}^\dagger(\xi)\,\hat{a}\,S(\xi)\,|\psi\rangle \tag{7.13}$$

and

$$\langle\psi_s|\,\hat{a}^2\,|\psi_s\rangle = \langle\psi|\,\hat{S}^\dagger(\xi)\,\hat{a}S(\xi)\,\hat{S}^\dagger(\xi)\,\hat{a}S(\xi)\,|\psi\rangle\,, \tag{7.14}$$

etc. For the special case where $|\psi\rangle$ is the vacuum state $|0\rangle$, $|\psi_s\rangle$ is the squeezed vacuum state, which we denote as $|\xi\rangle$:

$$|\xi\rangle = \hat{S}(\xi)\,|0\rangle\,. \tag{7.15}$$

Using Eqs. (7.12)–(7.15) we find that for the squeezed vacuum state

$$\left\langle(\Delta\hat{X}_1)^2\right\rangle = \frac{1}{4}[\cosh^2 r + \sinh^2 r - 2\sinh r\cosh r\cos\theta] \tag{7.16}$$

$$\left\langle(\Delta\hat{X}_2)^2\right\rangle = \frac{1}{4}[\cosh^2 r + \sinh^2 r + 2\sinh r\cosh r\cos\theta]. \tag{7.17}$$

For $\theta = 0$, these reduce to

$$\begin{aligned}\left\langle(\Delta\hat{X}_1)^2\right\rangle &= \frac{1}{4}e^{-2r},\\ \left\langle(\Delta\hat{X}_2)^2\right\rangle &= \frac{1}{4}e^{2r},\end{aligned} \tag{7.18}$$

and evidently squeezing exists in the $\hat{X}_1$ quadrature. For $\theta = \pi$, the squeezing will appear in the $\hat{X}_2$ quadrature. Note that the product of the uncertainties yields $1/16$ and thus for $\theta = 0$ or $\pi/2$ the squeezed vacuum states equalize the uncertainty relation of Eq. (2.56). Squeezed states need not, and generally do not, equalize the uncertainty relation.

There is a simple way to represent squeezing graphically. Recall that in Figs. 3.1 and 3.2 we presented phase-space representations of the noise associated with a coherent state (3.1) and a vacuum state (3.2) where in both cases the fluctuations of the quadrature operators were equal, $\Delta X_1 = \Delta X_2 = 1/2$. A representation of the squeezed vacuum state for $\theta = 0$, where the fluctuations in X_1 are reduced, is given in Fig. 7.2, whereas for $\theta = \pi$, see Fig. 7.3.

In the cases with $\theta = 0$ or π, the squeezing is along either $\hat{X}_1$ or $\hat{X}_2$. For other values of θ, let us define the rotated quadrature operators

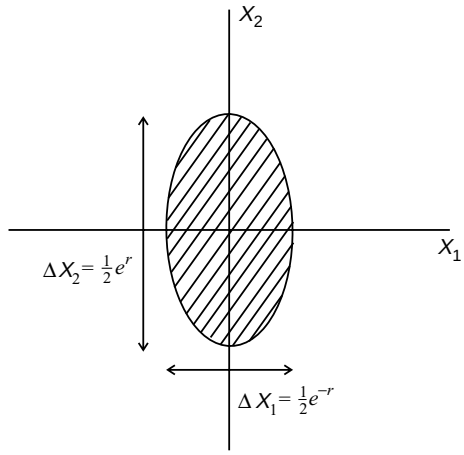


Fig. 7.2. Error ellipse for a squeezed vacuum state where the squeezing is in the X_1 quadrature.

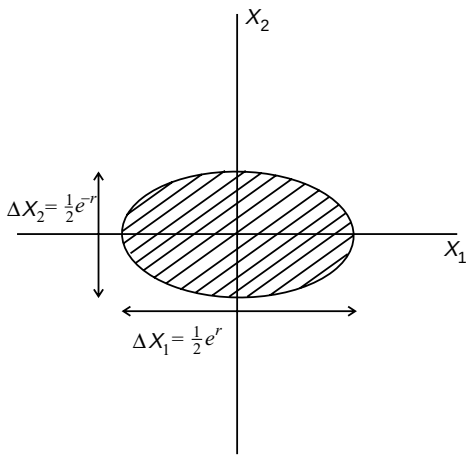


Fig. 7.3. Error ellipse for a squeezed vacuum state where the squeezing is in the X_2 quadrature.

$\hat{Y}_1$ and $\hat{Y}_2$ as

$$\begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 \end{pmatrix} = \begin{pmatrix} \cos\theta/2 & \sin\theta/2 \\ -\sin\theta/2 & \cos\theta/2 \end{pmatrix} \begin{pmatrix} \hat{X}_1 \\ \hat{X}_2 \end{pmatrix} \tag{7.19}$$

or equivalently

$$\hat{Y}_1 + i\,\hat{Y}_2 = (\hat{X}_1 + i\,\hat{X}_2)e^{-i\,\theta/2}. \tag{7.20}$$

For the squeezed vacuum state of Eq. (7.12) it can be shown that

$$\begin{aligned} \left\langle (\Delta\hat{Y}_1)^2 \right\rangle &= \frac{1}{4}e^{-2r} \\ \left\langle (\Delta\hat{Y}_2)^2 \right\rangle &= \frac{1}{4}e^{2r} \end{aligned} \tag{7.21}$$

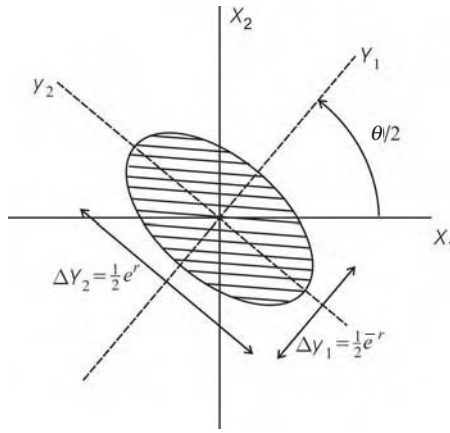


Fig. 7.4. Rotated error ellipse for the squeezed vacuum. The squeezing is along the $\theta/2$ direction.

which is represented graphically in Fig. 7.4. The important point here is that the squeezing need not be along X_1 or X_2.

A more general squeezed state may be obtained by applying the displacement operator to Eq. (7.12):

$$|\alpha, \xi\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle. \tag{7.22}$$

Obviously for $\xi = 0$ we obtain just a coherent state. Since the displacement operator effects the transformation

$$\begin{aligned} \hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha) &= \hat{a} + \alpha \\ \hat{D}^\dagger(\alpha)\hat{a}^+\hat{D}(\alpha) &= \hat{a}^\dagger + \alpha^*, \end{aligned} \tag{7.23}$$

the action of the product of the displacement and squeeze operators on $\hat{a}$ and $\hat{a}^\dagger$ can be obtained from Eqs. (7.12) and (7.23). Leaving the steps as an exercise, it can be shown that

$$\langle \hat{a} \rangle = \alpha, \tag{7.24}$$

which is independent of the squeeze parameter r, and that

$$\langle \hat{a}^2 \rangle = \alpha^2 - e^{i\theta} \sinh r \cosh r, \tag{7.25}$$

$$\langle \hat{a}^\dagger \hat{a} \rangle = |\alpha|^2 + \sinh^2 r. \tag{7.26}$$

Of course, the corresponding quantities for the coherent states are obtained for $r \to 0$ and for the squeezed vacuum as $\alpha \to 0$. Furthermore

$$\langle \hat{Y}_1 + i\hat{Y}_2 \rangle = \alpha e^{-i\theta/2} \tag{7.27}$$

and once again it follows that

$$\begin{aligned} \left\langle (\Delta \hat{Y}_1)^2 \right\rangle &= \frac{1}{4}e^{-2r} \\ \left\langle (\Delta \hat{Y}_2)^2 \right\rangle &= \frac{1}{4}e^{2r}. \end{aligned} \tag{7.28}$$

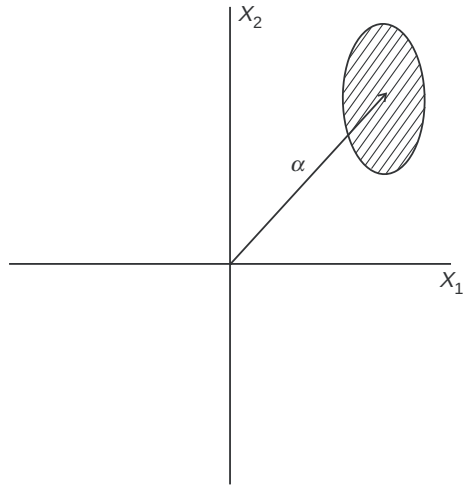


Fig. 7.5. Error ellipse of a displaced squeezed vacuum state with the squeezing in the X_1 quadrature.

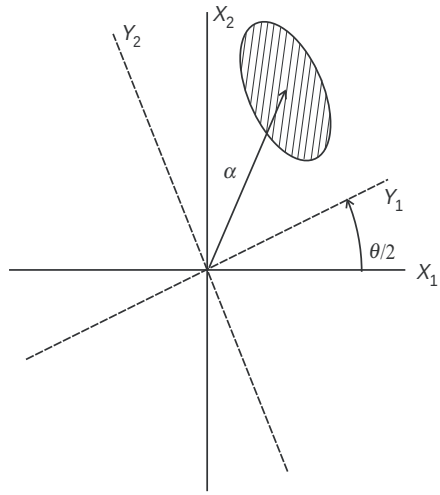


Fig. 7.6. Rotated error ellipse of a displaced squeezed vacuum state.

For the case that $\theta = 0$, the squeezed state is represented in phase space as in Fig. 7.5 where the "error ellipse" is essentially that of the squeezed vacuum displaced by α. Of course the more general case of θ will simply be a displaced rotated error ellipse as in Fig. 7.6.

Before going on, let us look at the electric field for a squeezed state. The single-mode field electric field operator of Eq. (2.15) may be written

$$\hat{E}(\chi) = \mathcal{E}_0 \sin(kz)(\hat{a}\, e^{-i\chi} + \hat{a}^\dagger e^{i\chi}) \tag{7.29}$$

where the operators $\hat{a}$ and $\hat{a}^\dagger$ are those at $t = 0$, and $\chi = \omega t$, the phase of the field, so that $\hat{E}(\chi)$ is the phase-dependent electric field. In terms of the quadrature

operators, Eq. (7.29) can be written as

$$\hat{E}(\chi) = 2\mathcal{E}_0 \sin(kz)[\hat{X}_1 \cos(\chi) + \hat{X}_2 \sin(\chi)]. \tag{7.30}$$

From the commutator $[\hat{X}_1, \hat{X}_2] = i/2$ we have

$$[\hat{E}(0), \hat{E}(\chi)] = i\mathcal{E}_0^2 \sin^2(kz) \sin\chi \tag{7.31}$$

and thus the electric field operators at different phases (or times) do not commute. From Eq. (7.28) it follows that

$$\left\langle (\Delta\hat{E}(0))^2 \right\rangle \left\langle (\Delta\hat{E}(\chi))^2 \right\rangle \geq \frac{1}{4}\left[\mathcal{E}_0^2 \sin^2(kz) \sin\chi\right]. \tag{7.32}$$

For many states of the field, the field uncertainty is independent of the phase χ (e.g. coherent states, thermal states). But if the uncertainty in the field for $\chi = 0$ is such that

$$\Delta E(0) < \frac{1}{\sqrt{2}}\mathcal{E}_0 |\sin(kz) \sin\chi| \tag{7.33}$$

then at some other phase χ we must have

$$\Delta E(\chi) > \frac{1}{\sqrt{2}}\mathcal{E}_0 |\sin(kz) \sin\chi|. \tag{7.34}$$

In this sense, squeezed light contains phase-dependent noise, reduced below that of the vacuum for some phases and enhanced above that of the vacuum for others. Quadrature squeezing, because it is phase sensitive, is related to the wave-like nature of light.

By using phasor diagrams, in Fig. 7.7 we illustrate the distribution of noise in the electric field for squeezed states. The case of the coherent state is illustrated in Fig. 3.8.

The fact that some parts of a light wave in a squeezed state are less noisy than for a field in a vacuum state has possible technological applications, particularly in the detection of weak signals. An important example is the detection of gravity waves with large-scale interferometers [4], as in the LIGO [5] and VIRGO [6] projects. Another possible application is in the area of optical communications and quantum information processing [7]. We shall not pursue these applications here.

To get a bit more insight into the squeezed states and ultimately find an expression for them in terms of the photon number states, let us start with the vacuum state $|0\rangle$ satisfying

$$\hat{a}\,|0\rangle = 0. \tag{7.35}$$

Multiplying by $\hat{S}(\xi)$ from the left and using the fact that this operator is unitary, we may write

$$\hat{S}(\xi)\,\hat{a}\,\hat{S}^\dagger(\xi)\,\hat{S}(\xi)\,|0\rangle = 0 \tag{7.36}$$

or

$$\hat{S}(\xi)\,\hat{a}\,\hat{S}^\dagger(\xi)\,|\xi\rangle = 0. \tag{7.37}$$

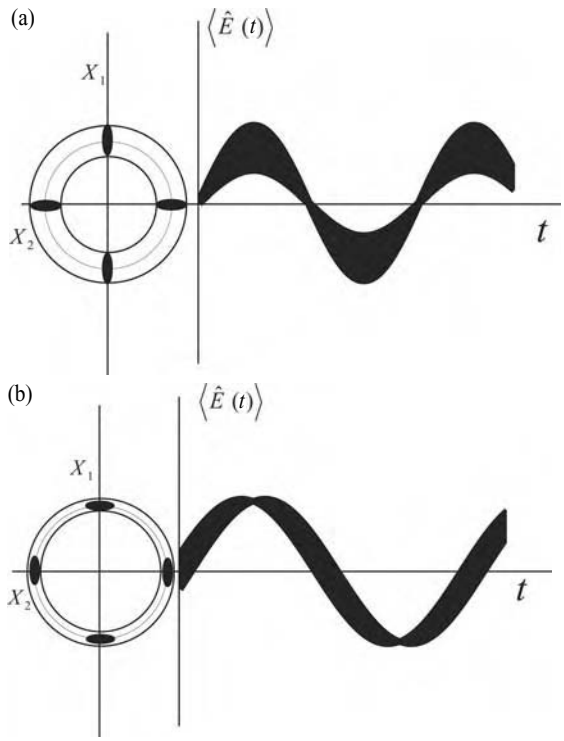


Fig. 7.7. Phasor diagram indicating the distribution of the noise for the electric field for a squeezed state for (a) squeezing initially in the X_1 quadrature, and (b) for squeezing initially in X_2.

Since

$$\hat{S}(\xi)\,\hat{a}\,\hat{S}^{\dagger}(\xi) = \hat{a}\cosh r + e^{i\theta}\hat{a}^{\dagger}\sinh r, \tag{7.38}$$

we may rewrite Eq. (7.37) as

$$\left(\hat{a}\mu + \hat{a}^{\dagger}\nu\right)|\xi\rangle = 0 \tag{7.39}$$

where $\mu = \cosh r$ and $\nu = e^{i\theta}\sinh r$. Thus the squeezed vacuum state is an eigenstate of the operator $\hat{a}\mu + \hat{a}^{+}\nu$ with eigenvalue zero. For the more general state of Eq. (7.22) we can similarly write

$$\hat{D}(\alpha)\,\hat{S}(\xi)\,\hat{a}\,\hat{S}^{\dagger}(\xi)\,\hat{D}^{\dagger}(\alpha)\,\hat{D}(\alpha)\,\hat{S}(\xi)\,|0\rangle = 0, \tag{7.40}$$

which, by using the relation

$$\hat{D}(\alpha)\hat{a}\,\hat{D}^{\dagger}(\alpha) = \hat{a} - \alpha \tag{7.41}$$

and Eq. (7.38), can be rewritten as the eigenvalue problem

$$(\hat{a}\mu + \hat{a}^{\dagger}\nu)|\alpha,\xi\rangle = \gamma\,|\alpha,\xi\rangle, \tag{7.42}$$

where

$$\gamma = \alpha \cosh r + \alpha^* e^{i\,\theta} \sinh r. \tag{7.43}$$

Obviously, for $r = 0$ this is just the eigenvalue problem for the ordinary coherent state and for $\alpha = 0$ it is for the squeezed vacuum state. Now writing

$$\hat{a} = \hat{X}_1 + i\,\hat{X}_2 = (\hat{Y}_1 + i\,\hat{Y}_2)\,e^{i\,\theta/2} \tag{7.44}$$

and substituting into Eq. (7.39) we obtain after some rearrangement

$$(\hat{Y}_1 + i\,\hat{Y}_2 e^{-2\,r})\,|\alpha, \xi\rangle = \beta|\alpha, \xi\rangle \tag{7.45}$$

where

$$\beta = \alpha\; e^{-r} e^{-i\,\theta/2} = \langle \hat{Y}_1 \rangle + i\langle \hat{Y}_2 \rangle e^{-2\,r}. \tag{7.46}$$

Equation (7.45) is precisely of the form of an eigenvalue equation for the eigenstate that equalizes the uncertainty product for the operators $\hat{Y}_1$ and $\hat{Y}_2$, which, since $[\hat{Y}_1, \hat{Y}_2] = i/2$, is just

$$\left\langle (\Delta \hat{Y}_1)^2 \right\rangle \left\langle (\Delta \hat{Y}_2)^2 \right\rangle = \frac{1}{16}. \tag{7.47}$$

Comparing Eq. (7.45) with (3.19), setting $\hat{A} = \hat{Y}_1$, $\hat{B} = \hat{Y}_2$ and $\hat{C} = 1/2$, we deduce that

$$\left\langle (\Delta \hat{Y}_2)^2 \right\rangle = \frac{1}{4} e^{2\,r} \tag{7.48}$$

and then from Eq. (7.43) that

$$\langle (\Delta \hat{Y}_1)^2 \rangle = \frac{1}{4} e^{-2\,r}. \tag{7.49}$$

Thus the state $|\alpha, \xi\rangle$ is an "intelligent" state with respect to the rotated quadrature operators $\hat{Y}_1$ and $\hat{Y}_2$.

In terms of the original quadrature operators $\hat{X}_1$ and $\hat{X}_2$, Eq. (7.39) reads

$$(\hat{X}_1 + i\lambda \hat{X}_2)|\alpha, \xi\rangle = \gamma\,|\alpha, \xi\rangle \tag{7.50}$$

where

$$\lambda = \left(\frac{\mu - \nu}{\mu + \nu}\right) \quad \text{and} \quad \gamma = \frac{\alpha}{\mu + \nu}. \tag{7.51}$$

For $\theta = 0$, Eq. (7.46) takes the form

$$(\hat{X}_1 + i\,\hat{X}_2 e^{-2\,r})|\alpha, \xi\rangle = \alpha\; e^{-r}|\alpha, \xi\rangle \tag{7.52}$$

and we obtain the results of Eqs. (7.15). But what about the case when $\theta \neq 0$? To understand the meaning of such a situation, we go back to the commutation relation $[\hat{A}, \hat{B}] = i\hat{C}$ and point out that the most general statement of the uncertainty

relation is actually not as given in Eq. (7.1) but rather it is

$$\left\langle (\Delta \hat{A})^2 \right\rangle \left\langle (\Delta \hat{B})^2 \right\rangle \geq \frac{1}{4} \left[\langle \hat{F} \rangle^2 + \langle \hat{C} \rangle^2 \right] \tag{7.53}$$

where

$$\langle \hat{F} \rangle = \langle \hat{A}\hat{B} + \hat{B}\hat{A} \rangle - 2\langle \hat{A} \rangle \langle \hat{B} \rangle \tag{7.54}$$

is the covariance, essentially the measure of the *correlations* between the observables $\hat{A}$ and $\hat{B}$. The equality in Eq. (7.53) is obeyed for a state $|\psi\rangle$ satisfying

$$\langle \hat{A} + i\lambda \hat{B} \rangle |\psi\rangle = (\langle \hat{A} \rangle + i\lambda \langle \hat{B} \rangle) |\psi\rangle \tag{7.55}$$

where λ may, in general, be complex. From Eq. (7.51) it is easy to show that

$$\left\langle (\Delta \hat{A})^2 \right\rangle - \lambda^2 \left\langle (\Delta \hat{B})^2 \right\rangle = i\lambda \left\langle \hat{F} \right\rangle \tag{7.56}$$

$$\left\langle (\Delta \hat{A})^2 \right\rangle + \lambda^2 \left\langle (\Delta \hat{B})^2 \right\rangle = \lambda \langle \hat{C} \rangle. \tag{7.57}$$

If λ is real, then from Eq. (7.56) we must have $\langle \hat{F} \rangle = 0$ and thus there are no correlations between $\hat{A}$ and $\hat{B}$. If λ is pure imaginary, from Eq. (7.57) we must have $\langle \hat{C} \rangle = 0$. But for $\hat{A} = \hat{X}_1$, $\hat{B} = \hat{X}_2$, it is not possible to have $\lambda = (\mu - \nu)/(\mu + \nu)$ pure imaginary. When it is real ($\theta = 0$) then

$$\langle \hat{F} \rangle = \langle \hat{X}_1 \hat{X}_2 + \hat{X}_2 \hat{X}_1 \rangle - 2\langle \hat{X}_1 \rangle \langle \hat{X}_2 \rangle = 0 \tag{7.58}$$

and there are no correlations between $\hat{X}_1$ and $\hat{X}_2$. But for $\theta \neq 0$ or 2π, λ is complex and it is easy to show that for this case $\langle \hat{F} \rangle \neq 0$. Thus the general squeezed state of the form $|\alpha, \xi\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle$ may be characterized by the presence of correlations between the observables $\hat{X}_1$ and $\hat{X}_2$. In the limit $\xi \to 0$ ($r \to 0$) we recover the coherent state $|\alpha\rangle$ for which there are no correlations between $\hat{X}_1$ and $\hat{X}_2$. We shall encounter other nonclassical states exhibiting correlations between the quadratures. Again, we emphasize that the squeezed state $|\alpha, \xi\rangle$ is explicitly constructed to have the property of squeezing, but many other pure states of the field may also exhibit squeezing at least for some ranges of the relevant parameters.

We now decompose the squeezed states into the photon number states in order to examine the photon statistics. We consider the squeezed vacuum state first. We write

$$|\xi\rangle = \sum_{n=0}^{\infty} C_n |n\rangle \tag{7.59}$$

which, upon substituting into Eq. (7.39), leads to the recursion relation

$$C_{n+1} = -\frac{\nu}{\mu} \left(\frac{n}{n+1} \right)^{1/2} C_{n-1}. \tag{7.60}$$

Note that this relation connects only every other photon state. In fact, there are two distinct solutions, one involving only even photon states and a second involving only odd photon states. Obviously, only the even solution contains the vacuum so we address only this case here. The solution of the recursion relation is

$$C_{2m} = (-1)^m (e^{i\theta} \tanh r)^m \left[\frac{(2m-1)!!}{(2m)!!} \right]^{1/2} C_0. \tag{7.61}$$

C_0 is determined from the normalization

$$\sum_{m=0}^{\infty} |C_{2m}|^2 = 1 \tag{7.62}$$

which leads to

$$|C_0|^2 \left(1 + \sum_{m=0}^{\infty} \frac{[\tanh r]^{2m} (2m-1)!!}{(2m)!!} \right) = 1. \tag{7.63}$$

Fortunately, there exists the mathematical identity

$$1 + \sum_{m=0}^{\infty} z^m \left(\frac{(2m-1)!!}{(2m)!!} \right) = (1-z)^{-1/2} \tag{7.64}$$

from which we easily obtain $C_0 = \sqrt{\cosh r}$. Finally we use the identities

$$(2m)!! = 2^m m! \tag{7.65}$$

$$(2m-1)!! = \frac{1}{2^m} \frac{(2m)!}{m!} \tag{7.66}$$

to obtain the most common form of the expansion coefficients for the squeezed vacuum state:

$$C_{2m} = (-1)^m \frac{\sqrt{(2m)!}}{2^m m!} \frac{(e^{i\theta} \tanh r)^m}{\sqrt{\cosh r}}. \tag{7.67}$$

Thus the squeezed vacuum state is

$$|\xi\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{m=0}^{\infty} (-1)^m \frac{\sqrt{(2m)!}}{2^m m!} e^{i\,m\,\theta} (\tanh r)^m \, |2m\rangle \,. \tag{7.68}$$

The probability of detecting $2m$ photons in the field is

$$P_{2m} = |\langle 2m \mid \xi \rangle|^2 = \frac{(2m)!}{2^{2m} (m!)^2} \frac{(\tanh r)^{2m}}{\cosh r} \tag{7.69}$$

while for detecting $2m+1$ states it is

$$P_{2m+1} = |\langle 2m+1 \mid \xi \rangle|^2 = 0. \tag{7.70}$$

Thus the photon probability distribution for a squeezed vacuum state is oscillatory, vanishing for all odd photon numbers. A typical distribution for the squeezed vacuum is given in Fig. 7.8.

Notice that, aside from the oscillatory nature of the distribution, the shape of it resembles that of thermal radiation. It must be remembered that the squeezed vacuum is a pure state whereas the thermal state is a mixed one.

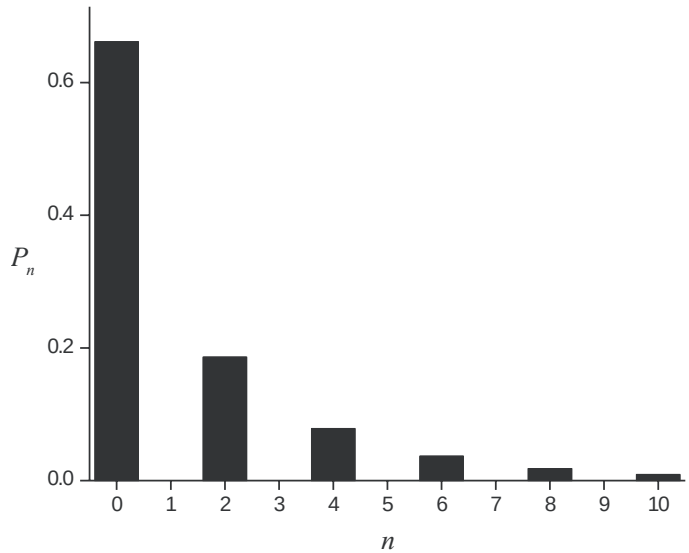


Fig. 7.8. Histogram for the photon number distribution for squeezed vacuum state.

We now seek solutions of Eq. (7.42) for the more general case for $\alpha \neq 0$. We again assume a solution of the form of Eq. (7.59), taking advantage of what we have learned in the case of $\alpha = 0$ to make the ansatz

$$C_n = \mathcal{N}(\cosh r)^{-1/2} \left[\frac{1}{2} e^{i\,\theta} \tanh r\right]^{n/2} f_n(x) \tag{7.71}$$

where $f_n(x)$ is as yet an unknown function and $\mathcal{N}$ is a normalization factor. The resulting recursion relation for $f_n(x)$ is

$$(n+1)^{1/2} f_{n+1}(x) - 2\gamma(e^{i\,\theta} \sinh(2r))^{-1/2} f_n(x) + 2n^{1/2} f_{n-1}(x) = 0, \tag{7.72}$$

where γ is given by Eq. (7.43). This is identical to the recursion relation for the Hermite polynomials $H_n(x)$:

$$H_{n+1}(x) - 2x H_n(x) + 2n H_{n-1}(x) = 0 \tag{7.73}$$

provided we make the identifications

$$f_n(x) = H_n(x)/\sqrt{n!} \text{ and } x = \gamma(e^{i\,\theta} \sinh(2r))^{-1/2}. \tag{7.74}$$

Thus we have

$$C_n = \mathcal{N}[n! \cosh r]^{-1/2} \left[\frac{1}{2} e^{i\theta} \tanh r\right]^{n/2} H_n\left[\gamma(e^{i\,\theta} \sinh(2r))^{-1/2}\right]. \tag{7.75}$$

To find $\mathcal{N}$ we set $n = 0$ in Eq. (7.75) to obtain

$$C_0 = \mathcal{N}[\cosh r]^{-1/2} \tag{7.76}$$

and note that

$$\begin{aligned} C_0 = \langle 0 \mid \alpha, \xi\rangle &= \langle 0|\hat{D}(\alpha)\hat{S}(\xi)|0\rangle \\ &= \langle 0|\hat{D}^{\dagger}(-\alpha)\hat{S}(\xi)|0\rangle \\ &= \langle -\alpha \mid \xi\rangle \end{aligned} \tag{7.77}$$

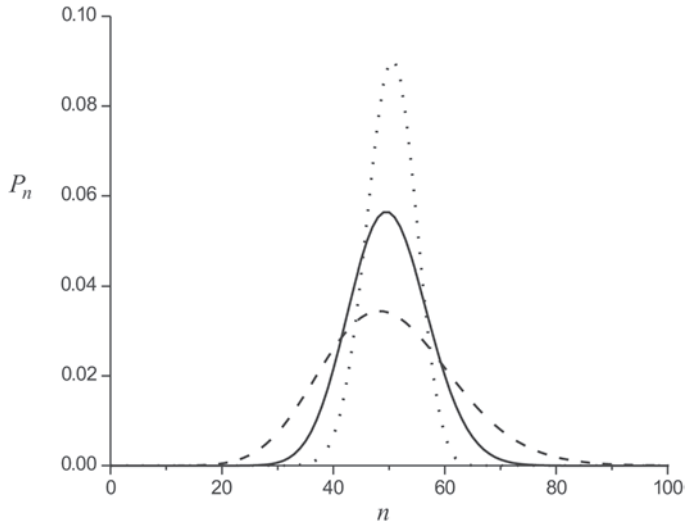


Fig. 7.9. Photon probability distributions (valid only at integers) for the coherent state with $|\alpha|^2 = 50$ (solid line) and the squeezed states for $|\alpha|^2 = 50$, $r = 0.5$, with $\psi - \theta/2 = 0$ (dotted line) and $\psi - \theta/2 = \pi/2$ (dashed line).

where the last line is the inner product of the coherent state $|-\alpha\rangle$ and the squeezed vacuum state $|\xi\rangle$. This product can be expressed as

$$\langle -\alpha \mid \xi \rangle = \exp\left(-\frac{1}{2}|\alpha|^2\right) \times \sum_{m=0}^{\infty} (\alpha^*)^{2m} [(2m)!]^{-1/2} C_{2m} \tag{7.78}$$

where the C_{2m} is given in Eq. (7.67). Thus from Eqs. (7.77)–(7.78) and using Eq. (7.67) we obtain

$$\mathcal{N} = (\cosh r)^{1/2} \langle -\alpha \mid \xi \rangle = \exp\left[-\frac{1}{2}|\alpha|^2 - \frac{1}{2}\alpha^{*2} e^{i\theta} \tanh r\right]. \tag{7.79}$$

Thus our squeezed state $|\alpha,\ \xi\rangle$ has the number state decomposition

$$|\alpha, \xi\rangle = \frac{1}{\sqrt{\cosh r}} \exp\left[-\frac{1}{2}|\alpha|^2 - \frac{1}{2}\alpha^{*2} e^{i\theta} \tanh r\right] \times \sum_{n=0}^{\infty} \frac{\left[\frac{1}{2} e^{i\theta} \tanh r\right]^{n/2}}{\sqrt{n!}} H_n\left[\gamma (e^{i\theta} \sinh(2r))^{-1/2}\right] |n\rangle\,. \tag{7.80}$$

The probability of finding n photons in the field is given by

$$\begin{aligned} P_n &= |\langle n \mid \alpha, \xi\rangle|^2 \\ &= \frac{(\frac{1}{2}\tanh r)^n}{n!\cosh r} \exp\left[-|\alpha|^2 - \frac{1}{2}(\alpha^{*2} e^{i\theta} + \alpha^2 e^{-i\theta}) \tanh r\right] \\ &\quad \times \left| H_n\left[\gamma (e^{i\theta} \sinh(2r))^{-1/2}\right]\right|^2. \end{aligned} \tag{7.81}$$

Recall that the average photon number for this state given by Eq. (7.26). If $|\alpha|^2 \gg \sinh^2 r$, we can say that the "coherent" part of the state dominates the squeezed part. But through Eq. (7.81) it is evident the distribution will depend on the phase of α. In Fig. 7.9 we show distributions for two values of $\psi - \theta/2$

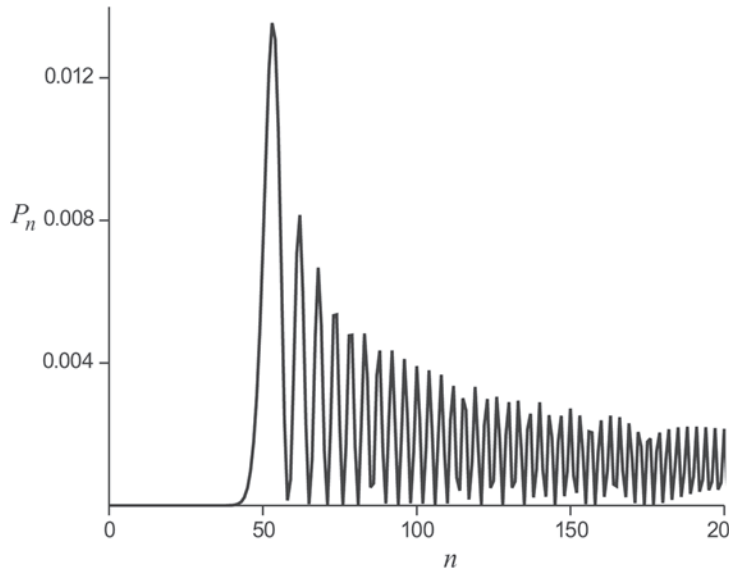


Fig. 7.10. Photon number probability distribution of a squeezed state with $\psi - \theta/2 = 0$ and with $|\alpha|^2 = 50$ and $r = 4$.

where ψ is the phase of α, i.e. $\alpha = |a|e^{i\,\psi}$, along with, for the sake of comparison, the distribution for a coherent state of average photon number $|\alpha|^2$.

Note that the distribution for $\psi - \theta/2 = 0$ is narrower than for the coherent state. This is in fact a manifestation of another form of squeezing sometimes called number squeezing, in which the photon number distribution is sub-Poissonian, i.e. narrower than the Poisson distribution for a coherent state. Like quadrature squeezing, it is a nonclassical effect. We shall return to this matter a bit later. In the case where $\psi - \theta/2 = \pi/2$, the distribution is broader than for a coherent state and is therefore called super-Poissonian. This is *not* a nonclassical effect. The case for which the squeezed part dominates, i.e. $\sinh^2 r > |\alpha|^2$, is pictured in Fig. 7.10 for $\psi - \theta/2 = 0$. Oscillations, some rather large scale, once again appear in the distribution. Oscillating photon number distributions have been interpreted by Schleich and Wheeler [8] as resulting from the interference of error contours in phase space. We will not pursue this interesting idea here but we refer the reader to the recent book by Schleich [9] which covers this and all other matters of interest in quantum optics from the point of view of phase space.

Before leaving this section, let us calculate some quasi-probability distribution functions associated with the squeezed states. We already know, of course, that the P function for a squeezed state is problematic, at least if we wish to write down something as an ordinary function. Rather, it turns out that the P function for a squeezed state, and other nonclassical states, is highly singular, containing such things as derivatives of a delta function. But the Q and Wigner functions are always well behaved. Using the coherent state $|\beta\rangle$, the Q function for the

squeezed state $|\alpha, \xi\rangle$ is given by

$$\begin{aligned} Q(\beta) &= \frac{1}{\pi}|\langle\beta \mid \alpha, \xi\rangle|^2 \\ &= \frac{1}{\pi \cosh r} \exp\{-(|\alpha|^2 + |\beta|^2) \\ &\quad + (\beta^*\alpha + \beta\alpha^*)/\cosh r \\ &\quad - \frac{1}{2}[e^{i\theta}(\beta^{*2} - \alpha^{*2}) + e^{-i\theta}(\beta^2 - \alpha^2)] \tanh r\}. \end{aligned} \tag{7.82}$$

In Fig. 7.11 we plot Q as a function $x = \mathrm{Re}(\beta)$ and $y = \mathrm{Im}(\beta)$, for $\theta = 0$ and for (a) $\alpha = 0$, and (b) $\alpha = \sqrt{5}$. For this choice of θ, the squeezing is along the X_1 direction, the narrowing of the Gaussian profile of the Q function being along the X_1 direction with a corresponding expansion along the X_2 direction. For $\alpha = 0$, the peak of the Q function is centered at $\beta = 0$, apropos a squeezed vacuum state, whereas for $\alpha = \sqrt{5}$, the peak is simply translated to $\beta = \sqrt{5}$. From the Q function we may obtain the antinormally ordered characteristic function $C_{\mathrm{A}}(\lambda)$ via the Fourier transform of Eq. (3.132). From Eq. (3.129) we obtain the Wigner characteristic function $C_{\mathrm{W}}(\lambda)$ and, finally, from Eq. (3.136) we obtain the Wigner function

$$W(\beta) = \frac{2}{\pi} \exp\left[-\frac{1}{2}x^2 e^{-2r} - \frac{1}{2}y^2 e^{2r}\right], \tag{7.83}$$

specialized to the case with $\theta = 0$. Once again we have Gaussians, narrowed in the direction of squeezing and expanded in the orthogonal direction. Note that for the squeezed state the Wigner function is non-negative. In fact, it can be shown that squeezed states of the form of $|\alpha, \xi\rangle$, including the special case of the coherent states ($\xi = 0$), are the only pure quantum states yielding non-negative Wigner functions [10].

7.2 Generation of quadrature squeezed light

Most schemes for generation of quadrature squeezed light are based on some sort of parametric process utilizing various types of nonlinear optical devices. Generally, one desires an interaction Hamiltonian quadratic in the annihilation and creation operators of the field mode to be squeezed. We consider a device known as a degenerate parametric down-converter. A certain kind of nonlinear medium is pumped by a field of frequency ω_{P} and some photons of that field are converted into pairs of identical photons, of frequency $\omega = \omega_{\mathrm{P}}/2$ each, into the "signal" field, the process known as *degenerate parametric down-conversion*. The Hamiltonian for this process is given by

$$\hat{H} = \hbar\omega\,\hat{a}^\dagger\hat{a} + \hbar\omega_{\mathrm{P}}\hat{b}^\dagger\hat{b} + i\hbar\chi^{(2)}(\hat{a}^2\hat{b}^\dagger - \hat{a}^{\dagger 2}\hat{b}), \tag{7.84}$$

where b is the pump mode and a is the signal mode. The object $\chi^{(2)}$ is a second-order nonlinear susceptibility. (For a discussion of nonlinear optics, the reader

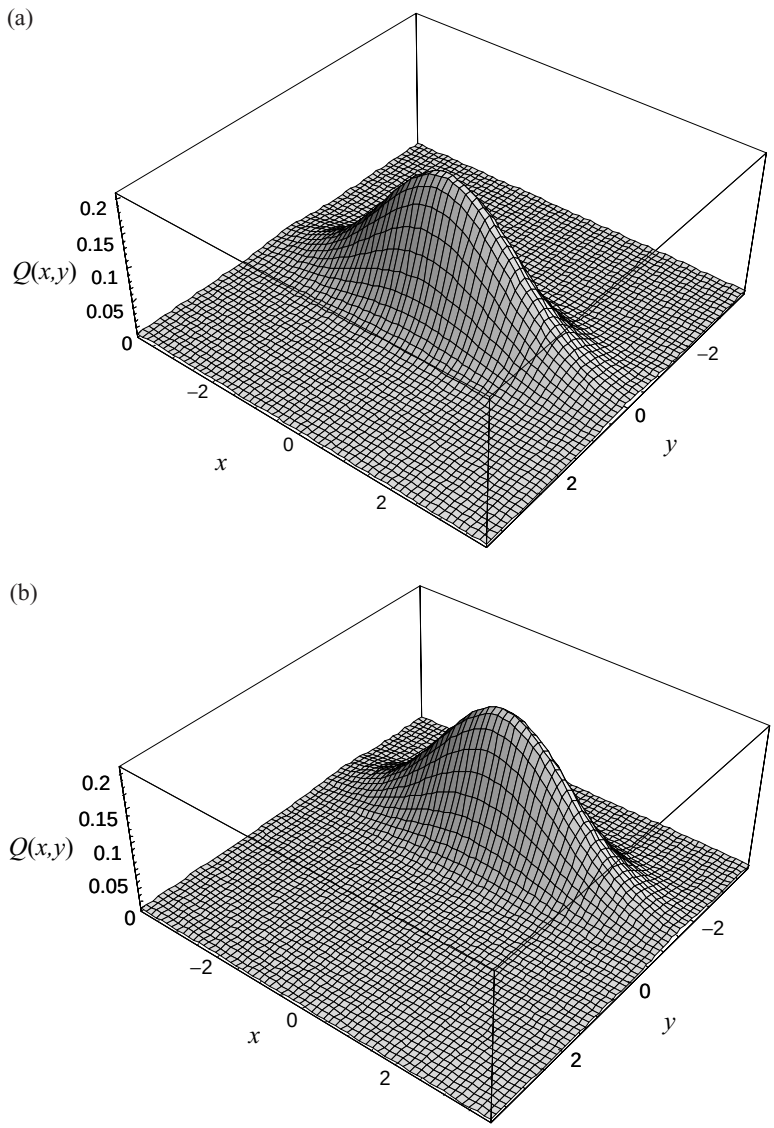


Fig. 7.11. Q function for (a) the squeezed vacuum, and (b) the displaced squeezed vacuum.

should consult, for example, the book by Boyd [11].) We now make the "parametric approximation" whereby we assume that the pump field is in a strong coherent classical field, which is strong enough to remain undepleted of photons over the relevant time scale. We assume that that field is in a coherent state $|\beta e^{-i\omega_{\rm P}t}\rangle$ and approximate the operators $\hat{b}$ and $\hat{b}^\dagger$ by $\beta e^{-i\omega_{\rm P}t}$ and $\beta^* e^{i\omega_{\rm P}t}$, respectively. Dropping irrelevant constant terms, the parametric approximation to the Hamiltonian in Eq. (7.84) is

$$\hat{H}^{(\rm PA)} = \hbar\omega\,\hat{a}^\dagger\hat{a} + i\,\hbar\left(\eta^*\hat{a}^2 e^{i\omega_{\rm P}t} - \eta\hat{a}^{\dagger 2}e^{-i\omega_{\rm P}t}\right), \tag{7.85}$$

where $\eta = \chi^{(2)}\beta$. Finally, transforming to the interaction picture, we obtain

$$\hat{H}_{\rm I}(t) = i\,\hbar\left[\eta^*\hat{a}^2 e^{i(\omega_{\rm P}-2\omega)t} - \eta\hat{a}^{\dagger 2}e^{-i(\omega_{\rm P}-2\omega)t}\right], \tag{7.86}$$

which is generally time dependent. But if $\omega_{\rm P}$ is chosen such that $\omega_{\rm P} = 2\omega$, we arrive at the time-independent interaction Hamiltonian

$$\hat{H}_{\rm I} = i\,\hbar(\eta^*\hat{a}^2 - \eta\hat{a}^{\dagger 2}). \tag{7.87}$$

The associated evolution operator

$$\hat{U}_{\rm I}(t,0) = \exp(-i\,\hat{H}_{\rm I}t/\hbar) = \exp(\eta^* t\hat{a}^2 - \eta t\hat{a}^{\dagger 2}) \tag{7.88}$$

obviously has the form of the squeeze operator of Eq. (7.10), $\hat{U}_{\rm I}(t,0) = \hat{S}(\xi)$, for $\xi = 2\eta t$.

There is another nonlinear process that gives rise to squeezed light, namely *degenerate four-wave mixing* in which two pump photons are converted into two signal photons of the same frequency. The fully quantized Hamiltonian for this process is

$$\hat{H} = \hbar\omega\,\hat{a}^\dagger\hat{a} + \hbar\omega\,\hat{b}^\dagger\hat{b} + i\,\hbar\chi^{(3)}(\hat{a}^2\hat{b}^{\dagger 2} - \hat{a}^{\dagger 2}\hat{b}^2) \tag{7.89}$$

where $\chi^{(3)}$ is a third-order nonlinear susceptibility. Going through similar arguments as above for the parametric down-converter, and once again assuming a strong classical pump field we obtain the parametric approximation of Eq. (7.87) but this time with $\eta = \chi^{(3)}\beta^2$.

7.3 Detection of quadrature squeezed light

It is, of course, not enough to generate squeezed light; one must be able to detect it. Several schemes for detection have been proposed and implemented. The general idea behind all proposed methods is to mix the signal field, presumed to contain the squeezing, with a strong coherent field, called the "local oscillator". Here we shall consider only one method, the one known as *balanced homodyne detection*.

A schematic of the method is shown in Fig. 7.12. Mode a contains the single field that is possibly squeezed. Mode b contains a strong coherent classical field which may be taken as a coherent state of amplitude β. The beam splitters are assumed to be 50:50 (hence the term *balanced* homodyne detection). Let us assume that the relation between the input $(\hat{a}, \hat{b})$ and output $(\hat{c}, \hat{d})$ operators is the same as in Eq. (6.10):

$$\begin{aligned}\hat{c} &= \frac{1}{\sqrt{2}}(\hat{a} + i\hat{b})\\ \hat{d} &= \frac{1}{\sqrt{2}}(\hat{b} + i\hat{a}).\end{aligned} \tag{7.90}$$

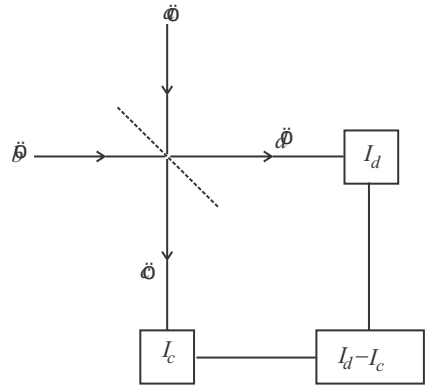


Fig. 7.12. Schematic of the balanced homodyne method for the detection squeezing. The field to be detected enters along $\hat{a}$ while a strong coherent field is injected along $\hat{b}$. The boxes in the lower left and upper right represent photo-detectors measuring the respective photo-currents. The box on the lower right represents a correlation device that subtracts the photo-currents.

The detectors placed in the output beams measure the intensities $I_c = \langle \hat{c}^\dagger \hat{c} \rangle$ and $I_d = \langle \hat{d}^\dagger \hat{d} \rangle$, and the difference in these intensities is

$$\begin{aligned} I_c - I_d &= \langle \hat{n}_{cd} \rangle = \langle \hat{c}^\dagger \hat{c} - \hat{d}^\dagger \hat{d} \rangle \\ &= i \langle \hat{a}^\dagger \hat{b} - \hat{a} \hat{b}^\dagger \rangle \end{aligned} \tag{7.91}$$

where we have used Eq. (7.90) to obtain the last line and where we have set $\hat{n}_{cd} = \hat{c}^\dagger \hat{c} - \hat{d}^\dagger \hat{d}$. Assuming the b mode to be in the coherent state $|\beta e^{-i\omega t}\rangle$, where $\beta = |\beta| e^{-i\psi}$, we have

$$\langle \hat{n}_{cd} \rangle = |\beta| \{ \hat{a} e^{i\omega t} e^{-i\theta} + \hat{a}^\dagger e^{-i\omega t} e^{i\theta} \} \tag{7.92}$$

where $\theta = \psi + \pi/2$. Assuming that the a mode light is also of frequency ω (and in practice *both* the a and b mode light fields derive from the same laser) we can set $\hat{a} = \hat{a}_0 e^{-i\omega t}$ so that we may write

$$\langle \hat{n}_{cd} \rangle = 2|\beta| \langle \hat{X}(\theta) \rangle \tag{7.93}$$

where

$$\hat{X}(\theta) = \frac{1}{2} \{ \hat{a}_0 e^{-i\,\theta} + \hat{a}_0^\dagger e^{i\,\theta} \} \tag{7.94}$$

is the field quadrature operator at the angle θ. By changing θ, which can be done by changing the phase ψ of the local oscillator, we can measure an arbitrary quadrature of the signal field. Of course, generally, θ will be chosen so as to achieve the maximum amount of quadrature squeezing. The variance of the output photon number difference operator $\hat{n}_{cd}$, in the limit of a strong local oscillator is

$$\left\langle (\Delta \hat{n}_{cd})^2 \right\rangle = 4|\beta|^2 \left\langle (\Delta \hat{X}(\theta))^2 \right\rangle. \tag{7.95}$$

For the input squeezing condition $\langle(\Delta\hat{X}(\theta))^2\rangle < \frac{1}{4}$ we have $\langle(\Delta\hat{n}_{cd})^2\rangle < |\beta|^2$. In an actual experiment, the signal beam is first blocked in order to obtain the shot-noise level. Furthermore, one has to account for the inefficiencies of the photo-detectors. We ignore this concern here and refer the reader to the relevant literature.

The first experimental realization of squeezed light, by Slusher *et al.* [12], used a four-wave mixing interaction. A noise reduction of about 20% of the allowable below the shot-noise level was achieved. A subsequent experiment by Wu *et al*. [13] achieved a reduction of 65% below the shot-noise level. Their method used a parametric amplifier. Even greater noise reduction has been achieved in more-recent years [14].

7.4 Amplitude (or number) squeezed states

Recall from Chapter 2 the number phase uncertainty relation $[\hat{n}, \hat{\varphi}] = i$ which, though technically unfounded, leads to the heuristically correct number phase uncertainty relation $\Delta n \Delta\varphi \geq 1/2$ valid in the regime of large average photon number. For a coherent state $|\alpha\rangle$ for which $\Delta n = \bar{n}^{1/2} (\bar{n} = |\alpha|^2)$, it can be shown (see Problem 3.2) that $\Delta\varphi = 1/(2\bar{n}^{1/2})$ if $\bar{n} \gg 1$ and thus the number phase uncertainty relation is equalized. But it is possible to envision, in analogy to quadrature squeezing, states that are squeezed in number, where $\Delta n < \bar{n}^{1/2}$, or in phase, where $\Delta\varphi < 1/(2\bar{n}^{1/2})$. States that are squeezed in phase may be difficult to categorize as nonclassical, in part because of the lack of a Hermitian operator representing phase, as discussed in Chapter 2, and because it is not too clear what classical limit has to be overcome. After all, the phase fluctuations of a coherent state can be made arbitrarily small simply by increasing $\bar{n}$. But for the photon number it is a different story. We can write the variance as

$$\langle(\Delta\hat{n})^2\rangle = \langle\bar{n}\rangle + (\langle\hat{a}^{\dagger 2}\hat{a}^2\rangle - \langle\hat{a}^\dagger\hat{a}\rangle^2) \tag{7.96}$$

which, in terms of the P function, is

$$\langle(\Delta\hat{n})^2\rangle = \langle\hat{n}\rangle + \int d^2\alpha P(\alpha)\Big[|\alpha|^2 - \langle\hat{a}^\dagger\hat{a}\rangle^2\Big]^2, \tag{7.97}$$

where

$$\langle\hat{a}^\dagger\hat{a}\rangle = \int d^2\alpha P(\alpha)|\alpha|^2. \tag{7.98}$$

Obviously the condition $\langle(\Delta\hat{n})^2\rangle < \langle\hat{n}\rangle$ for amplitude (or photon number) squeezing requires that $P(\alpha)$ take on negative values in some region of phase space. Hence amplitude squeezing is nonclassical. It is perhaps worth pointing out that the first term of Eq. (7.97) is sometimes referred to as the "particle-like" contribution while the second term is the "wave-like". Indeed, for thermal light, where $\langle(\Delta\hat{n})^2\rangle = \langle\hat{n}\rangle + \langle\hat{n}\rangle^2$, or coherent light, where $\langle(\Delta\hat{n})^2\rangle = \langle\hat{n}\rangle$, such a separation

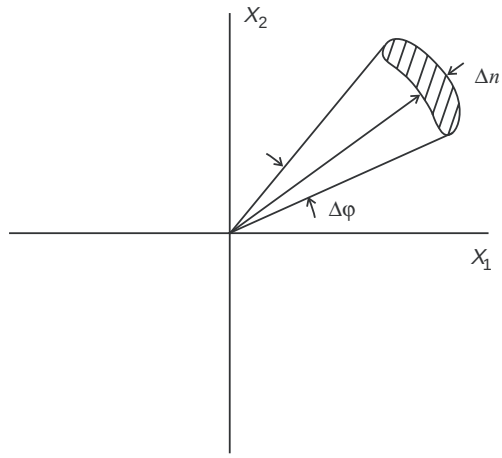


Fig. 7.13. The phase-space representation of a number squeezed state. Note that this particular state will also be quadrature squeezed.

holds up. But for amplitude squeezed light the interpretation of the second term as "wave-like" becomes a bit murky as it too takes on properties of a distinctly quantum-mechanical nature. A state exhibiting amplitude squeezing is said to possess sub-Poissonian statistics, the distribution being narrower than for a coherent state of the same average photon number. Not surprisingly, states with a photon number distribution broader than for a coherent state are said to have super-Poissonian statistics.

A simple way to gauge the nature of the photon statistics of any state is to calculate the so-called Q-parameter

$$Q = \frac{\langle(\Delta\hat{n})^2\rangle - \langle\hat{n}\rangle}{\langle\hat{n}\rangle} \tag{7.99}$$

(not to be confused with the Q function). For a state with Q in the range $-1 \leq Q < 0$ the statistics are sub-Poissonian, and if $Q > 0$, super-Poissonian. Obviously, $Q = 0$ for a coherent state.

In Fig. 7.13 we present a generic phase-space sketch of an amplitude squeezed state. For obvious reasons, such states are sometimes called "crescent" states. In the extreme case of a number state, as in Fig. 3.2, the "crescent" spreads out into a full circle and the Q-parameter goes to zero.

It is possible for some states of the field to exhibit simultaneously both quadrature and amplitude squeezing. Consider once again the states $|\alpha, \xi\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle$. As we have already discussed, it can be shown that

$$\langle\hat{n}\rangle = |\alpha|^2 + \sinh^2 r. \tag{7.100}$$

Also it can be shown that

$$\langle(\Delta\hat{n})^2\rangle = |\alpha\cosh r - 2^* e^{i\,\theta}\sinh r|^2 + 2\sinh^2 r\cosh^2 r. \tag{7.101}$$

Setting $\alpha = |\alpha|e^{i\,\psi}$ and choosing $\psi - \theta/2 = \pi/2$ we obtain

$$\langle(\Delta\hat{n})^2\rangle = |\alpha|^2 e^{2r} + 2\sinh^2 r\cosh^2 r, \tag{7.102}$$

which clearly does not exhibit amplitude squeezing. On the other hand, with $\psi = \theta/2$, we obtain

$$\left\langle (\Delta\hat{n})^2 \right\rangle = |\alpha|^2 e^{-2r} + 2\sinh^2 r \cosh^2 r, \tag{7.103}$$

which for certain choices of $|\alpha|$ and r may exhibit number squeezing. For example, for $|\alpha|^2$ very large and r small such that $\sinh r \approx 0$, $\langle \hat{n} \rangle \approx |\alpha|^2$ and $\langle (\Delta\hat{n})^2 \rangle \approx \langle \hat{n} \rangle e^{-2r}$ and thus $\langle (\hat{n})^2 \rangle < \langle \hat{n} \rangle$. The connection to quadrature squeezing can be made quite explicit by using Eq. (7.25) to write

$$\langle (\Delta\hat{n})^2 \rangle \approx 4\langle \hat{n} \rangle \left\langle (\Delta\hat{Y}_1)^2 \right\rangle. \tag{7.104}$$

On the experimental side, Short and Mandel [15] were the first to have observed, convincingly, sub-Poissonian statistics. This was in an experiment on single-atom resonance fluorescence.

7.5 Photon antibunching

In Section 5.4 we discussed the second-order coherence function $g^{(2)}(\tau)$ of Eq. (5.91) characterizing the joint probability of detecting one photon followed by another within the delay time τ at some fixed position. If $g^{(2)}(\tau) = 1$ then the photons arrive independently, and, in fact, we should expect that $g^{(2)}(\tau) \to 1$ as $\tau \to \infty$ for any field state, the "memory" of the first detected photon being lost after a long enough time. For coherent states, $g^{(2)}(\tau) = 1$. For $g^{(2)}(\tau) < g^{(2)}(0)$, the probability of detecting a second photon after delay time τ *decreases* and this indicates a *bunching* of photons. For a thermal field, $g^{(2)}(0) = 2$. On the other hand, for $g^{(2)}(\tau) > g^{(2)}(0)$, the probability of detecting a second photon *increases* with the delay time. This is photon antibunching which, as we have previously discussed, in Chapter 5, is a nonclassical effect. We can now explain why.

Let us first consider a single-mode field for which (as in Eq. (5.92))

$$\begin{aligned} g^{(2)}(\tau) = g^{(2)}(0) &= \frac{\left\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \right\rangle}{\left\langle \hat{a}^\dagger \hat{a} \right\rangle^2} \\ &= 1 + \frac{\left\langle (\Delta\hat{n})^2 \right\rangle - \langle \hat{n} \rangle}{\langle \hat{n} \rangle^2}. \end{aligned} \tag{7.105}$$

Strictly speaking, there cannot be any photon antibunching or bunching for a single-mode field as $g^{(2)}(\tau)$ is independent of the delay time τ. Nevertheless, in terms of the P function we can write Eq. (7.105) as

$$g^{(2)}(0) = 1 + \frac{\int P(\alpha) \left[|\alpha|^2 - \left\langle \hat{a}^\dagger \hat{a} \right\rangle \right]^2 d^2\alpha}{\left\langle \hat{a}^\dagger \hat{a} \right\rangle^2}, \tag{7.106}$$

where $\langle \hat{a}^\dagger \hat{a} \rangle$ is given by Eq. (7.98). For a classical field state where $P(\alpha) \geq 0$, we must have $g^{(2)}(0) \geq 1$. But for a nonclassical field state it is possible to have $g^{(2)}(0) < 1$, which, as previously stated, may be interpreted as a quantum-mechanical violation of the Cauchy inequality. The alert reader will have noticed

that the condition $g^{(2)}(0) < 1$ is the condition that the Q-parameter of Eq. (7.99) be negative; in other words, the condition for sub-Poissonian statistics. Indeed Q and $g^{(2)}(0)$, for a single-mode field, are simply related:

$$Q = \langle \hat{n} \rangle \left[g^{(2)}(0) - 1 \right]. \tag{7.107}$$

The fact that $Q < 0$ when $g^{(2)}(0) < 1$ has led to some confusion regarding the relationship of sub-Poissonian statistics and photon antibunching. Again, for a single-mode field, $g^{(2)}(\tau) = g^{(2)}(0) =$ constant and thus there can be no photon antibunching (or bunching for that matter). Bunching and antibunching occur only for multimode fields. For such fields in states with P function $P(\{\alpha_i\})$, where $\{\alpha_i\}$ denotes collectively the set of complex phase-space variables associated with each of the modes, the modes distinguished by the labels i and j, it can easily be shown that

$$g^{(2)}(0) = 1 + \frac{\int P(\{\alpha_i\}) \left[\sum_j |\alpha_j|^2 - \left\langle \hat{a}_j^\dagger \hat{a}_j \right\rangle \right]^2 d^2\{\alpha_i\}}{\left(\sum_j \left\langle \hat{a}_j^\dagger \hat{a}_j \right\rangle \right)^2} \tag{7.108}$$

where $d^2\{\alpha_i\} = d^2\alpha_1 d^2\alpha_2 \ldots$ and where

$$\left\langle \hat{a}_j^\dagger \hat{a}_j \right\rangle = \int P(\{\alpha_i\}) |\alpha_j|^2 d^2\{\alpha_i\}. \tag{7.109}$$

Again, for classical fields with $P(\{\alpha_i\}) \geq 0$, $g^{(2)}(0) \geq 1$. The Cauchy–Schwarz inequality applied to the corresponding classical coherence function $\gamma^{(2)}(\tau)$ implies that for classical fields it should always be the case that $g^{(2)}(\tau) \leq g^{(2)}(0)$, which does not allow for photon antibunching as shown in Chapter 5. The condition for antibunching, that $g^{(2)}(\tau) > g^{(2)}(0)$, is therefore an indication of nonclassical light as is the condition $g^{(2)}(0) < 1$. Remembering that $g^{(2)}(\tau) \to 1$ for $\tau \to \infty$, it follows that the condition $g^{(2)}(0) < 1$ implies photon antibunching, *except* (and this is an important exception) in the case of a single-mode field or, if $g^{(2)}(\tau)$ is constant, for some other reason. But the converse is not necessarily true. The condition for antibunching, $g^{(2)}(\tau) > g^{(2)}(0)$ does not imply sub-Poissonian statistics, $g^{(2)}(0) < 1$. Quite on the contrary, Zou and Mandel [16], in an article discussing the frequent confusion between the issues of sub-Poissonian statistics and photon antibunching, constructed a (somewhat artificial) two-mode state possessing simultaneously the properties of sub-Poissonian statistics and photon *bunching*.

In contrast to quadrature squeezing, sub-Poissonian statistics and photon antibunching are not phase dependent and are therefore related to the particle nature of light.

Photon antibunching was predicted in the resonance fluorescence from a two-level atom driven by a resonant laser field [17]. A full discussion of resonance

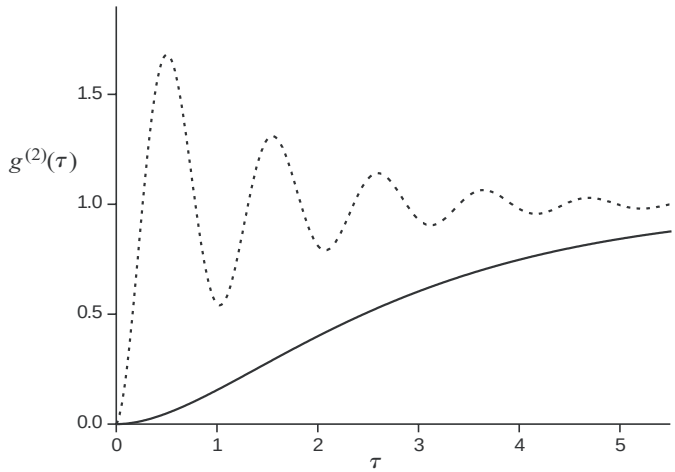


Fig. 7.14. Second-order coherence function for resonance fluorescence for the cases where $\Omega \gg \gamma$ (broken line) and $\Omega \ll \gamma$ (solid line).

fluorescence is beyond the scope of this book, but the results for the predicted second-order coherence function are easy to state and to interpret. With Ω the Rabi frequency of the driving field and γ the spontaneous emission decay rate, the stationary second-order correlation function $g^{(2)}(\tau)$ is given by

$$g^{(2)}(\tau) = [1 - \exp(-\gamma\tau/2)]^2, \quad \text{for} \quad \Omega \ll \gamma, \tag{7.110}$$

or

$$g^{(2)}(\tau) = 1 - \exp\left(-\frac{3\gamma}{4}\tau\right)\cos(\Omega\tau), \quad \text{for} \quad \Omega \gg \gamma. \tag{7.111}$$

These functions, for which $g^{(2)}(0) = 0$ in both limits, are plotted in Fig. 7.14.

The behaviors displayed in the figure may be interpreted as follows. Remember that $g^{(2)}(\tau)$ is essentially the joint probability of detecting a photon at time $t = 0$ and another photon at time $t = \tau$. But the source of the photons is a single atom driven by a coherent field. A detected fluorescent photon at $t = 0$ results, of course, from the electron "jump" from the excited to the ground state. But with the electron in the ground state, there is no possibility (i.e. the probability is zero) of emitting another photon at the time delay $\tau = 0$. There is a "dead time" before the laser field drives the electron back into the excited state with a finite probability at which time another photon may be emitted. The correlation function $g^{(2)}(\tau)$ is, in fact, proportional to the probability that at time τ the atom will be in the excited state given that it was initially in the ground state.

Experiments by Mandel and collaborators [18] have confirmed these predictions. These experiments involved a low-intensity beam of sodium atoms irradiated at a right angle by a stabilized dye laser resonantly tuned to a particular hypertime transition (the $3P_{3/2}$, $F = 3$, $M_F = 2$ to the $3^2S_{1/2}$, $F = 2$, $M_F = 2$). The beam intensity was so low that only one atom at a time occupied the region

of observation, that region being observed at right angles to both the atomic and laser beams.

7.6 Schrödinger cat states

As a last example of an important type of single-mode field state having strong nonclassical properties, we consider a class of states consisting of superpositions of two coherent states of equal amplitude but separated in phase by 180°, i.e. states of the form

$$|\psi\rangle = \mathcal{N}(|\alpha\rangle + e^{i\Phi}|-\alpha\rangle) \tag{7.112}$$

where the normalization factor $\mathcal{N}$ is given by

$$\mathcal{N} = [2 + 2\exp(-2\alpha^2)\cos\Phi]^{1/2} \tag{7.113}$$

(we have taken α real). For large $|\alpha|$, the states $|\alpha\rangle$ and $|-\alpha\rangle$ are macroscopically distinguishable and superpositions of the form of Eq. (7.112) are frequently referred to as the Schrödinger's cat states. Recall that the fate of Schrödinger's cat [19] was to end up in a superposition of macroscopically distinguishable states, the states of being alive and of being dead. It should be stressed that Schrödinger's intent was *not* to demonstrate how quantum weirdness could be elevated into the everyday classical world but rather to satirize the Copenhagen interpretation of quantum mechanics, which he, and many others, including Einstein, thought absurd. In spite of this, it has, in recent years, with the advancement of technology, become possible legitimately to consider laboratory realizations of superpositions of quantum states that are in some way macroscopically distinguishable. The driving force behind much of the activity is the attempt to address the question: where is the border between the quantum mechanical and classical worlds? Superposition states of the form of Eq. (7.112) are never seen in the macroscopic everyday world. Why not? Part of the answer seems to be as follows: no quantum system, particularly a macroscopic one, is truly an isolated system; it ultimately interacts with the rest of the universe, the "environment". The environment involves innumerable degrees of freedom which are not observed although, in some sense, the environment "observes" the system, effectively interacting with it in a dissipative fashion. In reality, of course, the entire universe is quantum mechanical and when a small part of it interacts with the rest, the two subsystems become entangled. As shown in Appendix A, tracing out the variables of the unobserved part of an entangled system leaves the system of interest in a mixed state. That's the general idea here. If a coherent macroscopic superposition state of the form of Eq. (7.112) can be created somehow, it should, upon interaction with the environment, quickly "decohere" into a statistical mixture of the form

$$\hat{\rho} = \frac{1}{2}(|\beta\rangle\langle\beta| + |-\beta\rangle\langle-\beta|) \tag{7.114}$$

where $\beta = \alpha\, e^{-\gamma\, t/2}$, and where, here, this γ is related to the rate of energy dissipation. Furthermore, it should be the case that the more macroscopic the components of initial superposition state, i.e. the greater $|\alpha|$, the more rapid should be the decoherence. We shall address these issues in more detail in Chapter 8. For the present, we concentrate on the properties of the cat states of Eq. (7.112) with a particular interest in those properties that are nonclassical. These are important as one must distinguish between superposition states and statistical mixtures, the latter having properties that are only classical.

There are three important cat states, depending on the choice of the relative phase Φ, that we shall consider. For $\Phi = 0$ we obtain the even coherent states

$$|\psi_{\rm e}\rangle = \mathcal{N}_{\rm e}(|\alpha\rangle + |-\alpha\rangle), \tag{7.115}$$

for $\Phi = \pi$, the odd coherent states

$$|\psi_{\rm o}\rangle = \mathcal{N}_0(|\alpha\rangle - |-\alpha\rangle), \tag{7.116}$$

where

$$\mathcal{N}_{\rm e} = \frac{1}{\sqrt{2}}[1 + \exp(-2\alpha^2)]^{-1/2}, \quad \mathcal{N}_0 = \frac{1}{\sqrt{2}}[1 - \exp(-2\alpha^2)]^{-1/2}, \tag{7.117}$$

are the respective normalization factors. These states were first introduced by Dodonov *et al*. [20]. For $\Phi = \pi/2$, we have the Yurke–Stoler states [21]

$$|\psi_{\rm ys}\rangle = \frac{1}{\sqrt{2}}\,(|\alpha\rangle + i|-\alpha\rangle). \tag{7.118}$$

All three states are eigenstates of the square of the annihilation operator with α^2 as the eigenvalue:

$$\hat{a}^2|\psi\rangle = \alpha^2|\psi\rangle. \tag{7.119}$$

We must be able to distinguish between the cat states Eqs. (7.115)–(7.119), which are coherent superpositions of $|\alpha\rangle$ and $|-\alpha\rangle$, and the statistical mixture

$$\hat{\rho}_{\rm mixture} = \frac{1}{2}(|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha|). \tag{7.120}$$

That is, as the corresponding density operator for a cat state, say, the even coherent state, is

$$\hat{\rho}_{\rm e} = |\psi_{\rm e}\rangle\langle\psi_{\rm e}| = |\mathcal{N}_{\rm e}|^2(|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha| + |-\alpha\rangle\langle\alpha| + |\alpha\rangle\langle-\alpha|), \tag{7.121}$$

we need to be able to detect the effects of the "coherence" terms $|\alpha\rangle\langle-\alpha|$ and $|-\alpha\rangle\langle\alpha|$ not present in the statistical mixture of Eq. (7.120).

Let us first consider the photon statistics. For $|\psi_{\rm e}\rangle$ we obtain

$$P_n = \begin{cases} \dfrac{2\exp(-\alpha^2)}{1+\exp(-\alpha^2)}\,\dfrac{\alpha^{2n}}{n!} & n \text{ even}, \\ 0 & n \text{ odd}, \end{cases} \tag{7.122}$$

and for the odd coherent state

$$P_n = \begin{cases} 0 & n \text{ even}, \\ \dfrac{2\exp(-\alpha^2)}{1+\exp(-\alpha^2)} & n \text{ odd}. \end{cases} \tag{7.123}$$

For the Yurke–Stoler state, P_n is just a Poisson distribution, identical to that of a coherent state $|\alpha\rangle$ *and* of the statistical mixture of Eq. (7.120), i.e. $P_n = \langle n|\hat{\rho}_{\text{mixture}}|n\rangle$. At least for the even and odd coherent states, oscillating photon number distributions distinguish those states from statistical mixtures. It is easy to see that the disappearance of the odd number states in $|\psi_e\rangle$ or the even in $|\psi_o\rangle$ is the result of quantum interference. The Q-parameter for the even coherent states is

$$Q = \frac{4\alpha^2\exp(-\alpha^2)}{1+\exp(-4\alpha^2)} > 0 \tag{7.124}$$

indicating super-Poissonian statistics for all α and for the odd coherent states

$$Q = -\frac{4\alpha^2\exp(-\alpha^2)}{1+\exp(-4\alpha^2)} < 0 \tag{7.125}$$

indicating sub-Poissonian statistics, again for all α. As α becomes large, $Q \to 0$ for both the even and odd coherent states. For the Yurke–Stoler states, for all α, $Q = 0$.

Now we look at quadrature squeezing. With respect to the even coherent state we find that

$$\left\langle (\Delta\hat{X}_1)^2 \right\rangle = \frac{1}{4} + \frac{\alpha^2}{1+\exp(-2\alpha^2)} \tag{7.126}$$

$$\left\langle (\Delta\hat{X}_2)^2 \right\rangle = \frac{1}{4} - \frac{\alpha^2\exp(-2\alpha^2)}{1+\exp(-2\alpha^2)} \tag{7.127}$$

so reduced fluctuations appear in the $\hat{X}_2$ quadrature for α not too large. For the odd coherent states we have

$$\left\langle (\Delta\hat{X}_1)^2 \right\rangle = \frac{1}{4} + \frac{\alpha^2}{1-\exp(-2\alpha^2)} \tag{7.128}$$

$$\left\langle (\Delta\hat{X}_2)^2 \right\rangle = \frac{1}{4} + \frac{\alpha^2\exp(-2\alpha^2)}{1-\exp(-2\alpha^2)} \tag{7.129}$$

and thus no squeezing is evident. (Note the role reversal here between the even and odd coherent states regarding quadrature squeezing and sub-Poissonian statistics.) For the Yurke–Stoler state we find

$$\left\langle (\Delta\hat{X}_1)^2 \right\rangle = \frac{1}{4} + \alpha^2 \tag{7.130}$$

$$\left\langle (\Delta\hat{X}_2)^2 \right\rangle = \frac{1}{4} - \alpha^2\ \exp(-4\alpha^2) \tag{7.131}$$

so squeezing appears in $\hat{X}_2$. Finally for the statistical mixture of Eq. (7.120) we find

$$\left\langle (\Delta \hat{X}_1)^2 \right\rangle = \frac{1}{4} + \alpha^2 \tag{7.132}$$

$$\left\langle (\Delta \hat{X}_2)^2 \right\rangle = \frac{1}{4} \tag{7.133}$$

which, as expected, exhibits no squeezing.

Another possible way to distinguish between the coherent superpositions and the statistical mixture is to examine the phase distribution. Using the phase states $|\theta\rangle$ introduced in Chapter 2, we will obtain, for the state of Eq. (7.112), in the limit of large α, the phase distribution (again with α real),

$$\begin{aligned} \mathcal{P}_{|\psi\rangle}(\theta) \approx \left(\frac{2\bar{n}}{\pi}\right)^{1/2} & |\mathcal{N}|^2 \{\exp[-2\bar{n}\theta^2] \\ & + \exp[-2\bar{n}(\theta - \pi)^2] \\ & + 2\cos(\bar{n}\pi - \Phi)\exp[-\bar{n}\theta^2 - \bar{n}(\theta - \pi)^2]\} \end{aligned} \tag{7.134}$$

where $\bar{n} = |\alpha|^2$. For the density operator of Eq. (7.120) we have

$$\mathcal{P}_{\rho}(\theta) \approx \frac{1}{2}\left(\frac{2\bar{n}}{\pi}\right)^{1/2} \{\exp[-\bar{n}\theta^2] + \exp[-\bar{n}(\theta - \pi)^2]\}. \tag{7.135}$$

Both distributions have peaks at $\theta = 0$ and $\theta = \pi$, as expected, but only the first has an interference term. Unfortunately, that interference term is very small, the Gaussians in the product having very little overlap.

As for the quasi-probability distributions, the P function is highly singular, involving an infinite sum of higher-order derivatives of a delta function. The Q function is, of course, always positive and, as such, does not give a clear signal for nonclassicality. But the Wigner function, when it takes on negative values, does exhibit a clear signal for nonclassicality. Using some results from Chapter 3, we find the Wigner function for the statistical mixture to be

$$\begin{aligned} W_{\mathrm{M}}(x, y) = \frac{1}{\pi} & \{\exp[-2(x - \alpha)^2 - 2y^2] \\ & + \exp[-2(x + \alpha)^2 - 2y^2]\} \end{aligned} \tag{7.136}$$

which is always positive and contains two Gaussian peaks centered at $x = \pm\alpha$ as shown in Fig. 7.15(a). But the Wigner function for the even coherent state is

$$\begin{aligned} W_{\mathrm{e}}(x, y) = & \frac{1}{\pi[1 + \exp(-2\alpha^2)]} \{\exp[-2(x - \alpha)^2 - 2y^2] \\ & + \exp[-2(x + \alpha)^2 - 2y^2] \\ & + 2\exp[-2x^2 - 2y^2]\cos(4y\alpha)\}. \end{aligned} \tag{7.137}$$

The last term results from the interference between the two states $|\alpha\rangle$ and $|-\alpha\rangle$. It causes the Wigner function to become highly oscillatory, akin to exhibiting

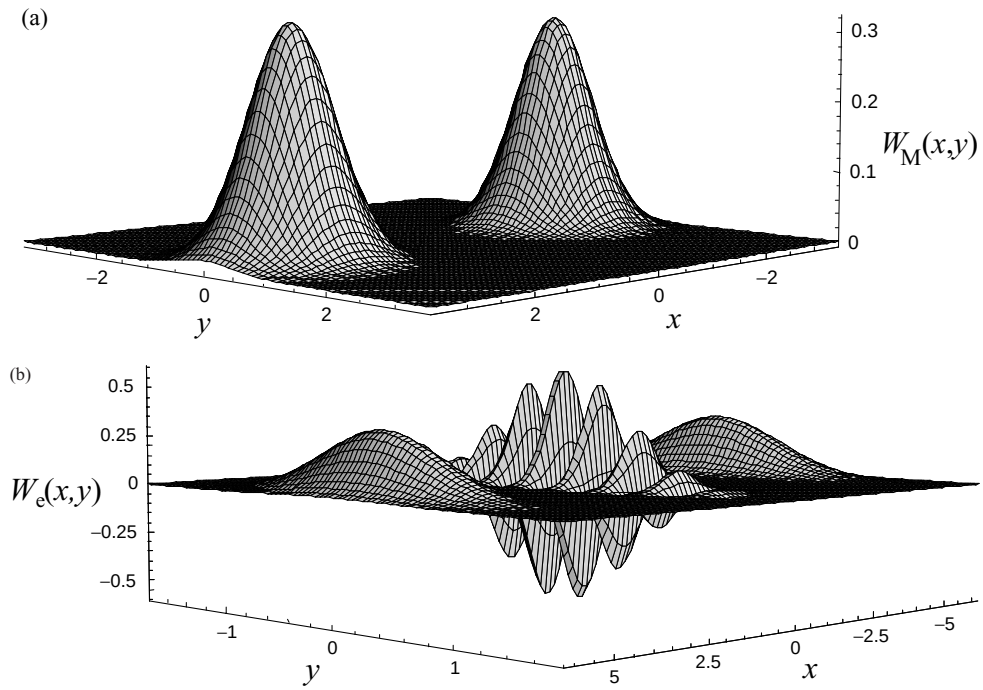


Fig. 7.15. Wigner function for (a) the statistical mixture of coherent states, and (b) the even coherent state, both for $\alpha = \sqrt{5}$. It is evident that the former is always positive and that the latter becomes negative and is highly oscillatory (i.e. displays interference fringes) in some regions of phase space.

interference fringes, and to go negative, as evident in Fig. 7.15(b). The negativity does not allow the Wigner function in this case to be interpreted as a probability distribution. For the sake of completeness, we give the corresponding Wigner functions for the odd coherent state and the Yurke–Stoler state:

$$W_0(x,y) = \frac{1}{\pi\left[1+\exp\left(-2\alpha^2\right)\right]}\left\{\exp\left[-2(x-\alpha)^2-2y^2\right]\right.$$
$$+\exp\left[-2(x+\alpha)^2-2y^2\right]$$
$$\left.-2\exp\left[-2x^2-2y^2\right]\cos(4y\alpha)\right\}, \tag{7.138}$$

$$W_{ys}(x,y) = \frac{1}{\pi}\left\{\exp\left[-2(x-\alpha)^2-2y^2\right]\right.$$
$$+\exp\left[-2(x+\alpha)^2-2y^2\right]$$
$$\left.-2\exp\left[-2x^2-2y^2\right]\sin(4y\alpha)\right\}. \tag{7.139}$$

These functions are similar in character to that of the even coherent state.

How can the Schrödinger cat states of Eqs. (7.115)–(7.118) be generated as a matter of principle (leaving aside for now the issue of decoherence)? We have earlier described how the squeezed states may be generated by parametric

processes. The states are generated by unitary evolution, the evolution operator being a realization of the squeeze operator. If the initial state is a vacuum, the squeezed vacuum evolves, that state populated by even photon numbers only. For the even coherent state of Eq. (7.115), again only even number states are populated but this time there does not seem to be a possible way to generate the states from the vacuum, or any other state, by unitary evolution. The same is true of the odd coherent state. Shortly, we shall describe a nonunitary method of generating such states, using projective state reduction. But the Yurke–Stoler state *can* be generated from the unitary evolution of a coherent state in a certain kind of nonlinear medium, namely a Kerr-like medium [21]. The interaction is a "self-interaction", i.e. involving no parametric driving fields, and is modeled by the interaction Hamiltonian

$$\hat{H}_{\rm I} = \hbar\, K(\hat{a}^\dagger \hat{a})^2 = \hbar\, K\, \hat{n}^2 \tag{7.140}$$

where K is proportional to a third-order nonlinear susceptibility $\chi^{(3)}$. An initial coherent state evolves into

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\,\hat{H}_{\rm I}\,t/\hbar}|\alpha\rangle \\ &= \exp(-|\alpha|^2/2)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\; e^{-iK\;n^2 t}|n\rangle. \end{aligned} \tag{7.141}$$

As n^2 is an integer, $|\psi(t)\rangle$ is periodic with period $T = 2\pi/K$. If $t = \pi/K$ we have $e^{-iKn^2 t} = e^{-i\pi n^2} = (-1)^n$ and thus $|\psi(\pi/K)\rangle = |-\alpha\rangle$. But for $t = \pi/2K$

$$e^{-iKn^2 t} = e^{-i\pi n^2/2} = \begin{cases} 1, & n \text{ even}, \\ i, & n \text{ odd}. \end{cases} \tag{7.142}$$

Thus we have

$$|\psi(\pi/2K)\rangle = \frac{1}{\sqrt{2}}\; e^{-i\,\pi/4}(|\alpha\rangle + i|-\alpha\rangle), \tag{7.143}$$

which, apart from an overall irrelevant phase factor, is obviously the Yurke–Stoler state of Eq. (7.118).

A more realistic form of the Kerr interaction is given by the Hamiltonian [22]

$$\hat{H}_{\rm I} = \hbar K \hat{a}^{\dagger 2}\hat{a}^2 = \hbar K(\hat{n}^2 - \hat{n}). \tag{7.144}$$

Its action on a coherent state is as

$$|\psi(t)\rangle = e^{-iKt(\hat{n}^2-\hat{n})}|\alpha\rangle = e^{-iKt\hat{n}^2}|\alpha e^{iK\,t}\rangle. \tag{7.145}$$

For $t = \pi/2K$ we have

$$|\psi(\pi/2K)\rangle = \frac{1}{\sqrt{2}}e^{-i\,\pi/4}(|\beta\rangle + i|-\beta\rangle), \tag{7.146}$$

where $\beta = i\,\alpha$, which is of the same form as Eq. (7.143), a Yurke–Stoler state apart from an irrelevant phase factor. The effect of the term linear in $\hat{n}$ is merely to produce a rotation of the amplitude of the coherent state.

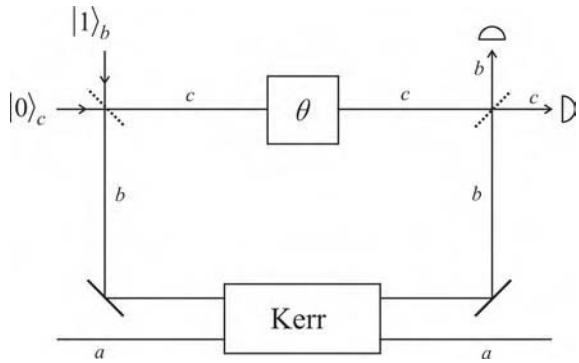


Fig. 7.16. Sketch of a proposed experiment for generating Schrödinger cat states. A single photon is injected into the first beam splitter of the interferometer and an external mode, coupled to an internal mode through a cross-Kerr interaction, is initially in a coherent state. Projective state reduction occurs upon detection of a photon in either of the detectors.

We now present a nonunitary method of generating the even and odd cat states, and the Yurke–Stoler states as well, for an optical field initially in a coherent state, as originally discussed by Gerry [23]. The central idea behind this approach is that of projective measurement where it is possible, in the case of an entangled states, to project one of the subsystems of the entangled states into a *pure* state through a measurement on the other subsystem (see Appendix B).

In Fig.7.16 we sketch the required experimental setup. Essentially we have a Mach–Zehnder interferometer whose modes we label b and c as indicated, coupled to an external mode a through a nonlinear medium realizing a so-called cross-Kerr interaction given by the Hamiltonian

$$\hat{H}_{\mathrm{CK}} = \hbar K\, \hat{a}^\dagger \hat{a}\, \hat{b}^\dagger \hat{b}, \tag{7.147}$$

where K is proportional to a third-order nonlinear susceptibility $\chi^{(3)}$. The action of the medium in the interferometer is given by the unitary evolution operator

$$\hat{U}_{\mathrm{CK}} = \exp\left(-itK\hat{a}^\dagger \hat{a}\, \hat{b}^\dagger \hat{b}\right), \tag{7.148}$$

where t is the interaction time $t = l/v$, where l is the length of the medium and v is the velocity of light in the medium. In the upper beam of the interferometer is placed a phase shift of θ which is generated by the operator

$$\hat{U}_{\mathrm{PS}}(\theta) = \exp\left(i\theta\, \hat{c}^\dagger \hat{c}\right). \tag{7.149}$$

The phase shift θ is to be adjusted. Beam splitters BS_1 and BS_2 we take to be of the type described by Eqs. (6.10). D_1 and D_2 are detectors placed in the output b and c beams. We assume that a coherent state $|\alpha\rangle$ is present in the a mode and

that a single photon is injected into BS_1 as indicated in Fig. 7.16. That is, the input state at BS_1 is $|1\rangle_b|0\rangle_c \equiv |10\rangle_{bc}$. Just after BS_1 we have the state

$$|\text{out}\rangle_{BS_1} = |\alpha\rangle_a \frac{1}{\sqrt{2}}(|10\rangle_{bc} + i|01\rangle_{bc}). \tag{7.150}$$

The actions of the phase shifter and the cross-Kerr medium yield the state

$$|K\,t, \theta\rangle = \frac{1}{\sqrt{2}}(|e^{-iKt}\alpha\rangle_a|10\rangle_{bc} + ie^{i\theta}|\alpha\rangle_a|01\rangle_{bc}), \tag{7.151}$$

with the obvious notation $|0\rangle_b|1\rangle_c \equiv |01\rangle_{bc}$. The rotated coherent state of the first term results from the presence of the single photon in the b mode. In the second term, the photon is in the c mode and so picks up the phase shift θ. We now assume that $Kt = \pi$ so that we have

$$|\pi, \theta\rangle = \frac{1}{\sqrt{2}}(|-\alpha\rangle_a|10\rangle_{bc} + ie^{i\theta}|\alpha\rangle_a|01\rangle_{bc}). \tag{7.152}$$

Next, as in Eqs. (6.21), BS_2 causes the transformations

$$\begin{aligned} |10\rangle_{bc} &\xrightarrow{BS_2} \frac{1}{\sqrt{2}}(|10\rangle_{bc} + i|01\rangle_{bc}), \\ |01\rangle_{bc} &\xrightarrow{BS_2} \frac{1}{\sqrt{2}}(|01\rangle_{bc} + i|10\rangle_{bc}), \end{aligned} \tag{7.153}$$

so that just after BS_2 we have the state

$$\begin{aligned} |\text{out}\rangle_{BS_2} &= \frac{1}{2}[|-\alpha\rangle_a(|10\rangle_{bc} + i|01\rangle_{bc}) + ie^{i\theta}|\alpha\rangle_a(|01\rangle_{bc} + i|10\rangle_{bc})] \\ &= \frac{1}{2}[(|-\alpha\rangle_a - e^{i\theta}|\alpha\rangle_a)|10\rangle_{bc} + i(|-\alpha\rangle_a + e^{i\theta}|\alpha\rangle_a)|01\rangle_{bc}]. \end{aligned} \tag{7.154}$$

We now choose the phase shift $\theta = \pi$ so that we have

$$|\text{out}\rangle_{BS_2} = \frac{1}{2}[(|\alpha\rangle_a + |-\alpha\rangle_a)|10\rangle_{bc} - i(|\alpha\rangle_a - |-\alpha\rangle_a)|01\rangle_{bc}]. \tag{7.155}$$

Now, if the detector D_1 fires, indicating a photon has emerged from the beam splitter in the b mode, the a mode is projected into the even coherent state $|\alpha\rangle_a + |-\alpha\rangle_a$ which upon normalization has the form of Eq. (7.115). But if detector D_2 fires, indicating that a photon has emerged from the beam splitter in mode c, then the a mode is projected into the odd coherent state $|\alpha\rangle_a - |-\alpha\rangle_a$, given in normalized form by Eq. (7.116). State reduction is a discontinuous, and thus nonunitary, process that we can use to great effect as a method for generating a variety of quantum states in entangled systems. Note that had we chosen to perform state reduction *before* the second beam splitter, i.e. with the system in the state of Eq. (7.152), we would not have been able to project onto the even and odd coherent states. The beam splitter effectively provides a means of scrambling the information on path and thus produces the required interference to generate superpositions upon measurement. In contrast, a measurement on the state of Eq. (7.152) produces path information and hence there is no interference. Note that had we chosen $\theta = \pi/2$, the projective measurements would have yielded

the states $|\alpha\rangle \pm i|-\alpha\rangle$, essentially Yurke–Stoler states. In Chapter 10, we shall describe a similar means of generating the cat states in the context of cavity QED.

7.7 Two-mode squeezed vacuum states

So far in this chapter, we have considered nonclassical states of a single-mode field. However, for multimode fields, the prospects for nonclassical effects become even richer as long as the field states are not merely product states of each of the modes. In other words, if we have *entangled* multimode field states we may expect, in general, strong nonclassical effects of greater richness than is possible for a single-mode field. In this section we shall consider the simple, and yet very important, example of the two-mode squeezed vacuum state.

We begin by introducing, in analogy with the single-mode squeeze operator of Eq. (7.10), the two-mode squeeze operator

$$\hat{S}_2(\xi) = \exp(\xi^* \hat{a}\hat{b} - \xi \hat{a}^\dagger \hat{b}^\dagger), \tag{7.156}$$

where $\xi = re^{i\theta}$ as before and where $\hat{a}$ and $\hat{b}$ are the operators for the two modes. Of course, $[\hat{a}, \hat{b}^\dagger] = 0$, etc. Note that $\hat{S}_2(\xi)$ does not factor into a product of the single-mode squeeze operators of the form of Eq. (7.10) for each mode. We now define the two-mode squeezed vacuum state by the action of $\hat{S}_2(\xi)$ on the two-mode vacuum $|0\rangle_a|0\rangle_b = |0, 0\rangle$:

$$\begin{aligned} |\xi\rangle_2 &= \hat{S}_2(\xi)|0, 0\rangle \\ &= \exp(\xi^* \hat{a}\hat{b} - \xi \hat{a}^\dagger \hat{b}^\dagger)|0, 0\rangle. \end{aligned} \tag{7.157}$$

Because $\hat{S}_2$ does not factor as a product of two single-mode squeeze operators, the two-mode squeezed vacuum state is *not* a product of two single-mode squeezed vacuum states, rather, as we shall later show, it is an entangled state containing strong correlations between the two modes. Notice that the two-mode squeeze operator contains terms that either create or annihilate photons pair-wise from each mode.

It turns out that because of the correlations between the modes, the squeezing of quantum fluctuations is not in the individual modes but rather in a superposition of the two-modes. We define the superposition quadrature operators as follows:

$$\begin{aligned} \hat{X}_1 &= \frac{1}{2^{3/2}}(\hat{a} + \hat{a}^\dagger + \hat{b} + \hat{b}^\dagger) \\ \hat{X}_2 &= \frac{1}{2^{3/2}i}(\hat{a} - \hat{a}^\dagger + \hat{b} - \hat{b}^\dagger). \end{aligned} \tag{7.158}$$

The operators satisfy the same commutation relations as for the single-mode case, $[\hat{X}_1, \hat{X}_2] = i/2$, and thus squeezing exists in the superposition quadratures if the conditions of Eq. (7.3) are satisfied. That such squeezing is nonclassical may be seen by writing the two-mode analogs to Eqs. (7.4) and (7.5) using a two-mode P function, $P(\alpha, \beta) = P(\alpha)P(\beta)$. For squeezing to exist, $P(\alpha, \beta)$ must

be nonpositive or singular in some regions of phase space. We omit the details here.

To investigate the squeezing properties of our state, we first use the Baker–Hausdorf lemma to obtain

$$\begin{aligned}\hat{S}_2^\dagger(\xi)\hat{a}\hat{S}_2(\xi) &= \hat{a}\cosh r - e^{i\theta}\hat{b}^\dagger \sinh r \\ \hat{S}_2^\dagger(\xi)\hat{b}\hat{S}_2(\xi) &= \hat{b}\cosh r - e^{i\theta}\hat{a}^\dagger \sinh r.\end{aligned} \tag{7.159}$$

By using these results, we can show (we leave the details as an exercise) that for the two-photon squeezed vacuum state

$$\langle \hat{X}_1\rangle = 0 = \langle \hat{X}_2\rangle \tag{7.160}$$

and that

$$\begin{aligned}\left\langle (\hat{X}_1)^2\right\rangle &= \frac{1}{4}[\cosh^2 r + \sinh^2 r - 2\sin r\cosh r\cos\theta] \\ \left\langle (\hat{X}_2)^2\right\rangle &= \frac{1}{4}[\cosh^2 r + \sinh^2 r + 2\sin r\cosh r\cos\theta].\end{aligned} \tag{7.161}$$

For $\theta = 0$ we obtain

$$\left\langle (\hat{X}_1)^2\right\rangle = \frac{1}{4}e^{-2r}, \qquad \left\langle (\hat{X}_2)^2\right\rangle = \frac{1}{4}e^{2r}. \tag{7.162}$$

The results in Eqs. (7.162) are mathematically identical to the case for a single-mode squeezed vacuum state but the interpretation here is in terms of the superposition of two modes.

We now seek a decomposition of our state $|\xi\rangle_2$ in terms of the two mode number states $|n\rangle_a \otimes |m\rangle_b \equiv |n, m\rangle$. We follow our earlier procedure by starting with

$$\hat{a}|0, 0\rangle = 0 \tag{7.163}$$

and using

$$\hat{S}_2(\xi)\hat{a}\hat{S}_2^\dagger(\xi) = \hat{a}\cosh r + e^{i\,\theta}\hat{b}^\dagger \sinh r \tag{7.164}$$

to write

$$\begin{aligned}&\hat{S}_2(\xi)\hat{a}\hat{S}_2^\dagger(\xi)\hat{S}_2(\xi)|0, 0\rangle \\ &\quad = (\mu\hat{a} + \nu\hat{b}^\dagger)|\xi\rangle_2 = 0\end{aligned} \tag{7.165}$$

where $\mu = \cosh r$ and $\nu = e^{i\theta}\sinh r$. Expanding $|\xi\rangle_2$ as

$$|\xi\rangle_2 = \sum_{n,m} C_{n,m}|n, m\rangle \tag{7.166}$$

the eigenvalue problem becomes

$$\sum_{n,m} C_{n,m}[\mu\sqrt{n}|n-1, m\rangle + \nu\sqrt{m+1}\;|n, m+1\rangle] = 0. \tag{7.167}$$

Of the many *possible* solutions, we are only interested in the one containing the two-mode vacuum state $|0, 0\rangle$. This solution has the form

$$C_{n,m} = C_{0,0}\left(-\frac{\nu}{\mu}\right)^n \delta_{n,m} = C_{0,0}(-1)^n e^{in\theta} \tanh^n r\, \delta_{n,m} \tag{7.168}$$

where $C_{0,0}$ is determined from normalization to be $C_{0,0} = (\cosh r)^{-1}$. Thus the two-mode squeezed vacuum states, in terms of the number states, are

$$|\xi\rangle_2 = \frac{1}{\cosh r} \sum_{n=0}^{\infty} (-1)^n e^{in\theta} (\tanh r)^n |n, n\rangle. \tag{7.169}$$

Aside from the obvious fact that our states are entangled, perhaps their most striking feature is the evident strong correlations between the two modes: only the paired states $|n, n\rangle$ occur in the superposition of Eq. (7.169). This is one reason why such states are often called "twin beams". Another way to express this is to note that $|\xi\rangle_2$ is also an eigenstate of the number difference operator $\hat{n}_a - \hat{n}_b$, where $\hat{n}_a = \hat{a}^\dagger \hat{a}$ and $\hat{n}_b = \hat{b}^\dagger \hat{b}$, with eigenvalue zero:

$$(\hat{n}_a - \hat{n}_b)|\xi\rangle_2 = 0. \tag{7.170}$$

Because of the correlations and of the symmetry between the two modes, the average photon number in each mode is the same and can easily be shown to be

$$\langle \hat{n}_a \rangle = \langle \hat{n}_b \rangle = \sinh^2 r, \tag{7.171}$$

the same expression as for the single-mode squeezed vacuum (Eq. (7.26) with $\alpha = 0$). The variances in the number operators are

$$\langle (\Delta \hat{n}_a)^2 \rangle = \langle (\Delta \hat{n}_b)^2 \rangle = \sinh^2 r \cosh^2 r = \frac{1}{4} \sinh^2(2r). \tag{7.172}$$

Obviously $\langle (\Delta \hat{n}_i) \rangle > \langle \hat{n}_i \rangle$, $i = a, b$, so both modes exhibit super-Poissonian photon statistics.

In order to quantify the quantum correlation between the two modes, we must examine operators acting on both systems. Here we take the combination photon number operators $\hat{n}_a \pm \hat{n}_b$ and examine the variances

$$\langle [\Delta(\hat{n}_a \pm \hat{n}_b)]^2 \rangle = \langle (\Delta \hat{n}_a)^2 \rangle + \langle (\Delta \hat{n}_b)^2 \rangle \pm 2\mathrm{cov}(\hat{n}_a, \hat{n}_b) \tag{7.173}$$

where the covariance of the photon numbers is defined as

$$\mathrm{cov}(\hat{n}_a, \hat{n}_b) \equiv \langle \hat{n}_a \hat{n}_b \rangle - \langle \hat{n}_a \rangle \langle \hat{n}_b \rangle. \tag{7.174}$$

For states containing *no* intermode correlations the expectation value of the product of the number operators factors as $\langle \hat{n}_a \hat{n}_b \rangle = \langle \hat{n}_a \rangle \langle \hat{n}_b \rangle$ and the covariance will be zero. But for the two-mode squeezed vacuum states, we know, by virtue of Eq. (7.170) that the variance of the number difference operator $\hat{n}_a - \hat{n}_b$ must be zero:

$$\langle [\Delta(\hat{n}_a - \hat{n}_b)]^2 \rangle = 0. \tag{7.175}$$

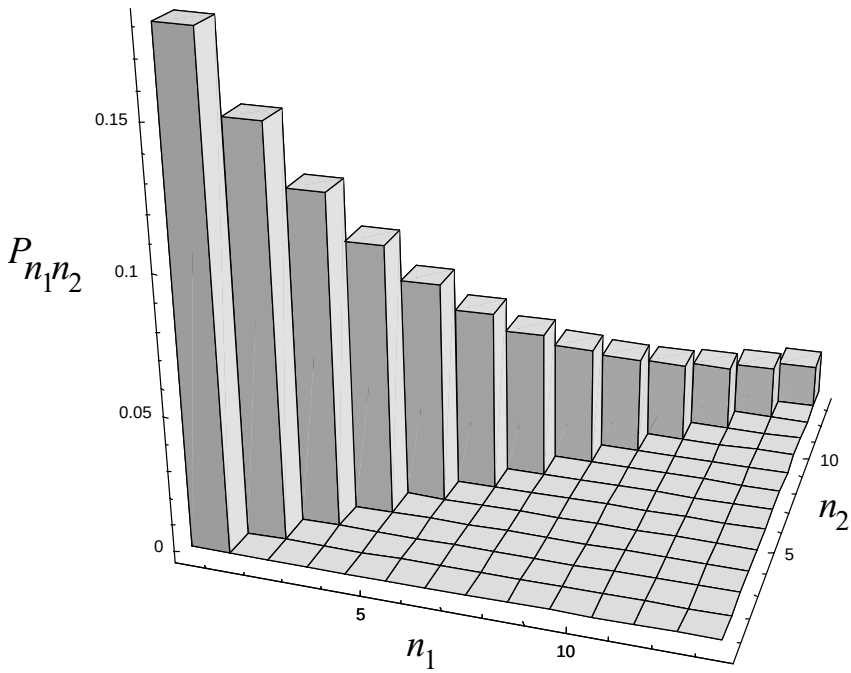


Fig. 7.17. $P_{n_1\, n_2}$ versus n_1 and n_2 for the two-mode squeezed vacuum state for $r = 2$.

It follows from Eqs. (7.172) and (7.173) that

$$\mathrm{cov}(\hat{n}_a, \hat{n}_b) = \frac{1}{4}\sinh^2(2r). \tag{7.176}$$

Thus the linear correlation coefficient, defined as

$$J(\hat{n}_a, \hat{n}_b) = \frac{\mathrm{cov}(\hat{n}_a, \hat{n}_b)}{\langle(\Delta\hat{n}_a)^2\rangle^{1/2}\ \langle(\Delta\hat{n}_b)^2\rangle^{1/2}}, \tag{7.177}$$

which, from Eqs. (7.172) and (7.176), takes the value

$$J(\hat{n}_a, \hat{n}_b) = 1, \tag{7.178}$$

the maximal value indicating strong intermode correlations.

From Eq. (7.169) we can obtain the joint probability of finding n_1 photons in mode a and n_2 photons in mode b, $P_{n_1 n_2}$:

$$\begin{aligned} P_{n_1\, n_2} &= |\langle n_1, n_2 \mid \xi\rangle_2|^2 \\ &= (\cosh r)^{-2}(\tanh r)^{2n}\delta_{n_1,n}\delta_{n_2,n}. \end{aligned} \tag{7.179}$$

If we plot $P_{n_1 n_2}$ versus n_1 and n_2, as in Fig. 7.17, it is clear that the joint probability is a monotonically decreasing function along the diagonal $n_1 = n_2$.

In order to examine properties of the individual modes it is useful to introduce the density operator for the state $|\xi\rangle_2$:

$$\hat{\rho}_{ab} = |\xi\rangle_{2\ 2}\langle\xi|. \tag{7.180}$$

The reduced density operators for the individual modes are easily found to be (see Appendix A)

$$\begin{aligned}\hat{\rho}_a &= \sum_{n=0}^{\infty} \frac{1}{(\cosh r)^2} (\tanh r)^{2n} |n\rangle_{a\,a}\langle n| \\ \hat{\rho}_b &= \sum_{n=0}^{\infty} \frac{1}{(\cosh r)^2} (\tanh r)^{2n} |n\rangle_{b\,b}\langle n|.\end{aligned} \tag{7.181}$$

In both cases, the probability of finding n photons in a single mode a or b is

$$P_n^{(i)} = {}_i\langle n|\hat{\rho}_i|n\rangle_i = \frac{(\tanh r)^{2n}}{(\cosh r)^2} \quad (i = a, b). \tag{7.182}$$

From Eq. (7.171) we can write

$$P_n^{(i)} = \frac{\langle \hat{n}_i \rangle^n}{(1 + \langle \hat{n}_i \rangle^{n+1})}, \tag{7.183}$$

which is a thermal distribution with an average photon number $\langle \hat{n}_i \rangle = \sinh^2 r$. In other words, the radiation field in one of the modes obtained by disregarding (i.e. making no measurement on) the other mode of the two-mode squeezed vacuum, is indistinguishable from the radiation from a thermal source [24]. The effective temperature of this "source" can be obtained from Eq. (2.141) by replacing $\bar{n}$ with $\langle \hat{n}_i \rangle$ to yield

$$T_{\text{eff}} = \frac{\hbar\omega_i}{2k_{\text{B}} \ln(\coth r)} \tag{7.184}$$

where ω_i is the frequency of the mode.

The alert reader will have noted that the expansion in Eq. (7.169) is already of the form of a Schmidt decomposition, in this case containing an infinite number of terms. Furthermore, we already have the reduced density operators of both modes and they are identical in form. Thus we can use the results in Appendix A to obtain the von Neumann entropy, a measure of the degree of entanglement, as

$$S(\hat{\rho}_a) = -\sum_{n=0}^{\infty} \left[\frac{\tanh^{2n} r}{\cosh^2 r}\right] \ln\left[\frac{\tanh^{2n} r}{\cosh^2 r}\right] = S(\hat{\rho}_b). \tag{7.185}$$

In Fig. 7.18 we plot the von Neumann entropy against the squeeze parameter r. Evidently, the degree of entanglement increases with increasing r.

Our two-mode squeezed vacuum state, as for the single-mode squeezed vacuum state, may be generated by a parametrically driven nonlinear medium. The form of the Hamiltonian for this case, with all modes quantized is

$$\hat{H} = \hbar\omega_a \hat{a}^\dagger \hat{a} + \hbar\omega_b \hat{b}^\dagger \hat{b} + \hbar\omega_{\text{P}} \hat{c}^\dagger \hat{c} + i\hbar\chi^{(2)}(\hat{a}\hat{b}\hat{c}^\dagger - \hat{a}^\dagger \hat{b}^\dagger \hat{c}), \tag{7.186}$$

where the c mode contains the pump field. Again we assume the pump field to be a strong, undepleted, coherent state $|\gamma e^{-i\omega_{\text{P}} t}\rangle$. The parametric approximation

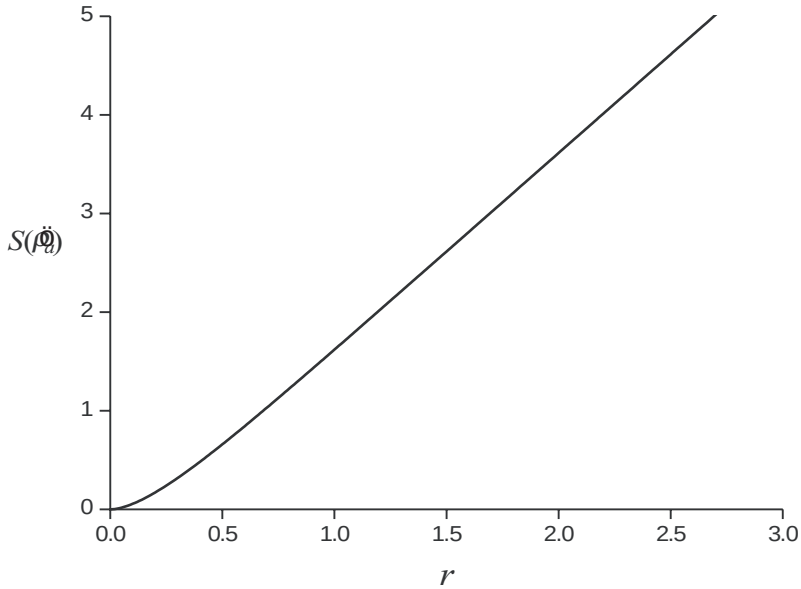


Fig. 7.18. A plot of the von Neumann entropy $S(\hat{\rho}_a)$ against the squeeze parameter r.

to Eq. (7.186) is then

$$\hat{H}^{(\mathrm{PA})} = \hbar\omega_a \hat{a}^\dagger \hat{a} + \hbar\omega_b \hat{b}^\dagger \hat{b} + i\,\hbar(\eta^* e^{i\omega_\mathrm{P} t} \hat{a}\hat{b} - \eta e^{-i\omega_\mathrm{P} t} \hat{a}^\dagger \hat{b}^\dagger), \tag{7.187}$$

where $\eta = \chi^{(2)}\gamma$. Transforming to the interaction picture we obtain the time-dependent interaction Hamiltonian

$$\hat{H}_\mathrm{I}(t) = i\,\hbar\Big[\eta^* e^{i\,(\omega_\mathrm{P}-\omega_a-\omega_b)\,t} \hat{a}\hat{b} - \eta e^{-i\,(\omega_\mathrm{P}-\omega_a-\omega_b)\,t} \hat{a}^\dagger \hat{b}^\dagger\Big]. \tag{7.188}$$

With the pump field frequency adjusted so that $\omega_\mathrm{P} = \omega_a + \omega_b$, we obtain the time-independent interaction Hamiltonian

$$\hat{H}_\mathrm{I} = i\,\hbar(\eta^* \hat{a}\hat{b} - \eta \hat{a}^\dagger \hat{b}^\dagger). \tag{7.189}$$

This describes the process of *nondegenerate parametric down-conversion* whereby a photon of the pump field is converted into two photons, one for each of the modes a and b. Obviously the associated evolution operator is

$$\hat{U}_\mathrm{I}(t, 0) = \exp[-i\,\hat{H}_\mathrm{I} t/\hbar] = \hat{S}_2(\xi) \tag{7.190}$$

where $\xi = \eta t$.

The states produced by the nondegenerate down-conversion process (i.e. the two-mode squeezed vacuum states) are entangled and, as such, they have been much discussed and have even been employed experimentally for fundamental tests of quantum mechanics, some aspects of which will be addressed in Chapter 9. They have also been studied in connection with various applications in quantum information processing, notably for purposes of quantum teleportation [25].

7.8 Higher-order squeezing

We have by no means exhausted all possibilities when it comes to nonclassical states and criteria under which the states are, in fact, nonclassical. It is possible to have states that exhibit quadrature squeezing but not number squeezing, or vice versa. It is even possible to have nonclassical states that are neither quadrature nor number squeezed.

Here we shall introduce one form of so-called higher-order squeezing, specifically the form introduced by Hong and Mandel [26]. Other forms will be addressed in the problems. The Hong–Mandel higher-order squeezing refers to squeezing in the higher-order moments of a quadrature operator. Setting

$$\hat{X}(\vartheta) = \frac{1}{2}(\hat{a}\, e^{-i\vartheta} + \hat{a}^{\dagger} e^{i\vartheta}) \tag{7.191}$$

as our generic quadrature operator, we say that a field is squeezed to the Nth order if for some angle ϑ, $\langle(\Delta\hat{X})^N\rangle < \langle(\Delta\hat{X})^N\rangle_{\rm coh}$, where the right-hand side of the inequality is evaluated for a coherent state. This definition is an obvious generalization of the condition for the usual second-order quadrature squeezing, $\langle(\Delta\hat{X})^2\rangle < \langle(\Delta\hat{X})^2\rangle_{\rm coh} = \frac{1}{4}$. But although this latter condition can be written as $\langle{:}(\Delta\hat{X})^2{:}\rangle < 0$, one *cannot* write $\langle{:}(\Delta\hat{X})^N{:}\rangle < 0$ as the condition for Nth-order squeezing, the relationship between the moments $\langle(\Delta\hat{X})^N\rangle$ and the normally ordered moments $\langle{:}(\Delta\hat{X})^N{:}\rangle$ being much more complicated in cases when $N > 2$. To obtain the relationship, we use the Baker–Hausdorf–Campbell theorem the reader will have encountered in Problem 3 of Chapter 2. Recall that for operators $\hat{A}$ and $\hat{B}$, which *may* not commute, i.e. $[\hat{A}, \hat{B}] \neq 0$, but that do commute with their commutator, $[\hat{A}, [\hat{A}, \hat{B}]] = 0 = [\hat{B}[\hat{A}, \hat{B}]]$, then

$$e^{\hat{A}+\hat{B}} = \exp(-\tfrac{1}{2}[\hat{A}, \hat{B}])e^{\hat{A}} e^{\hat{B}}. \tag{7.192}$$

With this formula we can show that

$$\exp(y\Delta\hat{X}) = {:}\exp(y\Delta\hat{X}){:}\exp(y^2/8) \tag{7.193}$$

where y is just a c-number parameter. Taking the expectation value of this operator relation, expanding in Taylor series and equating coefficients of like powers of y on both sides, we obtain

$$\begin{aligned}\left\langle(\Delta\hat{X})^N\right\rangle = \left\langle{:}(\Delta\hat{X})^N{:}\right\rangle &+ \frac{N^{(2)}}{1!}\left(\frac{1}{8}\right)\left\langle{:}(\Delta\hat{X})^{N-2}{:}\right\rangle \\ &+ \frac{N^{(4)}}{2!}\left(\frac{1}{8}\right)^2\left\langle{:}(\Delta\hat{X})^{N-4}{:}\right\rangle + \cdots \\ \cdots + &\begin{cases} (N-1)!!\left(\frac{1}{4}\right)^{\frac{1}{2}N-\frac{3}{2}}, & N \text{ even}, \\ \dfrac{N!}{3!2^{\frac{1}{2}N-\frac{3}{2}}}\dfrac{\left(\frac{1}{4}\right)^{\frac{1}{2}N-\frac{3}{2}}}{\left(\frac{1}{2}N-\frac{3}{2}\right)}\left\langle{:}(\Delta\hat{X})^3{:}\right\rangle, & N \text{ odd}, \end{cases}\end{aligned} \tag{7.194}$$

where $N^{(r)} = N(N-1)\ldots(N-r+1)$. For a coherent state, all the normally ordered moments vanish: $\langle{:}(\Delta\hat{X})^N{:}\rangle = 0$. So in the case of N odd,

$N = 2l + 1, l = 0, 1, 2, \ldots$, we obtain $\langle(\Delta\hat{X})^{2l+1}\rangle_{\text{coh}} = 0$. As it is not possible for the moments to go negative, there can be no higher-order squeezing in the odd moments of the quadrature operator. On the other hand, for N even, $N = 2l$, we have

$$\left\langle(\Delta\hat{X})^{2l}\right\rangle_{\text{coh}} = (2l-1)!!\left(\frac{1}{4}\right)^{l}. \tag{7.195}$$

Thus we can say that the field state is squeezed to the order $2l$ whenever

$$\left\langle(\Delta\hat{X})^{2l}\right\rangle < (2l-1)!!\left(\frac{1}{4}\right)^{l}. \tag{7.196}$$

Squeezing exists to second order if

$$\left\langle{:}(\Delta\hat{X})^{2}{:}\right\rangle < 0, \tag{7.197}$$

to fourth order if

$$\left\langle{:}(\Delta\hat{X})^{4}{:}\right\rangle + \frac{3}{2}\left\langle{:}(\Delta\hat{X})^{2}{:}\right\rangle < 0, \tag{7.198}$$

to sixth order if

$$\left\langle{:}(\Delta\hat{X})^{6}{:}\right\rangle + \frac{15}{4}\left\langle{:}(\Delta\hat{X})^{4}{:}\right\rangle + \frac{45}{16}\left\langle{:}(\Delta\hat{X})^{2}{:}\right\rangle < 0, \tag{7.199}$$

etc. There is the possibility that Eq. (7.196) is satisfied even though $\langle{:}(\Delta\hat{X})^{N}{:}\rangle$ is not negative beyond $N = 2$ owing to the domination of the series by the second-order term $\langle{:}(\Delta\hat{X})^{2}{:}\rangle$. In such a case, there is only *intrinsic* second-order squeezing.

As an extension of the parameter introduced in Eq. (7.8) to characterize second-order squeezing, Nth-order squeezing may be conveniently characterized by the parameter

$$s_N(\vartheta) = \frac{\left\langle(\Delta\hat{X})^{N}\right\rangle - (N-1)!!(0.25)^{N/2}}{(N-1)!!(0.25)^{N/2}}. \tag{7.200}$$

Specific examples of states manifesting higher-order squeezing are considered in the problems below.

7.9 Broadband squeezed light

In our discussion of nonclassical light, we have, so far, confined our interest to a single-mode field or a double-mode field with widely separated frequencies. Of course, these are highly ideal situations, difficult to realize experimentally. In a more realistic approach, the presence of other modes of frequencies near the one of interest, or those in the bandwidth of the sensitivity of the detectors, must be taken into account. To give a sense of what is involved, we consider here, briefly, broadband squeezed light.

To keep things as uncluttered as possible, we assume that all the modes are identically polarized and thus treat the fields as scalars. We further assume that some frequency ω_0 is the mid frequency of a quasi-monochromatic field. The

positive frequency part of the field, from Eq. (2.125), is given by the collective annihilation operator

$$\hat{E}^{(+)}(\mathbf{r},t)=\sum_{[\mathbf{k}]} l(\omega_k)\hat{a}_{\mathbf{k}}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}, \tag{7.201}$$

where $l(\omega_k)=i(\hbar\omega_k/2\varepsilon_0 V)^{1/2}$ and where $[\mathbf{k}]$ represents the set of relevant plane-wave modes. Of course, the collective creation operator is just $\hat{E}^{(-)}(\mathbf{r},t)=[\hat{E}^{(+)}(\mathbf{r},t)]^{\dagger}$. These operators satisfy the equal time commutation relation

$$\left[\hat{E}^{(+)}(t),\hat{E}^{(-)}(t)\right]=\sum_{[\mathbf{k}]}|l(\omega_k)|^2\equiv C, \tag{7.202}$$

where we have suppressed the position vector for simplicity. We may now define two collective quadrature operators according to

$$\begin{aligned}\hat{X}_1^{(c)}(t)&=\tfrac{1}{2}\left[\hat{E}^{(+)}(t)+\hat{E}^{(-)}(t)\right],\\ \hat{X}_2^{(c)}(t)&=\tfrac{1}{2i}\left[\hat{E}^{(+)}(t)-\hat{E}^{(-)}(t)\right],\end{aligned} \tag{7.203}$$

satisfying

$$\left[\hat{X}_1^{(c)},\hat{X}_2^{(c)}\right]=\frac{i}{2}C, \tag{7.204}$$

from which it follows that

$$\left\langle\left(\Delta\hat{X}_1^{(c)}\right)^2\right\rangle\left\langle\left(\Delta\hat{X}_2^{(c)}\right)^2\right\rangle\geq\frac{1}{16}|C|^2. \tag{7.205}$$

Broadband squeezing exists then if

$$\left\langle\left(\Delta\hat{X}_i^{(c)}\right)^2\right\rangle<\frac{1}{4}|C|,\ i=1\text{ or }2. \tag{7.206}$$

Other forms of higher-order squeezing have been proposed, for states of both single- and double-mode fields. Some of these are taken up in the Problems.

Problems

1. Investigate the time evolution of the wave packets associated with the squeezed vacuum state and of the more general squeezed state. Graphically display the wave packets (you should actually consider the probability distribution over the harmonic oscillator potential) at different times. Show that the squeezed vacuum state is a "breather", that is, that its centroid is stationary but that its width oscillates at twice the frequency of the harmonic oscillator potential.
2. Carry out the steps to obtain the squeezed vacuum Wigner function of Eq. (7.83).
3. Obtain the Wigner function for the displaced squeezed vacuum.
4. We defined a general squeezed state in Eq. (7.22) as a displaced squeezed vacuum state, i.e. the state generated by the operator product $\hat{D}(\alpha)\hat{S}(\xi)$ acting on the vacuum. Consider now the state generated by the reversed operator product, i.e. $\hat{S}(\xi)\hat{D}(\alpha)$, acting on the vacuum, the squeezed coherent state. Find the appropriate eigenvalue problem and solve it via a recursion relation. Investigate the properties of the state, e.g. the

photon statistics, and quadrature and number squeezing. How is this state related to the displaced squeezed vacuum?

5. Operating on a coherent state with the annihilation operator has no effect on the state itself, i.e. $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$. But consider now the effect of operating on a coherent state with the creation operator: $\hat{a}^\dagger|\alpha\rangle$. Obtain the normalized state and investigate its properties. Is it a nonclassical state?
6. Investigate (numerically!) the question of whether or not squeezing (both in quadrature and in number) occurs in the dynamics of the Jaynes–Cummings model. Assume the atom to always be initially in the excited state and the field initially in a coherent state. Examine the effects of increasing the average photon number from as low as $\bar{n} = 5$ to as high as $\bar{n} = 100$.
7. Study the dynamics of the Jaynes–Cummings model where the initial coherent state of the field is replaced by the squeezed state of Eq. (7.80). Check the limiting cases where $r \to 0$ (the coherent state) and where $\alpha \to 0$ (the squeezed vacuum state of Eq. (7.68)).
8. Consider the Kerr interaction as described by Eq. (7.144).
   (a) Obtain the Heisenberg equation for the annihilation operator and solve it.
   (b) Show that for an initial coherent state, the photon-counting statistics remain Poissonian for all times.
   (c) Examine the possible occurrence of quadrature squeezing. We know that for times $t = \pi/2K$, the condition for generating Yurke–Stoler states, that squeezing will be present in $\hat{X}_2$ as per Eq. (7.131), but what about at other, shorter, times? Make sure to examine both quadratures.
9. Consider the so-called "real" and "imaginary" states consisting of the superpositions of the forms $|\alpha\rangle \pm |\alpha^*\rangle$. Assuming α is complex, normalize the superposition and then investigate the states for quadrature and number squeezing. Obtain the Wigner function for these states.
10. Consider the set of operators defined as

$$\hat{K}_1 = \frac{1}{2}(\hat{a}^{\dagger 2} + \hat{a}^2), \quad \hat{K}_2 = \frac{1}{2i}(\hat{a}^{\dagger 2} - \hat{a}^2), \quad \text{and} \quad K_3 = \frac{1}{2}\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right).$$

    The first two are sometimes referred to as the quadratures of the square of the field [27]. (a) Show that these operators form a closed set under commutation; that is, the commutator of any two yields the third. (b) Obtain the uncertainty product for the operators K_1 and K_2. (c) Use the coherent states $|\alpha\rangle$ to see if they equalize this product. (d) Examine the question of squared field squeezing as a form of higher-order squeezing: is it a nonclassical effect? (e) If this is a nonclassical effect, check to see if the Schrödinger cat states, either the even, odd, or Yurke–Stoler types, are squeezed in this sense.
11. Justify the statement that all normally ordered moments of the quadrature operator vanish, i.e. that $\langle :(\Delta\hat{X})^N: \rangle = 0$ for $\hat{X}$ given by Eq. (7.191).
12. Derive Eq. (7.194).
13. Establish, through the use of the P function, that the Hong–Mandel definition of higher-order squeezing is, indeed, a nonclassical effect.
14. Work out the conditions for higher-order squeezing in a broadband field.

15. Consider the pair coherent states $|\eta, q\rangle$, which are defined as simultaneous eigenstates of the pair annihilation operator $\hat{a}\,\hat{b}$ with eigenvalue η (a complex number) and the number difference operator $\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}$ with eigenvalue $q = \ldots, -2, -1-, 0, 1, 2, \ldots$. Solve the eigenvalue problem under the assumption that $q \geq 0$. Then examine the properties of the states, specifically the joint photon probability distribution, the intermode correlations, and two-mode squeezing.
16. Consider the inequality $\langle \hat{a}^{\dagger 2}\hat{a}^2\rangle\langle \hat{b}^{\dagger 2}\hat{b}^2\rangle \geq \langle \hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}\rangle^2$. This is a form of the Cauchy–Schwarz inequality [28]. For a product of coherent states, it is easy to see that the *equality* is satisfied. Violation of this inequality indicates a strong nonclassical correlation. Do the two-mode squeezed vacuum states violate the inequality? What about the pair coherent states of Problem 15?
17. Obtain the von Neumann entropy in the case of pair coherent states. Examine how the degree of entanglement varies with the eigenvalue η and the parameter q.
18. Consider a Mach–Zehnder interferometer, as described in Chapter 6, and assume that one input mode is in a coherent state and that the other is in a squeezed vacuum state. Show that the phase fluctuations at the output are reduced below the standard quantum limit ($\Delta\varphi_{\rm SQL} = 1/\sqrt{\bar{n}}$) to $\Delta\varphi = e^{-r}/\sqrt{\bar{n}}$.
19. Reconsider the previous problem but with the squeezed vacuum replaced by even or odd coherent states. What effect, if any, do these states have on the reduction of noise at the output of the interferometer?
20. Consider the Hamiltonian

$$\hat{H} = \hbar\omega_{\rm P}\,\hat{a}^\dagger\hat{a} + \hbar\omega_b\hat{b}^\dagger\hat{b} + \hbar\omega_c\hat{c}^\dagger\hat{c} + i\,\hbar\,\chi^{(2)}(\hat{a}\hat{b}\hat{c}^\dagger - \hat{a}^\dagger\hat{b}^\dagger\hat{c}),$$

which is similar to that of Eq. (7.186) but where we have taken the a mode as the pump. In this situation the nonlinear medium will perform *frequency conversion*.

(a) Make the parametric approximation and show that, under the condition that $\omega_{\rm P} = \omega_b - \omega_c$, the interaction picture Hamiltonian has the form

$$\hat{H}_{\rm I} = -i\,\hbar(\eta\hat{b}^\dagger\hat{c} - \eta^*\hat{b}\hat{c}^\dagger).$$

(b) Using the evolution operator $\hat{U}_{\rm fc}(t) = \exp(-i\,\hat{H}_{\rm I}t/\hbar) = \exp[t(\eta\hat{b}^\dagger\hat{c} - \eta^*\hat{b}\hat{c}^\dagger)]$, obtain expressions for the $\hat{b}$ and $\hat{c}$ operators at times $t > 0$.

(c) It should be clear that under this interaction, a vacuum state, i.e. $|0\rangle_b|0\rangle_c$, remains a vacuum state. But what about a state of the form $|0\rangle_b|N\rangle_c$? (Hint: start with the relation $\hat{b}|0\rangle_b|N\rangle_c = 0$ and obtain an eigenvalue problem as we did for nondegenerate down-conversion.) Show that the joint photon number distribution is a binomial distribution of the N photons over the two modes. Is this an entangled state? Are there intermode correlations for this state?

(d) Consider the initial state to be $|0\rangle_b|\alpha\rangle_c$. Are there intermode correlations for the state at times $t > 0$?

References

[1] J. R. Oppenheimer as paraphrased by B. R. Frieden, *Probability, Statistical Optics, and Data Testing* (Berlin: Springer, 1991), p. 363.

[2] M. Hillery, *Phys. Lett. A*, **111** (1985), 409.

[3] Early papers on squeezing in optical systems are: B. R. Mollow and R. J. Glauber, *Phys. Rev*., **160** (1967), 1076; *Phys. Rev.,* **160** (1967), 1097; D. Stoler, *Phys. Rev. D*, **1** (1970), 3217; E. C. Y. Lu, *Lett. Nuovo Cimento*, **2** (1971), 1241; H. P. Yuen, *Phys. Rev. A*, **13** (1976), 2226.

[4] P. Hello, in *Progress in Optics* XXXVIII, E. Wolf (editor) (Amsterdam: Elsevier, 1998), p. 85.

[5] B. Caron *et al*., *Class. Quantum. Grav*., **14** (1997), 1461.

[6] See Y. Yamamoto and H. A. Haus, *Rev. Mod. Phys*., **58** (1986), 1001, and references therein.

[7] R. E. Slusher and B. Yurke, *IEEE J. Lightwave Tech*., **8** (1990), 466. See also E. Giacobino, C. Fabre and G. Leuchs, *Physics World*, **2** (Feb. 1989), 31.

[8] W. Schleich and J. A. Wheeler, *Nature*, **326** (1987), 574; *J. Opt. Soc. Am. B*, **4** (1987), 1715.

[9] W. Schleich, *Quantum Optics in Phase Space* (Berlin: Wiley-VCH, 2001); see chapters 7 and 8.

[10] R. L. Hudson, *Rep. Math. Phys*., **6** (1974), 249.

[11] R. W. Boyd, *Nonlinear Optics*, 2nd edition (New York: Academic Press, 2003).

[12] R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz and J. F. Valley, *Phys. Rev. Lett.*, **55** (1985), 2409.

[13] L.-A. Wu, H. J. Kimble, J. L. Hall and H. Wu, *Phys. Rev. Lett.*, **57** (1987), 2520.

[14] G. Breitenbach, T. Mueller, S. F. Pereira, J.-Ph. Poizat, S. Schiller and J. Mlynek, *J. Opt. Soc. Am. B*, **12** (1995), 2304.

[15] R. Short and L. Mandel, *Phys. Rev. Lett*., **51** (1983), 384.

[16] X. T. Zou and L. Mandel, *Phys. Rev. A*, **41** (1990), 475.

[17] H. J. Carmichael and D. F. Walls, *J. Phys.*, **9B** (1976), L43; C. Cohen-Tannoudji, in *Frontiers in Laser Spectroscopy*, edited by R. Balian, S. Haroche and S. Liberman (Amsterdam: North Holland, 1977); H. J. Kimble and L. Mandel, *Phys. Rev. A*, **13** (1976), 2123.

[18] H. J. Kimble, M. Dagenais and L. Mandel, *Phys. Rev. Lett*., **39** (1977), 691.

[19] E. Schrödinger's original articles on the so-called "cat paradox", originally published in *Naturwissenschaften* in 1935, can be found, in an English translation, in *Quantum Theory of Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton: Princeton University Press, 1983), pp. 152–167.

[20] V. V. Dodonov, I. A. Malkin and V. I. Man'ko, *Physica*, **72** (1974), 597. See also V. Bužek, A. Vidiella-Barranco and P. L. Knight, *Phys. Rev. A*, **45** (1992), 6570; C. C. Gerry, *J. Mod. Opt*., **40** (1993), 1053.

[21] B. Yurke and D. Stoler, *Phys. Rev. Lett*., **57** (1986), 13.

[22] N. Imoto, H. A. Haus and Y. Yamamoto, *Phys. Rev. A*, **32** (1985), 2287; M. Kitagawa and Y. Yamamoto, *Phys. Rev. A*, **34** (1986), 3974.

[23] C. C. Gerry, *Phys. Rev. A*, **59** (1999), 4095.

[24] S. M. Barnett and P. L. Knight, *J. Opt. Soc. Am. B*, **2** (1985), 467; B. Yurke and M. Potasek, *Phys. Rev. A*, **36** (1987), 3464; S. M. Barnett and P. L. Knight, *Phys. Rev. A*, **38** (1988), 1657.

[25] See, for example, P. T. Cochrane, G. J. Milburn and W. J. Munro, *Phys. Rev. A*, **62** (2000), 307.

[26] C. K. Hong and L. Mandel, *Phys. Rev. Lett*., **54** (1985), 323; *Phys. Rev. A*, **32** (1985), 974.

[27] M. Hillery, *Phys. Rev. A*, ***36*** (1987), 3796.
[28] R. Loudon, *Rep. Prog. Phys.*, **43** (1980), 58.

Bibliography

Below is a selection of early to recent review articles on nonclassical light.

D. F. Walls, "Evidence for the quantum nature of light", *Nature*, **280** (1979), 451.
D. F. Walls, "Squeezed states of light", *Nature*, **306** (1983), 141.
L. Mandel, "Non-classical states of the electromagnetic field", *Physica Scripta*, **T12** (1985), 34.
R. W. Henry and S. C. Glotzer, "A squeezed-state primer", *Am. J. Phys.*, **56** (1988), 318.
J. J. Gong and P. K. Aravind, "Expansion coefficients of a squeezed coherent state in a number state basis", *Am. J. Phys*., **58** (1990), 1003.
M. C. Teich and B. E. A. Saleh, "Tutorial: Squeezed states of light", *Quantum Opt*., **1** (1989), 153.
P. L. Knight, "Quantum fluctuations in optical systems", lectures in *Les Houches, Session LXIII, 1995, Quantum Fluctuations*, edited by S. Reynaud, E. Giacobini and J. Zinn-Justin (Amsterdam: Elsevier, 1997).
V. Bužek and P. L. Knight, "Quantum interference, superposition states of light, and nonclassical effects", in *Progress in Optics XXXIV*, edited by E. Wolf (Amsterdam: Elsevier, 1995), p. 1.
L. Davidovich, "Sub-Poissonian processes in quantum optics", *Rev. Mod. Phys*., **68** (1996), 127.
D. N. Klyshko, "The nonclassical light", *Physics-Uspeki*, **39** (1996), 573.
C. C. Gerry and P. L. Knight, "Quantum superpositions and Schrödinger cat states in quantum optics", *Am. J. Phys*., **65** (1997), 964.

A recent book that reviews many aspects of nonclassical states of light, including their connection, at least for some states, to some simple Lie groups, namely SU(2) *and* SU(1,1) *is*:

V. V. Dodonov and V. I. Man'ko (editors), *Theory of Nonclassical States of Light* (London: Taylor and Francis, 2003).

A book dedicated to a discussion of experimental techniques in quantum optics, including the generation of squeezed light, is:

H.-A. Bachor, *A Guide to Experiments in Quantum Optics* (Weinhein: Wiley-VCH, 1998).

The application of squeezed light to interferometry is discussed by:

C. M. Caves, "Quantum-mechanical noise in an interferometer", *Phys. Rev. D*, **23** (1981), 1693.

Another, more recent, application of nonclassical light is in the area of quantum imaging. See:

M. I. Kolobov, "The spatial behavior of nonclassical light", *Rev. Mod. Phys*., **71** (1999), 1539.

Chapter 8
Dissipative interactions and decoherence

8.1 Introduction

So far, we have discussed closed systems involving a single quantized mode of the field interacting with atoms, as for example in the Jaynes–Cummings model in Chapter 4. As we saw in this model, the transition dynamics are coherent and reversible: the atom and field mode exchange excitation to and fro without loss of energy. As we add more modes for the atom to interact with, the coherent dynamics become more complicated as the relevant atom–field states come in and out of phase and beat together to determine the total state occupation probabilities. As time goes on, these beats get out of phase, leading to an apparent decay of the initial state occupation probability. But at later times, the beating eigenfrequencies get back in phase in a manner rather reminiscent of the Jaynes–Cummings revival discussed earlier in this book, and this leads to a partial recurrence or revival of the initial state probability. The time scale for this partial revival depends on the number of participating electromagnetic field modes and as these increase to the level appropriate for an open system in free space the recurrence disappears off to the remote future, and the exponential decay law appropriate for decay is recovered as an excellent approximation [1].

We have already discussed the origin of spontaneous emission and the Einstein A coefficient using perturbation theory in Chapter 4. For many purposes, all that we need to do to include decay in our discussion of the dynamics of coherently excited atoms is to add a loss term to upper-state populations, a term which reflects the gain in population in lower states to which the excited atom can decay, and to allow for an appropriate decay for the coherences (or dipole moments). The resultant equations are known as the optical Bloch equations [2]. The dynamics is often well approximated by an ansatz which makes the upper unstable state energy complex, with an imaginary part which reflects the lifetime of the upper state (but care needs to be exercised in the use of this simple ansatz applied to probability amplitudes as it cannot allow for the repopulation of the lower level in any convincing way).

More rigorous models of spontaneous emission decay have to take into account the quantized nature of the free-space vacuum with which the atom interacts. One

can develop a multimode Schrödinger picture approach to this problem, which involves coupled equations of motion for the initial excited state of the atom with no photons present and the infinite number of states where there is one excitation in the field and the atom is in its ground state. These equations cannot be solved analytically, but when the coupling between the atom and field is weak enough for only small changes in excitation probabilities to be generated within the coherence time of the relevant radiation (and for the vacuum this is a very short time indeed!), the exponential decay law is obtained. This is the essence of the standard theory of spontaneous emission due to Weisskopf and Wigner; if instead we worked with the density matrix, the equivalent approach is called the Born–Markov approximation. An account of these models is beyond the scope of this book but details of these can be found in the bibliography.

8.2 Single realizations or ensembles?

Quantum mechanics is usually introduced as a theory for ensembles, but the invention of ion traps, for example, offers the possibility to observe and manipulate single particles, where observability of quantum jumps, which are not be seen directly in the ensemble, lead to conceptual problems of how to describe single realizations of these systems [3]. Usually Bloch equations or Einstein rate equations are used to describe the time evolution of ensembles of atoms or ions driven by light. New approaches via conditional time evolution, given say when no photon has been emitted, have been developed to describe single experimental realizations of quantum systems. This leads to a description of the system via wave functions instead of density matrices. This conditional "quantum trajectory" approach [4] is still an ensemble description but for a subensemble, where we know when photons have been emitted. The jumps that occur in this description can be considered as the result of the increase of our knowledge about the system which is represented by the wave function (or the density operator) describing the system. In the formalism to be presented one usually imagines that Gedanken measurements are performed in a rapid succession, for example, on the emitted radiation field. These will either have the result that a photon has been found in the environment or that no photon has been found. A sudden change in our information about the radiation field (for example through the detection of a photon emitted by the system into the environment) leads to a sudden change of the wave function of the system. However, it is not only the detection of a photon that leads to an increase of information, but so does the failure to detect a photon (i.e. a null result). New insights have been obtained into atomic dynamics and into dissipative processes, and new powerful theoretical approaches developed. Apart from the new insights into physics these methods also allow the simulation of complicated problems, e.g. in laser cooling, that were completely intractable with the master equation approach.

Quantum mechanics is a statistical theory that makes probabilistic predictions of the behaviour of ensembles (an ideally infinite number of identically prepared quantum systems) by using density operators. This description was completely sufficient for the first 60 years of the existence of quantum mechanics because it was generally regarded as completely impossible to observe and manipulate single quantum systems. For example, Schrödinger, in 1952, wrote [5]:

> *. . . we never experiment with just one electron or atom or (small) molecule. In thought experiments we sometimes assume that we do; this invariably entails ridiculous consequences . . . In the first place it is fair to state that we are not experimenting with single particles, any more than we can raise ichthyosauria in the zoo.*

This (rather extreme) opinion was challenged by a remarkable idea of Dehmelt, which he first made public in 1975 [6]. He considered the problem of high-precision spectroscopy, where one wants to measure the transition frequency of an optical transition as accurately as possible, e.g. by observing the resonance fluorescence from that transition as part (say) of an optical frequency standard. However, the accuracy of such a measurement is fundamentally limited by the spectral width of the observed transition. The spectral width is caused by spontaneous emission from the upper level of the transition, which leads to a finite lifetime τ of the upper level. Basic Fourier considerations then imply a spectral width of the scattered photons on the order of $1/\tau$. To obtain a precise value of the transition frequency, it would therefore be advantageous to excite a metastable transition that scatters only a few photons within the measurement time. On the other hand, one then has the problem of detecting these few photons and this turns out to be practically impossible by direct observation. So obviously one has arrived at a major dilemma. Dehmelt's proposal, however, suggests a solution to these problems, provided one would be able to observe and manipulate single ions or atoms, which became possible with the invention of single ion traps [7]. We illustrate Dehmelt's idea in its original simplified rate equation picture. It runs as follows. Instead of observing the photons emitted on the metastable two-level system directly, he proposed to use an optical double-resonance scheme as depicted in Fig. 8.1.

One laser drives the metastable $0 \leftrightarrow 2$ transition while a second strong laser saturates the strong $0 \leftrightarrow 1$; the lifetime of the upper level 1 is, for example, 10 ns while that of level 2 is of the order of 1s. If the initial state of the system is the lower state 0 then the strong laser will start to excite the system to the rapidly decaying level 1, which will then lead to the emission of a photon after a time that is usually very short (of the order of the lifetime of level 1). This emission restores the system back to the lower level 0; the strong laser can start to excite the system again to level 1, which will emit a photon on the strong transition again. This procedure repeats until at some random time the laser on the weak transition

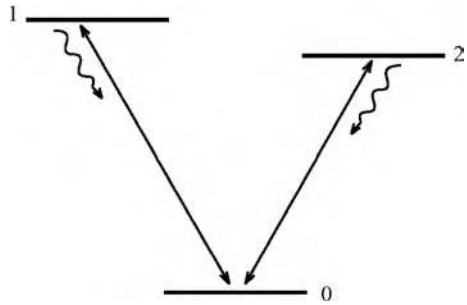


Fig. 8.1. The *V* system. Two upper levels 1 and 2 couple to a common ground state 0. The transition frequencies are assumed to be largely different so that each of the two lasers driving the system couples to only one of the transitions. The 0 ↔ 1 transition is assumed to be strong while the 0 ↔ 2 transition is weak.

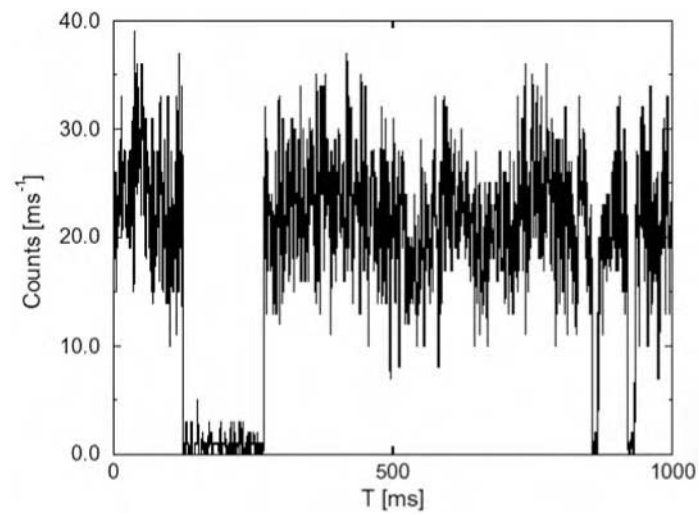


Fig. 8.2. Recorded resonance fluorescence signal exhibiting quantum jumps from a laser-excited Mg ion [8]. Periods of high photon count rate are interrupted by periods with negligible count rate (except for an unavoidable dark count rate).

manages to excite the system into its metastable state 2 where it remains shelved for a long time, until it jumps back to the ground state either by spontaneous emission or by stimulated emission due to the laser on the 0 ↔ 2 transition. During the time the electron remains in the metastable state 2, no photons will be scattered on the strong transition and only when the electron jumps back to state 0 can the fluorescence on the strong transition start again. Therefore from the switching on and off of the resonance fluorescence on the strong transition (which is easily observable) we can infer the extremely rare transitions on the 0 ↔ 2 transition. Therefore we have a method to monitor rare quantum jumps (transitions) on the metastable 0 ↔ 2 transition by observation of the fluorescence from the strong 0 ↔ 1 transition. A typical experimental fluorescence signal is depicted in Fig. 8.2 [8] where the fluorescence intensity $I(t)$ is plotted.

However, this scheme works only if we observe a single quantum system, because if we observe a large number of systems simultaneously the random nature of the transitions between levels 0 and 2 implies that some systems will be able to scatter photons on the strong transition while others are not able to do so because they are in their metastable state at that moment. From a large collection of ions observed simultaneously one would then obtain a more or less constant intensity of photons emitted on the strong transition. The calculation of single system properties, such as the distribution of the lengths of the periods of strong fluorescence, required some effort that eventually led to the development of the quantum-jump approach. Apart from the interesting theoretical implications for the study of individual quantum systems, Dehmelt's proposal obviously has important practical applications. An often-cited example is the realization of a new time standard using a single atom in a trap. The key idea here is to use either the instantaneous intensity or the photon statistics of the emitted radiation on the strong transition (the statistics of the bright and dark periods) to stabilize the frequency of the laser on the weak transition. This is possible because the photon statistics of the strong radiation depends on the detuning of the laser on the weak transition. Therefore a change in the statistics of bright and dark periods indicates that the frequency of the weak laser has shifted and has to be adjusted. However, for continuously radiating lasers this frequency shift will also depend on the intensity of the laser on the strong transition [9]. Therefore in practice pulsed schemes are preferable for frequency standards [10]. Owing to the inability of experimentalists to store, manipulate and observe single quantum systems (ions) at the time of Dehmelt's proposal, both the practical as well as the theoretical implications of his proposal were not immediately investigated. It was about ten years later that this situation changed. At that time Cook and Kimble published a paper [11] in which they made the first attempt to analyse the situation described above theoretically. Their advance was stimulated by the fact that by that time it had become possible to actually store *single* ions in an ion trap. In their simplified rate-equation approach Cook and Kimble started with the rate equations for an incoherently driven three-level system as shown in Fig. 8.1 and assumed that the strong $0 \leftrightarrow 1$ transition is driven to saturation. (By saturation we mean that the populations of the states 0 and 1 are so strongly influenced by the stimulating radiation field that no further increase in the transition probabilities can be achieved by increasing the intensity.) They consequently simplify their rate equations, introducing the probabilities P_+ of being in the metastable state and P_- of being in the strongly fluorescing $0 \leftrightarrow 1$ transition. This simplification now allows the description of the resonance fluorescence to be reduced to that of a two-state random telegraph process. Either the atomic population is in the levels 0 and 1 and therefore the ion is strongly radiating (on), or the population rests in the metastable level 2 and no fluorescence is observed (off). They then proceed to calculate the distributions for the lengths of bright and dark periods and find that their distribution is Poissonian. Their

analysis that we have outlined very briefly here is of course very much simplified in many respects. The most important point is certainly the fact that Cook and Kimble assume incoherent driving and therefore adopt a rate-equation model. In a real experiment coherent radiation from lasers is used. The complications arising in coherent excitation finally led to the development of the quantum-jump approach. Despite these problems the analysis of Cook and Kimble showed the possibility of direct observation of quantum jumps in the fluorescence of single ions, a prediction that was confirmed shortly afterwards in a number of experiments [12]. The following effort of a great number of physicists eventually culminated in the development of the quantum-jump approach [13].

8.3 Individual realizations

The quantum-jump method provides a simple description of decay processes which avoids the technicalities of more formal approaches. The procedure adopted in quantum-jump simulations, a kind of Monte Carlo procedure, can be summarized as follows [3, 14] for the simple case of photons in a single-mode cavity field, where photons are lost from the field and absorbed by the walls of the cavity, which we assume to be at $T = 0$ K. When we speak of "emission" in what follows, we shall mean emission *from the field*, the loss of photons from the field, the emission being represented mathematically by the action of an annihilation operator acting on the field state. The rate that photons are lost by the field (the emission rate) we denote as γ. Then the jump procedure is as follows.

1. Determine the current probability ΔP of emission of a photon within the time interval δt: this depends on the decay rate and the photon occupation number according to
$$\Delta P = \gamma \langle \Psi | \hat{a}^\dagger \hat{a} | \Psi \rangle \, \delta t, \tag{8.1}$$
$|\Psi\rangle$ being a particular member of the ensemble.
2. Obtain a random number r between zero and one, compare this with ΔP and decide on emission as follows.
3. Emit if $r < \Delta P$; so that the system jumps to the renormalized form:
$$|\Psi\rangle \to |\Psi_{\text{emit}}\rangle = \frac{\hat{a} |\Psi\rangle}{\langle \Psi | \hat{a}^\dagger \hat{a} | \Psi \rangle^{1/2}}, \tag{8.2}$$
where the annihilation operator takes care of the loss of a cavity field photon emitted by the wall external to the cavity.
4. Or no emission if $r > \Delta P$; so the system evolves according to
$$\begin{aligned} |\Psi\rangle \to |\Psi_{\text{no emit}}\rangle &= \frac{e^{-i\,\delta t \hat{H}_{\text{eff}}} |\Psi\rangle}{\left[\langle \Psi | e^{i\,\delta t \hat{H}^\dagger_{\text{eff}}} e^{-i\,\delta t \hat{H}_{\text{eff}}} |\Psi\rangle \right]^{1/2}} \\ &= \frac{e^{-i\,\delta t \hat{H}_{\text{eff}}} |\Psi\rangle}{[\langle \Psi | e^{-\delta t \hat{a}^\dagger \hat{a}} |\Psi\rangle]^{1/2}} \\ &\approx \frac{\{1 - (i/\hbar)\hat{H}\delta t - (\gamma/2)\delta t \, \hat{a}^\dagger \hat{a}\}|\Psi\rangle}{(1 - \Delta P)^{1/2}}, \end{aligned} \tag{8.3}$$

valid to order δt, where $\hat{H}_{\text{eff}} = \hat{H} - i\,\hbar(\gamma/2)\hat{a}^\dagger\hat{a}$ is the effective non-Hermitian Hamiltonian describing nonunitary evolution, and where the second term represents the decay in the energy of the cavity field. We refer the reader to the literature [8] for a more complete discussion of the no-jump evolution.

5. Repeat this whole process to obtain an individual trajectory, or history that will describe the conditional evolution of one member of the ensemble of decaying systems, contingent of us recording jump events at times $t_1, t_2, \ldots$
6. Average observables over many such trajectories to obtain a description of how the whole ensemble behaves on average, where all possible histories are included.

To reassure ourselves that this is all true, we note that the history of a particular trajectory splits into two alternatives in a time Δt short enough to allow at most one jump: either a jump is recorded or it is not:

$$|\Psi\rangle \rightarrow \begin{cases} |\Psi_{\text{emit}}\rangle \text{ with probability } \Delta P, \\ |\Psi_{\text{no emit}}\rangle \text{ with probability } 1 - \Delta P. \end{cases} \tag{8.4}$$

Then in terms of the density operator of the associated pure state, $|\Psi\rangle\langle\Psi|$, the evolution for a short time step of length δt, where $|\Psi(t)\rangle \equiv |\Psi\rangle$ and $|\Psi(\delta t)\rangle \equiv |\Psi(t+\delta t)\rangle$, becomes a sum of the two possible outcomes,

$$\begin{aligned} |\Psi\rangle\langle\Psi| \rightarrow\ & |\Psi(\delta t)\rangle\langle\Psi(\delta t)| \\ =\ & \Delta P\,|\Psi_{\text{emit}}\rangle\langle\Psi_{\text{emit}}| + (1-\Delta P)\,|\Psi_{\text{no emit}}\rangle\langle\Psi_{\text{no emit}}| \\ =\ & \gamma\,\delta t\,\hat{a}\,|\Psi\rangle\langle\Psi|\,\hat{a}^\dagger \\ & + \{1 - (i/\hbar)\hat{H}\delta t - (\gamma/2)\delta t\,\hat{a}^\dagger\hat{a}\}\,|\Psi\rangle\langle\Psi| \\ & \times \{1 + (i/\hbar)\hat{H}\delta t - (\gamma/2)\delta t\,\hat{a}^\dagger\hat{a}\} \\ \approx\ & |\Psi\rangle\langle\Psi| - \frac{i}{\hbar}\delta t[\hat{H}, |\Psi\rangle\langle\Psi|] \\ & + \frac{\gamma}{2}\delta t\{2\hat{a}\,|\Psi\rangle\langle\Psi|\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\,|\Psi\rangle\langle\Psi| - |\Psi\rangle\langle\Psi|\,\hat{a}^\dagger\hat{a}\}. \end{aligned} \tag{8.5}$$

Averaging over all members of the ensemble for this short time evolution we obtain

$$\hat{\rho}\,(t+\delta t) = \hat{\rho}\,(t) - \frac{i}{\hbar}\delta t[\hat{H}, \hat{\rho}] + \frac{\gamma}{2}\delta t\{2\hat{a}\hat{\rho}\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho} - \hat{\rho}\,\hat{a}^\dagger\hat{a}\}, \tag{8.6}$$

$\hat{\rho}(t)$ being the density operator of the ensemble. This result has the appearance of Euler's approximation for obtaining the solution of an ordinary differential equation. Indeed, in the limit $\delta t \rightarrow 0$, we do obtain a first-order ordinary differential equation for the density operator:

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \frac{\gamma}{2}\{2\hat{a}\hat{\rho}\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho} - \hat{\rho}\,\hat{a}^\dagger\hat{a}\}. \tag{8.7}$$

This we refer to as the master equation, and it is precisely of the form obtained from more-complicated approaches for this zero-temperature ($T = 0$ K) field damping problem.

Before proceeding, let us look at the implications of the quantum jumps on some familiar quantum states of the field. We assume that no interactions occur other than the dissipative one so that we have $\hat{H} = 0$. Suppose the cavity field is prepared in a coherent state $|\alpha\rangle$. When a quantum jump occurs and a photon is emitted by the field, then because the coherent state is a right eigenstate of the annihilation operator, $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$, the field is unchanged and thus the coherent state remains a coherent state. Then, the no-jump evolution results in

$$\frac{e^{-(\gamma/2)\delta t\,\hat{a}^\dagger\hat{a}}\,|\alpha\rangle}{\langle\alpha|\,e^{-\gamma\,\delta t\,\hat{a}^\dagger\hat{a}}\,|\alpha\rangle} = |\alpha\,e^{-\gamma\,\delta t/2}\rangle \tag{8.8}$$

for δt small. Thus it appears that the coherent state, the state we already know to be very much a classical state, remains coherent in spite of the dissipative interaction although the amplitude of the state undergoes an exponential decay.

On the other hand, suppose the cavity field has been prepared in, say, the even cat state

$$|\psi_{\rm e}\rangle = \mathcal{N}_{\rm e}\,(|\alpha\rangle + |-\alpha\rangle)\,, \tag{8.9}$$

as discussed in Chapter 7. The occurrence of a quantum jump has the following effect:

$$\hat{a}\,(|\alpha\rangle + |-\alpha\rangle) = \alpha\,(|\alpha\rangle - |-\alpha\rangle)\,.$$

The emission of a photon by the cavity field converts the even cat state into the odd cat

$$|\psi_{\rm o}\rangle = \mathcal{N}_{\rm o}\,(|\alpha\rangle - |-\alpha\rangle)\,. \tag{8.10}$$

The abrupt transition of an even cat state to an odd cat state is sometimes referred to as a "jumping cat" transition. For a no-jump transition, then the field state undergoes decay:

$$\exp(-\gamma\,\delta t\,\hat{a}^\dagger\hat{a}/2)\,(|\alpha\rangle \pm |-\alpha\rangle) = (|\alpha e^{-\gamma\,\delta t/2}\rangle \pm |-\alpha e^{-\gamma\,\delta t/2}\rangle), \tag{8.11}$$

thus showing that the cat states "shrink" during the intervals of no jumping. The jumping between the even and the odd cat states is particularly visible in the behavior of the Wigner function because the interference fringes change sign. This transformation is seen in going from Fig. 8.3(a) to 8.3(b). Subsequently, the cat state shrinks under "no jump" evolution in going to Fig. 8.3(c). After another jump, the interference fringes change sign again as shown in Fig. 8.3(d). Adding the Wigner functions of the even and odd cat states, as happens during averaging to describe the ensemble of cavity fields, causes the cancellation of the interference fringes altogether. The consequence of these uncontrollable jumps, when ensemble averaged, is the *decoherence* of a quantum superposition state into a classical statistical mixture. Decoherence will be addressed further in Section 8.5.

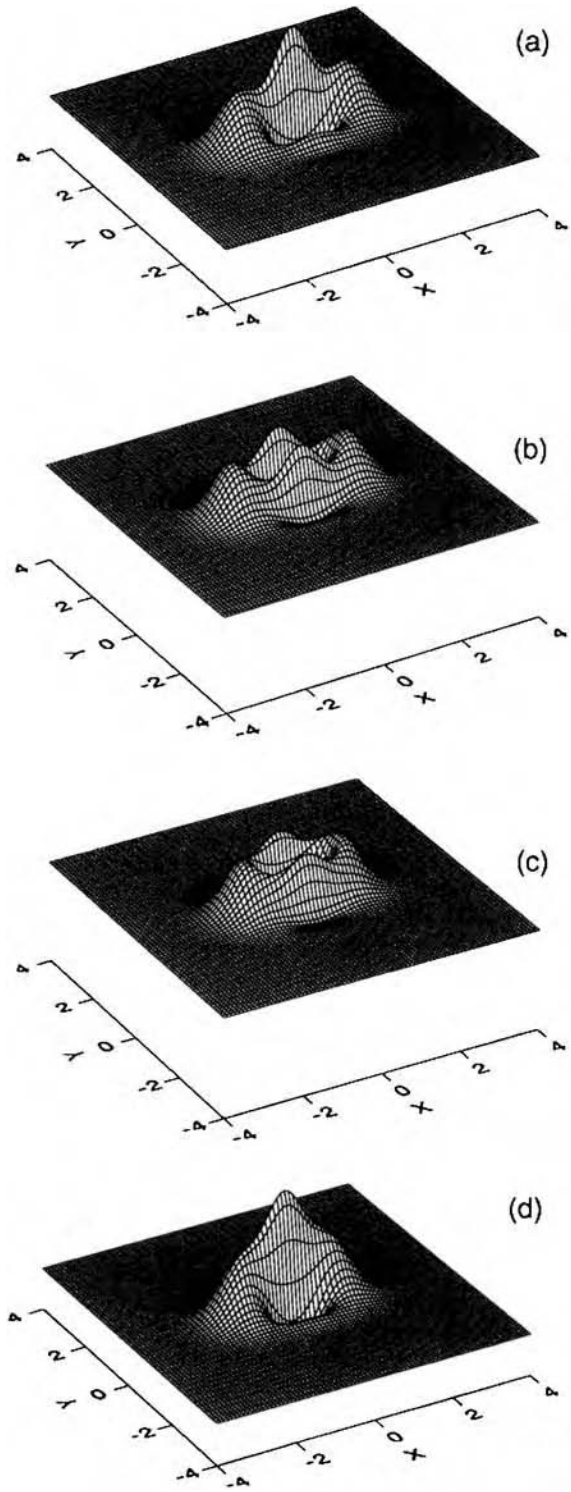


Fig. 8.3. The Wigner functions for a "jumping cat": (a) is for the initial even cat state and (b) is the odd cat obtained after the first jump. In (c) we show the Wigner function shrinking and (d) shows the result of a second jump. Notice that the jumps reverse the sign of the interference fringes.

8.4 Shelving and telegraph dynamics in three-level atoms

We now study in slightly more detail how the dynamics of the system determines the statistics of bright and dark periods. Again assume a three-level system as shown in Fig. 8.1. Provided the $0 \leftrightarrow 1$ and $0 \leftrightarrow 2$ Rabi frequencies are small compared with the decay rates, one finds for the population in the strongly fluorescing level 1 as a function of time something like the behaviour shown in Fig. 8.3 (we derive this in detail later in this chapter).

Before we develop detailed theoretical models to describe individual quantum trajectories (i.e. state evolution conditioned on a particular sequence of observed events), it is useful to examine how the entire ensemble evolves. This is in line with the historical development as initially it was tried to find quantum jump characteristics in the ensemble behaviour of the system. We do this in detail for the particular three-level V-configuration (shown in Fig. 8.1) appropriate for Dehmelt's quantum-jump phenomena. For simplicity, we examine the case of *incoherent* excitation. Studies for coherent excitation using Bloch equations can be found [15]. But for our purposes, the following Einstein rate equations for the elements of the atomic density operator demonstrate the key effects [16]:

$$\begin{aligned}
\frac{d\rho_{11}}{dt} &= -(A_1 + B_1 W_1)\,\rho_{11} + B_1 W_1 \rho_{00}, \\
\frac{d\rho_{22}}{dt} &= -(A_2 + B_2 W_2)\,\rho_{22} + B_2 W_2 \rho_{00}, \\
\frac{d\rho_{00}}{dt} &= -(B_1 W_1 + B_2 W_2)\,\rho_{00} + (A_1 + B_1 W_1)\,\rho_{11} + (A_2 + B_2 W_2)\,\rho_{22},
\end{aligned} \tag{8.12}$$

where A_i, B_i are the Einstein A and B coefficients for the relevant spontaneous and induced transitions, W_i the applied radiation field energy density at the relevant transition frequency, and ρ_{ii} is the relative population in state $|i\rangle$ where $\rho_{00} + \rho_{11} + \rho_{22} = 1$ for this *closed* system. In shelving, we assume that both $B_1 W_1$ and A_1 are much larger than $B_2 W_2$ and A_2, and furthermore that $B_2 W_2 \gg A_2$. The steady-state solutions of these rate equations are straightforward to obtain, and we find

$$\begin{aligned}
\rho_{11}(t \to \infty) &= \frac{B_1 W_1 (A_2 + B_2 W_2)}{A_1 (A_2 + 2B_2 W_2) + B_1 W_1 (2A_2 + 3B_2 W_2)}, \\
\rho_{22}(t \to \infty) &= \frac{B_2 W_2 (A_1 + B_1 W_1)}{A_1 (A_2 + 2B_2 W_2) + B_1 W_1 (2A_2 + 3B_2 W_2)}.
\end{aligned} \tag{8.13}$$

The expression for $\rho_{00}(t \to \infty)$ can be obtained from the relation $\rho_{00} + \rho_{11} + \rho_{22} = 1$, which must hold at all times. Now if the allowed $0 \leftrightarrow 1$ transition is saturated so that $B_2 W_2 \gg A_2$ we have

$$\rho_{11}(t \to \infty) \approx \rho_{00}(t \to \infty) \approx \frac{A_2 + B_2 W_2}{2A_2 + 3B_2 W_2}, \tag{8.14}$$

and

$$\rho_{22}\,(t \to \infty) \approx \frac{B_2 W_2}{2A_2 + 3B_2 W}. \tag{8.15}$$

Now we see that a small $B_2 W_2$ transition rate to the shelf state has a major effect on the dynamics. Note that if the induced rates are much larger than the spontaneous rates, the steady-state populations are $\rho_{00} = \rho_{11} = \rho_{22} = 1/3$: the populations are evenly distributed amongst the constituent states of the transition. However, the dynamics reveals a different story to that suggested by the steady-state populations. Again, if the allowed transition is saturated, then the time-dependent solutions of the excited-state rate equations tell us that for $\rho_{00}(0) = 1$ we have

$$\rho_{11}\,(t) = \frac{B_2 W_2}{2\,(2A_2 + 3B_2 W_2)} e^{-(A_2 + 3B_2 W_2/2)t} - \frac{1}{2} e^{-(2B_1 W_1 + A_1 + B_2 W_2/2)t} + \frac{A_2 + B_2 W_2}{2A_2 + 3B_2 W_2} \tag{8.16}$$

and

$$\rho_{22}(t) = \frac{B_2 W_2}{2A_2 + 3B_2 W_2} \left\{1 - e^{-(A_2 + 3B_2 W_2/2)t}\right\}. \tag{8.17}$$

Note that these expressions are good only for strong driving of the $0 \leftrightarrow 1$ transition. For a very long-lived shelf state 2, we see that for saturated transitions ($B_i W_i \gg A_i$)

$$\rho_{11}(t) \approx \frac{1}{3} \left\{1 + \frac{1}{2}\left(e^{-3B_2 W_2 t/2} - 3e^{-2B_1 W_1 t}\right)\right\} \tag{8.18}$$

and

$$\rho_{22}\,(t) \approx \frac{1}{3}\{1 - e^{-3B_2 W_2 t}\}. \tag{8.19}$$

These innocuous-looking expressions contain a lot of physics. We remember that state 1 is the strongly fluorescing state. On a time scale short compared with $(B_2 W_2)^{-1}$, we see that the populations attain a quasi-steady state appropriate to the *two-level* $0 \leftrightarrow 1$ dynamics, which for strong driving equalizes the populations in the states 0 and 1:

$$\rho_{11}\,(t,\ \text{short}) \sim \frac{1}{2}\{1 - e^{-2B_1 W_1 t}\} \to \frac{1}{2}. \tag{8.20}$$

For truly long times the third, shelving, state makes its effect and in the steady state for strong driving, the populations will equilibrate equally among the three levels so that

$$\rho_{11}\,(t, \text{long}) \sim \frac{1}{3} \tag{8.21}$$

as we see qualitatively in Figs. 8.4 and 8.5.

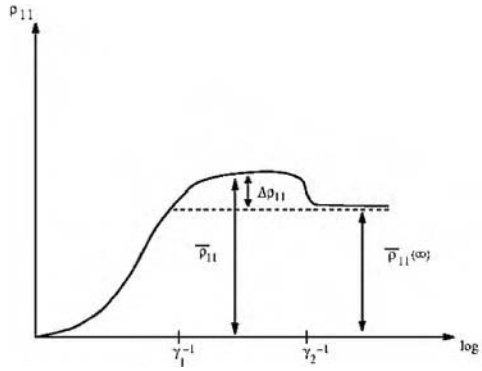


Fig. 8.4. Time-evolution of the population in the strongly fluorescing level 1 of the three-level ion shown in Fig. 8.1. The lifetimes γ_1^{-1} and γ_2^{-1} are marked on the time axis. The short-time quasi-steady-state density matrix element is denoted $\bar{\rho}_{11}$ where the transition dynamics is three-level like; $\bar{\rho}_{11}(\infty)$ is the true three-level steady-state matrix element; and $\Delta\rho_{11}$ is the difference between these two. What is crucial here is the "hump" $\Delta\rho_{11}$: this is a signature of the telegraphic nature of the fluorescence. As we saw earlier in the discussion of Eq. (8.3), the change from two-level to three-level dynamics already gives us a signature of quantum jumps and telegraphic fluorescence provided we are wise enough to recognize the signs. This figure illustrates the change from two- to three-level dynamics.

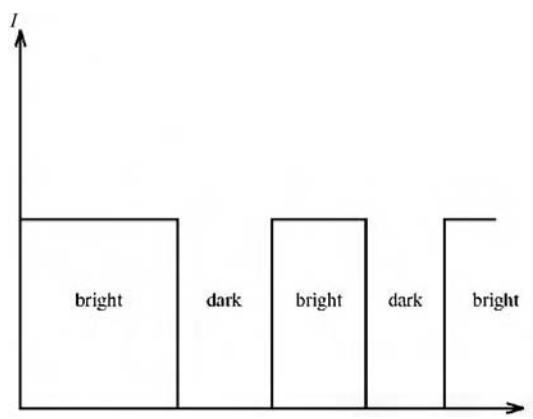


Fig. 8.5. A few periods of bright and dark sequences in the fluorescence intensity I from a three-level system. The bright periods last on average T_B; and the dark periods T_b. In the mid 1980s when studies of quantum-jump dynamics of laser-driven three-level atoms began in earnest, a great deal of effort was expended in determining the relationship between joint probabilities of detection of a photon at time t and the **next** or **any** photon a time τ later. This was addressed in detail by many physicists (see [3] and references therein).

8.5 Decoherence

Decoherence, as already mentioned, is the transformation, over time, of a quantum-mechanical superposition state into a classical statistical mixture as a result of the quantum system interacting with the "environment". As a simple (perhaps even simple-minded) model demonstrating this effect, consider a cavity field initially prepared in the even cat state of Eq. (8.9). The initial state of the cavity-field plus environment system is then

$$|\Psi(0)\rangle \sim (|\alpha\rangle + |-\alpha\rangle)\,|\mathcal{E}\rangle\,, \tag{8.22}$$

where $|\mathcal{E}\rangle$ represents the quantum state of the environment. The environment, of course, consists of many systems and many states; but we are not interested here in such details. At some later time the system will have evolved [17] to

$$|\Psi(t)\rangle \sim |\beta(t)\rangle\,|\mathcal{E}_1\rangle + |-\beta(t)\rangle\,|\mathcal{E}_2\rangle\,, \tag{8.23}$$

where $\beta(t) = \alpha\, e^{-\gamma t/2}$, as we know from our earlier discussion. We see that the field state is entangled with states of the environment. The corresponding density operator is

$$\begin{aligned}\hat{\rho}(t) &\sim [|\beta(t)\rangle\,|\mathcal{E}_1\rangle + |-\beta(t)\rangle\,|\mathcal{E}_2\rangle]\,[\langle\beta(t)|\,\langle\mathcal{E}_1| + \langle-\beta(t)|\,\langle\mathcal{E}_2|] \\ &= |\beta(t)\rangle\,\langle\beta(t)| \otimes |\mathcal{E}_1\rangle\,\langle\mathcal{E}_1| + |-\beta(t)\rangle\,\langle-\beta(t)| \otimes |\mathcal{E}_2\rangle\,\langle\mathcal{E}_2| \\ &\quad + |\beta(t)\rangle\,\langle-\beta(t)| \otimes |\mathcal{E}_1\rangle\,\langle\mathcal{E}_2| + |-\beta(t)\rangle\,\langle\beta(t)| \otimes |\mathcal{E}_2\rangle\,\langle\mathcal{E}_1|\,. \end{aligned} \tag{8.24}$$

Tracing over the states of the environment, and assuming that $\langle\mathcal{E}_i \mid \mathcal{E}_j\rangle = \delta_{i,j}$, we obtain the reduced density operator of the field as

$$\begin{aligned}\hat{\rho}_{\mathrm{F}}(t) &\sim \mathrm{Tr}_{\mathrm{Env}}\,\hat{\rho}(t) = \sum_i \langle\mathcal{E}_i|\,\hat{\rho}(t)\,|\mathcal{E}_i\rangle \\ &= |\beta(t)\rangle\,\langle\beta(t)| + |-\beta(t)\rangle\,\langle-\beta(t)|\,, \end{aligned} \tag{8.25}$$

a statistical mixture of coherent states. Note that at *very* long times, we obtain just the vacuum state $|0\rangle\langle 0|$, again a pure state. The master equation, Eq. (8.7), brings about the nonunitary evolution that yields the results given by Eq. (8.25) as we shall now demonstrate.

To be clear that we are studying the case with $\hat{H} = 0$, we rewrite the relevant master equation as

$$\frac{d\hat{\rho}}{dt} = \frac{\gamma}{2}\{2\hat{a}\hat{\rho}\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho} - \hat{\rho}\,\hat{a}^\dagger\hat{a}\}. \tag{8.26}$$

Consider first an initial coherent state $\hat{\rho}(0) = |\alpha\rangle\langle\alpha|$. It is easy to show, by substitution, that

$$\hat{\rho}(t) = |\alpha\, e^{-\gamma t/2}\rangle\langle\alpha\, e^{-\gamma t/2}| \tag{8.27}$$

is the correct solution. Evidently, under dissipative forces at temperature $T = 0$ K, an initial coherent state remains a coherent state but with an exponentially

decreasing amplitude. The average photon number of the decaying state is given by

$$\bar{n}(t) = |\alpha|^2 \, e^{-\gamma t} \tag{8.28}$$

and thus the decay time of the field energy is $T_{\rm decay} = 1/\gamma$. Now let us consider as initial state the even cat state of Eq. (8.9). The initial field density operator will be

$$\hat{\rho}_{\rm F}(0) = |\psi_{\rm e}\rangle \langle \psi_{\rm e}| = \mathcal{N}_{\rm e}^2 \left[|\alpha\rangle \langle \alpha| + |-\alpha\rangle \langle -\alpha| + |\alpha\rangle \langle -\alpha| + |\alpha\rangle \langle -\alpha| \right]. \tag{8.29}$$

We recall the normally ordered characteristic function of Eq. (3.128b)

$$C_{\rm N}(\lambda, t) = {\rm Tr}[\hat{\rho}_{\rm F}(t) \, e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}}]. \tag{8.30}$$

For $\hat{\rho}(t)$ satisfying Eq.(8.26), the characteristic function satisfies the property [18, 19]

$$C_{\rm N}(\lambda, t) = C_{\rm N}(\lambda \, e^{-\kappa t/2}, 0). \tag{8.31}$$

Thus by a simple comparison, we may deduce the solution

$$\begin{aligned} \hat{\rho}_{\rm F}(t) = \mathcal{N}_{\rm e}^2 \{ & |\alpha \, e^{-\gamma t/2}\rangle \langle \alpha \, e^{-\gamma t/2}| + |-\alpha \, e^{-\gamma t/2}\rangle \langle -\alpha \, e^{-\gamma t/2}| \\ & + e^{-2|\alpha|^2(1-e^{-\gamma t})} [|\alpha \, e^{-\gamma t/2}\rangle \langle -\alpha \, e^{-\gamma t/2}| + |-\alpha \, e^{-\gamma t/2}\rangle \langle \alpha \, e^{-\gamma t/2}|] \}. \end{aligned} \tag{8.32}$$

The energy decays at the same rate as for a coherent state, but there is another quantity that is now undergoing a decay, namely the quantity $\exp[-2|\alpha|^2(1 - e^{-\gamma t})]$ that multiplies the "coherences", or the "off-diagonal" terms $|\pm\beta\rangle\langle\mp\beta|$. For short times we may expand to yield $\exp[-2|\alpha|^2(1 - e^{-\gamma t})] \approx \exp[-2\gamma \, t|\alpha|^2]$ from which we obtain the decoherence time $T_{\rm decoh} = 1/(2\gamma|\alpha|^2) = T_{\rm decay}/(2|\alpha|^2)$. This is the characteristic time over which the coherences tend to vanish. Note that the coherences decay faster than the energy as long as $|\alpha|^2$ is sufficiently high.

In Fig. 8.6 we follow the evolution of the Wigner function for an even cat state. At sufficiently long time, the interference fringes disappear while the Gaussians associated with each of the coherent state components of the original superposition remain distinct, though at very long times these two merge to form the Wigner function of the vacuum.

For more details on the phenomenon of decoherence, the reader should consult a recent article by Zurek [20] and the many references therein.

8.6 Generation of coherent states from decoherence: optical balance

In this chapter we have shown how dissipative effects tend to destroy the quantum coherence inherent in quantum superposition effects. It may therefore come as a bit of surprise to learn that, under some circumstances, quantum coherence can be *created* by dissipative effects. If we look at the master equation of Eq. (8.7), we see that both unitary and nonunitary evolutionary terms are present, the former

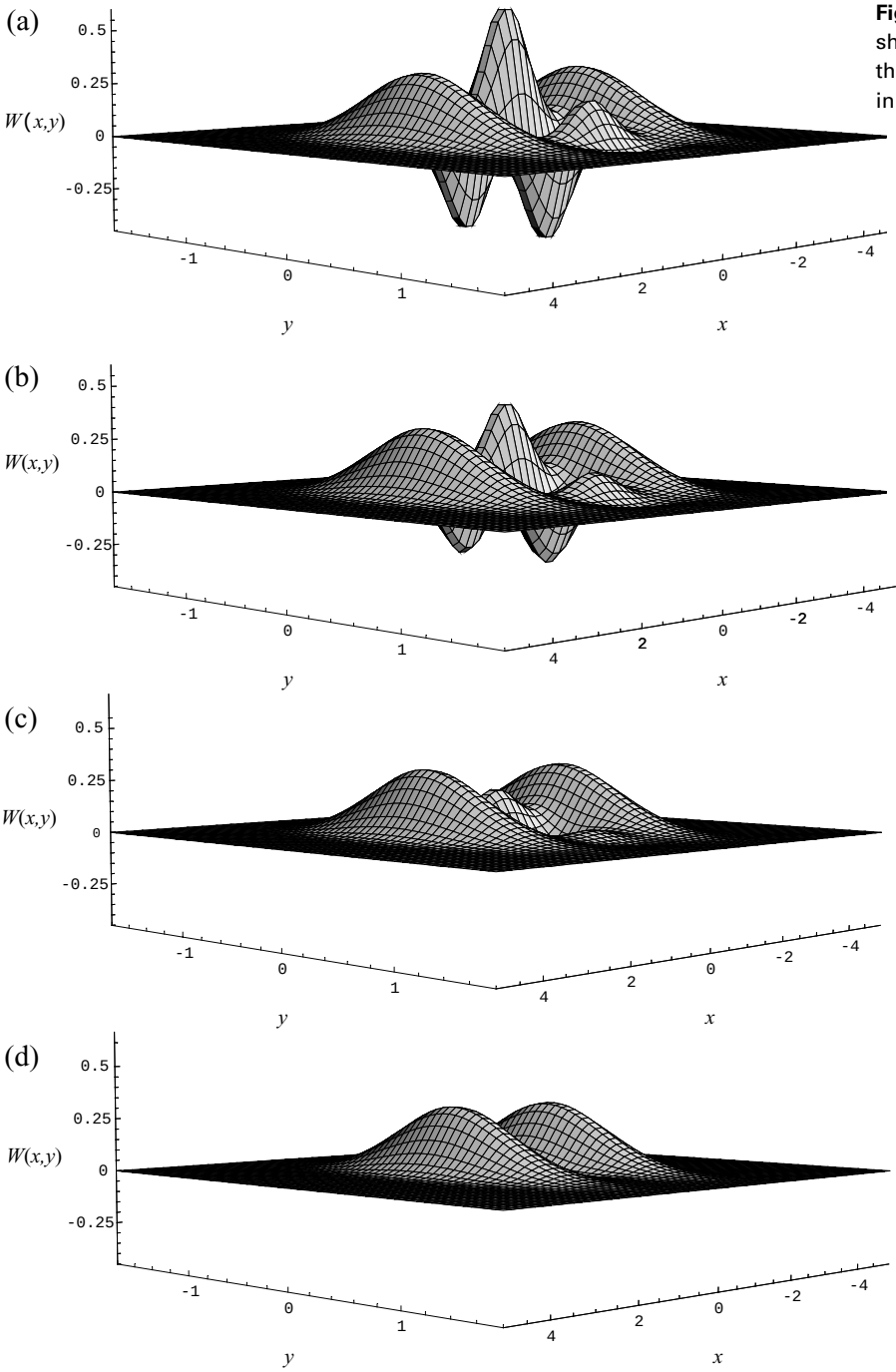


Fig. 8.6. A sequence showing the evolution of the Wigner function for an initial even cat state.

being the term with the commutation and the latter the dissipative terms, which are nonlinear in the annihilation and creation operators and associated with incoherent losses. These interactions are in competition with each other and thus it becomes possible that a system might evolve, as $t \to \infty$, into some state other than the vacuum. This condition just described is known as "optical balance" and we demonstrate it by assuming that a pumped single-mode field, described by the interaction Hamiltonian

$$\hat{H}_{\rm I} = \hbar G(\hat{a} + \hat{a}^\dagger), \tag{8.33}$$

where G represents the strength of the classical pump field, is propagating in a photon-absorbing medium where the photon absorption is described by the dissipative terms in Eq. (8.7). Thus our master equation becomes

$$\frac{d\hat{\rho}}{d\,t} = -i\,G[\hat{a} + \hat{a}^\dagger, \hat{\rho}] + \frac{\gamma}{2}\{2\hat{a}\hat{\rho}\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho} - \hat{\rho}\,\hat{a}^\dagger\hat{a}\}. \tag{8.34}$$

We introduce the new operator $\hat{b} = \hat{a} + i\,G/\gamma$ so that we may rewrite the previous equation suggestively as

$$\frac{d\hat{\rho}}{d\,t} = -\frac{\gamma}{2}\{2\hat{b}\hat{\rho}\,\hat{b}^\dagger - \hat{b}^\dagger\hat{b}\hat{\rho} - \hat{\rho}\,\hat{b}^\dagger\hat{b}\}. \tag{8.35}$$

In the steady state we have $d\hat{\rho}/dt = 0$ and thus it follows that we must have $\hat{b}\hat{\rho} = 0 = \hat{\rho}\,\hat{b}^\dagger$, which in turn means that

$$\left(\hat{a} + \frac{iG}{\gamma}\right)\hat{\rho} = 0 = \hat{\rho}\left(\hat{a}^\dagger - \frac{iG}{\gamma}\right). \tag{8.36}$$

These last equations are solved by

$$\hat{\rho} = |\alpha\rangle\langle\alpha| \quad \text{where} \quad \alpha = -i\,G/\gamma. \tag{8.37}$$

Thus competing coherent and incoherent processes lead to coherent states in the steady-state regime [21]. Note that if the initial state is a vacuum state, $\hat{\rho}(0) = |0\rangle\langle 0|$, then at times short enough to ignore the effects of dissipation, the evolution will be unitary and will also bring about a coherent state:

$$\begin{aligned}\hat{\rho}\,(t) &\approx e^{-iGt(\hat{a}+\hat{a}^\dagger)}\hat{\rho}(0)e^{iGt(\hat{a}+\hat{a}^\dagger)} \\ &= \hat{D}\,(-i\,G\,t)\,|0\rangle\,\langle 0|\,\hat{D}^\dagger\,(-i\,G\,t) \\ &= |-i\,G\,t\rangle\,\langle -i\,G\,t|\,,\end{aligned} \tag{8.38}$$

valid for times t very short.

8.7 Conclusions

In this chapter we have shown how quantum optics can address the interesting question of how one may understand the dynamics of single quantum systems. In current work on quantum computing, such questions are at the heart of how we can address, manipulate and read out the state of quantum systems.

Problems

1. Consider a cavity field initially in a number state $|n\rangle$. Describe what happens to this state during the various steps of the quantum jump algorithm.
2. Suppose the cavity field is prepared in the superposition state $(|0\rangle + |10\rangle)/\sqrt{2}$. What is the average photon number for this state? Suppose now that a quantum jump occurs. What is the average number of photons in the field now that a single photon has been emitted? Does the result make any sense?
3. Obtain a set of coupled first-order differential equations for the elements of the density matrix for the density operator satisfying the master equation of Eq. (8.26).
4. Solve these equations, numerically if necessary, if the initial state of the system is the number state $|5\rangle$. Obtain the Wigner function, making 3-D plots, for the field state at various stages in the evolution.
5. Perform a numerical integration of the master equation for the case of an initial even cat state with $|\alpha|^2 = 2$. As the state evolves, monitor (i.e. plot) $\mathrm{Tr}\hat{\rho}(t)$ vs. time to check that it maintains a value of unity (a good way to check on the numerics), and monitor $\mathrm{Tr}\hat{\rho}^2(t)$ as a measure of decoherence.
6. Show that the solution to the master equation of Eq. (8.26) can be given in iterative form as [22]

$$\rho_{mn}(t) = \exp\left(-\frac{\gamma\, t\,(m+n)}{2}\right)\sum_{l}\left(\frac{(m+l)!\,(n+l)!}{m!n!}\right)^{1/2} \times \frac{(1-\exp(-\gamma\, t))^{l}}{l!}\rho_{m+l,n+l}(0)\,, \tag{8.39}$$

   where the $\rho_{mn} = \langle m|\hat{\rho}|n\rangle$.
7. Use the iterative method to obtain the exact solutions of the master equation for initial coherent states and even cat states, in agreement with the above solutions.
8. Numerically integrate the master equation of Eq. (8.34) assuming the field initially in a vacuum state. Verify numerically in some way the existence of coherent states for both short-time and steady-state regimes. Also address the question of whether or not the evolving state is always a coherent state. Is it a pure state at all times?

References

[1] M. Ligare and R. Oliveri, *Am. J. Phys.*, **70** (2002), 58; V. Buzek, G. Drobny, M. G. Kim, M. Havukainen and P. L. Knight, *Phys. Rev. A*, **60** (1999), 582.

[2] See L. Allen and J. H. Eberly, *Optical Resonance and the Two-Level Atom* (New York: Wiley, 1975; Mineoloa: Dover, 1987).

[3] M. B. Plenio and P. L. Knight, *Rev. Mod. Phys.*, **70** (1998), 101.

[4] T. A. Brun, *Am. J. Phys.*, **70** (2002), 719.

[5] E. Schrödinger, *Br. J. Philos. Sci.*, **3** (1952), 109; 233.

[6] H. Dehmelt, *Bull. Am. Phys. Soc.*, **20** (1975), 60.

[7] G. Z. K. Horvath, P. L. Knight and R. C. Thompson, *Contemp. Phys.*, **38** (1997), 25.

[8] R. C. Thompson, private communication (1996).

[9] This is merely the a.c. Stark effect: the strong radiation shifts the atomic energy levels in an intensity-dependent fashion.
[10] H. A. Klein, G. P. Barwood, P. Gill and G. Huang, *Physica Scripta T*, **86** (2000), 33.
[11] R. J. Cook and H. J. Kimble, "Possibility of Direct Observation of Quantum Jumps", *Phys. Rev. Lett.*, **54** (1985), 1023.
[12] J. C. Bergquist, R. B. Hulet, W. M. Itano and D. J. Wineland, *Phys. Rev. Lett*., **57** (1986), 1699; W. Nagourney, J. Sandberg and H. G. Dehmelt, *Phys. Rev. Lett*., **56** (1986), 2797; T. Sauter, R. Blatt, W. Neuhauser and P. E. Toschek, *Opt. Comm*., **60** (1986), 287.
[13] See, for example, P. L. Knight, in *Quantum Fluctuations Les Houches Session LXIII*, S. Reynaud, E. Giacobino and J. Zinn-Justin (editors) (Amsterdam: Elsevier Science, 1997).
[14] J. Dalibard, Y. Castin and K. Mølmer, *Phys. Rev. Lett*., **68** (1992), 580; G. C. Hegerfeldt and T. S. Wilser, in *Proceedings of the II International Wigner Symposium*, H. D. Doebner, W. Scherer and F. Schroek (editors) (Singapore: World Scientific, 1991).
[15] See, for example, M. S. Kim and P. L. Knight, *Phys. Rev. A*, **36** (1987), 5265; see also references in [3].
[16] D. T. Pegg, R. Loudon and P. L. Knight, *Phys. Rev. A*, **33** (1986), 4085.
[17] See W. Zurek, *Physics Today*, **44**: 10 (1991), 36.
[18] J. Perina, *Quantum Statistics of Linear and Nonlinear Optical Phenomena* (Dordrecht: Reidel, 1984).
[19] M. Brune, S. Haroche, J. M. Raimond, L. Davidovich and N. Zagury, *Phys. Rev. A*, **45** (1992), 5193.
[20] W. Zurek, *Rev. Mod. Phys*., **75** (2003), 715.
[21] See G. S. Agarwal, *J. Opt. Soc. Am. B*, **5** (1988), 1940.
[22] J. Škvarček and M. Hillery, *Acta Phys. Slovaca*, **49** (1999), 756.

Bibliography

H. J. Carmichael, *An Open Systems Approach to Quantum Optics*, Lecture Notes in Physics (Berlin: Springer-Verlag, 1993).
P. Ghosh, *Testing Quantum Mechanics on New Ground* (Cambridge: Cambridge University Press, 1999).
W. H. Louisell, *Quantum Statistical Properties of Radiation* (New York: Wiley, 1973).
I. Percival, *Quantum State Diffusion* (Cambridge: Cambridge University Press, 1998).
S. Stenholm and M. Wilkens, "Jumps in quantum theory", *Contemp. Phys.*, **38** (1997), 257.

Chapter 9
Optical test of quantum mechanics

Over the past three decades or so, experiments of the type called Gedanken have become real. Recall the Schrödinger quote from Chapter 8: ". . . we never experiment with just one electron or atom or (small) molecule." This is no longer true. We can do experiments involving single atoms or molecules and even on single photons, and thus it becomes possible to demonstrate that the "ridiculous consequences" alluded to by Schrödinger are, in fact, quite real. We have already discussed some examples of single-photon experiments in Chapter 6, and in Chapter 10 we shall discuss experiments performed with single atoms and single trapped ions. In the present chapter, we shall elaborate further on experimental tests of fundamentals of quantum mechanics involving a small number of photons. By fundamental tests we mean tests of quantum mechanics against the predictions of local realistic theories (i.e. hidden variable theories). Specifically, we discuss optical experiments demonstrating violations of Bell's inequalities, violations originally discussed by Bell in the context of two spin-one-half particles [1]. Such violations, if observed experimentally, falsify local realistic hidden-variable theories. Locality refers to the notion, familiar in classical physics, that there cannot be a causal relationship between events with space-like separations. That is, the events cannot be connected by any signal moving at, or less than, the speed of light; i.e. the events are outside the light-cone. But in quantum mechanics, it appears that nonlocal effects, effects seemly violating the classical notion of locality in a certain restricted sense, are possible. For example, the fact that a measurement on one part of an entangled system seems to instantaneously project the other part of the system into a particular state, even though the parts may be widely separated, is a nonlocal quantum effect. "Realism" in the context of hidden variable theories means that a quantum system has objectively definite attributes (quantum numbers) for all observables at all times. For example, a particle with spin $+\,\frac{1}{2}$ along the z-direction is known to be in a superposition state with spin $\pm\,\frac{1}{2}$ along the x-direction. The standard Copenhagen interpretation of quantum mechanics holds that the particle's spin along the x-direction is objectively indefinite, i.e. has no particular value of spin as a matter of principle until a measurement reduces the state vector to one of the possible states in the superposition. In a hidden-variable theory, the spin along the x-axis is assumed to be definite,

though unknown, and the experiment merely reveals the already existing state of the particle. But there are two types of hidden-variable theory, local and nonlocal. Nonlocal hidden-variable theories of the type considered by Bohm [2] reproduce the predictions of standard quantum mechanics. But they exhibit nonlocality, the feature of quantum mechanics that seems to have bothered Einstein even more than the apparent lack of realism. But local hidden-variable theories do make predictions different from those of standard quantum theory, and Bell's theorem provides a way to make such tests. In this chapter, we shall discuss optical tests of local hidden-variable theories using polarization-entangled photons.

In what follows we shall first discuss modern sources of paired and entangled single-photon states obtained from down-conversion processes (as opposed to the technique of cascaded emissions from atomic transitions as discussed in Chapter 6). We then review some one- and two-photon interference experiments, introduce the notion of the "quantum eraser" in this context, and then discuss an experiment on induced quantum coherence. Next, an experiment on photon tunneling exhibiting superluminal effects is discussed. Finally we describe two experiments on tests of Bell's inequalities, one involving polarization-entangled photons and a second (Franson's experiment) based on the time–energy uncertainty relation.

9.1 Photon sources: spontaneous parametric down-conversion

We have already discussed, in Chapter 7, the generation of nonclassical light via parametrically driven nonlinear media characterized by a second-order nonlinear susceptibility $\chi^{(2)}$. We focus on the nondegenerate case, and assuming the pump field to be quantized, the interaction Hamiltonian takes the form

$$\hat{H}_I \sim \chi^{(2)} \hat{a}_{\rm p} \hat{a}_{\rm s}^{\dagger} \hat{a}_{\rm i}^{\dagger} + H.c., \tag{9.1}$$

where we have altered our notation from Chapter 7 so that now $\hat{a}_{\rm p}$ is the annihilation operator of the pump beam and $\hat{a}_{\rm s}^{\dagger}$ and $\hat{a}_{\rm i}^{\dagger}$ are the creation operators of the "signal" and "idler" beams respectively. The denotations "signal" and "idler" appear for historical reasons and have no special significance, the choice of beam labels being somewhat arbitrary. In the simplest case, with the signal and idler beams initially in vacuum states, a single pump beam photon, typically in the ultraviolet spectral range, is converted into two optical photons, one in the signal beam, the other in the idler:

$$|1\rangle_{\rm p} |0\rangle_{\rm s} |0\rangle_{\rm i} \Rightarrow \hat{a}_{\rm p} \hat{a}_{\rm s}^{\dagger} \hat{a}_{\rm i}^{\dagger} |1\rangle_{\rm p} |0\rangle_{\rm s} |0\rangle_{\rm i} = |0\rangle_{\rm p} |1\rangle_{\rm s} |1\rangle_{\rm i} . \tag{9.2}$$

As the signal and idler modes are initially in vacuum states, the process is "spontaneous". Note that the photons produced in the signal and idler modes are assumed to be generated simultaneously. That this is the case was demonstrated many years ago by Burnham and Weinberg [3] who used coincidence counting with detectors arranged to satisfy momentum and energy conservation and to have equal

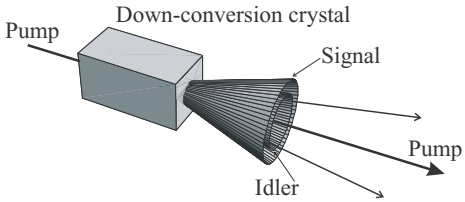


Fig. 9.1. Type I down-conversion. Photons from the pump beam are converted into signal and idler photons that emerge from the crystal along different directions. The signal and idler photons have identical polarization but orthogonal to that of the pump. The possible directions form a set of concentric cones. The light from the different cones is of different colors, typically orange near the center to a deep red at wider angles. The pump beam is in the ultraviolet.

time delays. The simultaneous production of signal and idler photons is key to the applications of such parametric sources to the fundamental test of quantum mechanics to be described. In order for the down-conversion process to go forward however, certain other conditions must be satisfied. Letting $\omega_{\rm p}$, $\omega_{\rm s}$, and $\omega_{\rm i}$ represent the frequencies of the pump, signal, and idler respectively, energy conservation requires that

$$\hbar\omega_{\rm p} = \hbar\omega_{\rm s} + \hbar\omega_{\rm i}. \tag{9.3}$$

Further, if $\mathbf{k}_{\rm p}$, $\mathbf{k}_{\rm s}$, and $\mathbf{k}_{\rm i}$ represent the respective wave numbers, then we must have, inside the crystal,

$$\hbar\mathbf{k}_{\rm p} \approx \hbar\mathbf{k}_{\rm s} + \hbar\mathbf{k}_{\rm i}, \tag{9.4}$$

where the $\approx$ sign appears as the result of an uncertainty given by the reciprocal of the length of the nonlinear medium [4]. Equations (9.3) and (9.4) are known as the "phase matching" conditions and they can be attained in certain types of nonlinear media, such as noncentrosymmetric crystals [5]. Only noncentrosymmetric crystals have a nonvanishing $\chi^{(2)}$. The most commonly used crystals are KDP (KD_2PO_4) and BBO (β-BaB_2O_4). The connection to nonlinear optics is given in Appendix D.

There are, in fact, two types of SPDC process. In type I, the signal and idler photons have the same polarization but these are orthogonal to that of the pump. The interaction Hamiltonian for this process is given by

$$\hat{H}_{\rm I} = \hbar\eta\,\hat{a}_{\rm s}^{\dagger}\hat{a}_{\rm i}^{\dagger} + H.c., \tag{9.5}$$

where the parametric approximation has been made and where $\eta \propto \chi^{(2)}\mathcal{E}_{\rm p}$,where $\mathcal{E}_{\rm p}$ is the amplitude of the classical coherent field. The phase-matching condition of Eq. (9.4) imposes a constraint such that the signal and idler photons (conjugate photons) must emerge from the crystal on opposite sides of concentric cones centered on the direction of the pump beam as shown in Fig. 9.1. Evidently there are an infinite number of ways of selecting the signal and idler beams. Examples

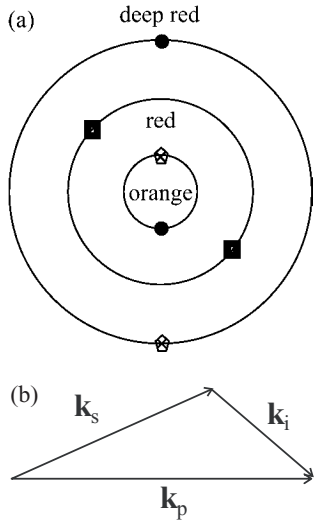


Fig. 9.2. (a) A cross-section of the cones of light emerging from a type I down-converter. Like symbols represent conjugate photons satisfying the phase-matching condition. Note that those on the middle circle are degenerate in frequency. (b) A graphical view of the phase-matching condition.

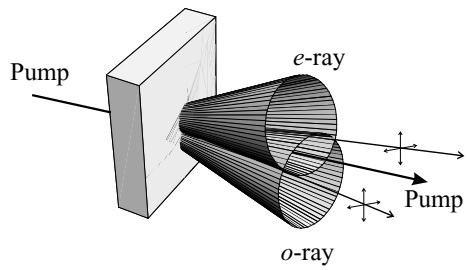


Fig. 9.3. Type II down-conversion. The signal and idler photons have orthogonal polarizations. Birefringence effects cause the photons to be emitted along two intersecting cones, one of the ordinary (*o*) ray and the extraordinary (*e*) ray.

are shown in Fig. 9.2. The Hamiltonian of Eq. (9.5) in practice represents a particular "post-selection" of the momenta of the output beams.

In type II down-conversion, the signal and idler photons have orthogonal polarizations. Because of birefringence effects, the generated photons are emitted along two cones, one for the ordinary (*o*) wave and another for the extraordinary (*e*) wave, as indicated in Fig. 9.3. The intersection of the cones provides a means for generating polarization-entangled states. We use the notation $|V\rangle$ and $|H\rangle$ to represent vertically and horizontally polarized *single-photon* states. For photons that emerge along the intersections of the cones, with photons from other parts of the cones being excluded by screens with pinholes in front of the intersection points, there will be an ambiguity as to whether the signal or idler photons will be vertically or horizontally polarized, as indicated in Fig. 9.4. The Hamiltonian describing this is given by

$$\hat{H}_{\rm I} = \hbar\eta\left(\hat{a}^{\dagger}_{\rm Vs}\hat{a}^{\dagger}_{\rm Hi} + \hat{a}^{\dagger}_{\rm Hs}\hat{a}^{\dagger}_{\rm Vi}\right) + H.c., \tag{9.6}$$

where the operators $\hat{a}^{\dagger}_{\rm Vs}$, $\hat{a}^{\dagger}_{\rm Hs}$, $\hat{a}^{\dagger}_{\rm Vi}$, and $\hat{a}^{\dagger}_{\rm Hi}$ are the creation operators for photons with vertical and horizontal polarization for the signal and idler beams,

respectively. Again, this Hamiltonian represents the post-selection obtained by placing a screen in front of the source with pinholes located just in the regions of the overlapping beams.

Let us take the initial state of the signal and idler modes to be represented by $|\Psi_0\rangle = |\{0\}\rangle$, which is the collective vacuum state for either type I or type II down-conversion. In either case, the state vector evolves according to

$$|\Psi(t)\rangle = \exp(-it\hat{H}_{\rm I}/\hbar)\,|\Psi_0\rangle\,, \tag{9.7}$$

which we expand, since $\hat{H}_{\rm I}$ has no explicit time dependence, as

$$|\Psi(t)\rangle \approx [1 - it\hat{H}_{\rm I}/\hbar + \tfrac{1}{2}(-it\hat{H}_{\rm I}/\hbar)^2]|\Psi_0\rangle \tag{9.8}$$

to second order in time. If we consider the type I SPDC then $|\Psi_0\rangle = |0\rangle_{\rm s}|0\rangle_{\rm i}$ and we have

$$|\Psi(t)\rangle = (1 - \mu^2/2)|0\rangle_{\rm s}|0\rangle_{\rm i} - i\mu|1\rangle_{\rm s}|1\rangle_{\rm i} \tag{9.9}$$

where $\mu = \eta t$. This state vector is normalized to first order in μ and we have dropped the term of order μ^2 containing the state $|2\rangle_{\rm s}|2\rangle_{\rm i}$. In the case of type II SPDC with the initial state $|\Psi_0\rangle = |0\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|0\rangle_{\rm Hi}$, we have

$$\begin{aligned}|\Psi(t)\rangle &= (1 - \mu^2/2)|0\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|0\rangle_{\rm Hi} \\ &\quad -i\mu\frac{1}{\sqrt{2}}(|1\rangle_{\rm Vs}|0\rangle_{\rm Hs}|0\rangle_{\rm Vi}|1\rangle_{\rm Hi} + |0\rangle_{\rm Vs}|1\rangle_{\rm Hs}|1\rangle_{\rm Vi}|0\rangle_{\rm Hi}).\end{aligned} \tag{9.10}$$

We now define the vertically and horizontally polarized vacuum and single photon states as $|0\rangle := |0\rangle_{\rm V}|0\rangle_{\rm H}$, $|V\rangle := |1\rangle_{\rm V}|0\rangle_{\rm H}$, and $|H\rangle := |0\rangle_{\rm V}|1\rangle_{\rm H}$, so that we may write

$$\begin{aligned}|\Psi(t)\rangle &= (1 - \mu^2/2)|0\rangle_{\rm s}|0\rangle_{\rm i} \\ &\quad -i\mu(|V\rangle_{\rm s}|H\rangle_{\rm i} + |H\rangle_{\rm s}|V\rangle_{\rm i}).\end{aligned} \tag{9.11}$$

The state in the second term, which when normalized reads

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|V\rangle_{\rm s}|H\rangle_{\rm i} + |H\rangle_{\rm s}|V\rangle_{\rm i}), \tag{9.12}$$

is one member out of a set of four states known as Bell states. The full set of Bell states is

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 \pm |V\rangle_1|H\rangle_2), \tag{9.13}$$

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|H\rangle_2 \pm |V\rangle_1|V\rangle_2). \tag{9.14}$$

We shall discuss these states, and their implications, in Section 9.6.

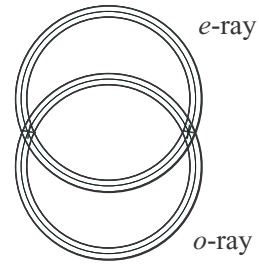


Fig. 9.4. The intersections of the cones for the *o*-ray and the *e*-ray are the sources for polarization entangled light. From these points it is not possible to tell from which beam a photon is obtained. The Hamiltonian in Eq. (9.6) describes the light taken from both intersections.

9.2 The Hong–Ou–Mandel interferometer

In Chapter 6 we discussed what happens when twin single-photon states are simultaneously incident at each input port of a 50:50 beam splitter: the photons

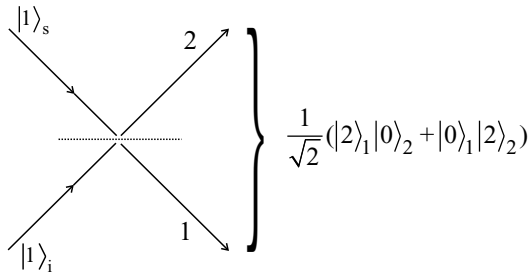


Fig. 9.5. Labeling of the output modes of a 50:50 beam splitter with inputs from the signal and idler beams of a type I down-converter. If single-photon states simultaneously fall onto the beam splitter, the output does not contain the term $|1\rangle_1|1\rangle_2$, for reasons discussed in Chapter 6.

emerge together in one output beam or the other; no single photons ever emerge in both beams. Recall that for the input state $|1\rangle_s|1\rangle_i$ one has after the beam splitter the state

$$|\psi_{BS}\rangle = \frac{1}{\sqrt{2}}(|2\rangle_1|0\rangle_2 + |0\rangle_1|2\rangle_2), \tag{9.15}$$

where we have labeled the output modes 1 and 2 in accordance with Fig. 9.5. Detectors placed at the outputs should never register simultaneous counts. In fact, one could take the lack of simultaneous counts as an indication that the photons are incident on the beam splitter simultaneously. The first demonstration of this effect was by Hong, Ou and Mandel (HOM) in a now-classic experiment performed in 1987 [6]. In fact, the experiment was designed to measure the time separation between the two photons striking a beam splitter. A sketch of their experiment is given in Fig. 9.6. The nonlinear crystal is pumped to produce, assuming type I down-conversion, twin single-photon states whose beams are then directed to the input ports of a 50:50 beam splitter. Photon detectors are placed at the outputs of the beam splitter and the count signals are fed into a correlator. Changing the position of the beam splitter causes a slight time delay between the times the photons fall onto the beam splitter. With a slight nonzero time delay, the photons independently reflect or transmit through the beam splitter causing both detectors sometimes to fire within a short time of each other. It can be shown [6] that the rate of coincident detections, R_{coin}, has the form

$$R_{coin} \sim \left[1 - e^{-(\Delta\omega)^2(\tau_s - \tau_i)^2}\right], \tag{9.16}$$

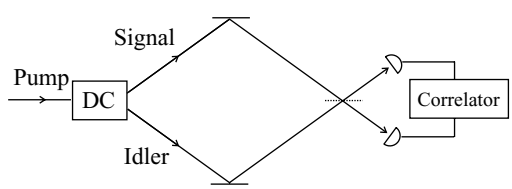


Fig. 9.6. The Hong–Ou–Mandel experiment. When the path lengths are equal, no coincident counts are detected.

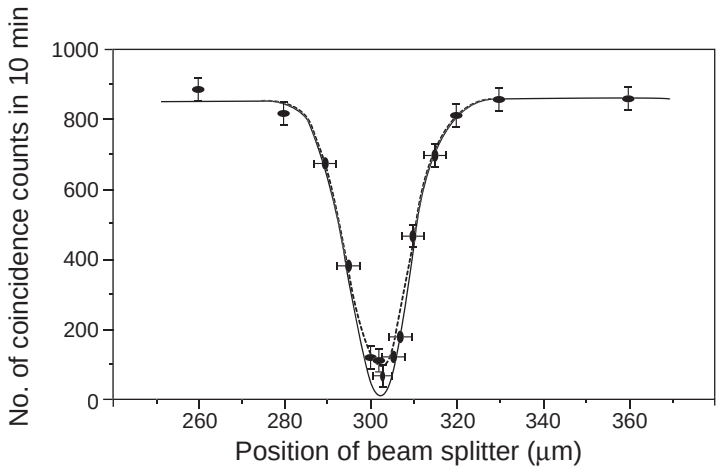


Fig. 9.7. Redrawn from the Hong–Ou–Mandel paper, the number of coincident counts in a 10 minute time interval as a function of the position of the beam splitter. The count rate does not go to zero exactly, owing to the fact that the beams do not perfectly overlap at the beam splitter. (Reprinted with permission.)

where $\Delta\omega$ is the bandwidth of the light and $c\tau_s$ and $c\tau_i$ are the distances that the signal and idler photons respectively travel from the down-converter to the beam splitter. The band width $\Delta\omega$ incorporates the reality that the signal and idler beams are not monochromatic, and its appearance in Eq. (9.16) results from the assumption that the spectral distribution is Gaussian. Obviously, for $\tau_s - \tau_i = 0$ we have $R_{coin} = 0$. The rate of coincidence counts rises to a maximum for $|\tau_s - \tau_i| \gg \tau_{corr}$, where $\tau_{corr} = 1/\Delta\omega$ is the correlation time of the photons. The correlation time is of the order of a few nanoseconds, which is hard to measure with conventional techniques as detectors commonly used do not often have short enough resolving times. But this correlation time can be measured with the HOM interferometer. The experimental results are plotted in Fig. 9.7, reprinted from Reference [6]. Plotted here is the number of counts over an interval of 10 min against the position of the beam splitter (essentially the time separation) with the solid line representing the theoretical prediction. The experimental data do not go exactly to the predicted minimum because it is not possible for the beams precisely to overlap at the beam splitter. From the distribution of the counts, the correlation time of the two photons can be determined to be $\sim$100 fs.

9.3 The quantum eraser

In the HOM experiment just described, the fact that the photons are indistinguishable is the key to understanding the results. Because type I SPDC is used as the source, both photons have the same polarization. They may have slightly different energies but the photon detectors are not really sensitive to the energy difference. We have not needed to specify the polarization direction of the photons; it was only important that they be the same. But now suppose, for the sake of definiteness, we take them to be horizontally polarized (see Fig. 9.8), denoting

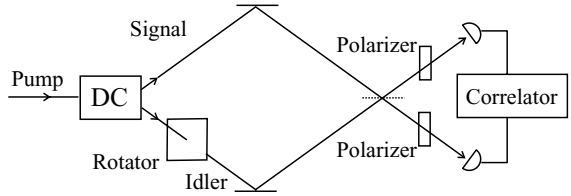


Fig. 9.8. A quantum erasure experiment. The rotator in the idler beam rotated the polarization of idler photons thus marking them and yielding path information. This destroys the quantum interference exhibited in the Hong–Ou–Mandel experiment. Inserting polarizers in the beam after the interferometer restores quantum interferences.

the twin single-photon states as $|H\rangle_s|H\rangle_i$. With such an input state to the beam splitter (assumed to not affect the polarization of the photons), the output state can be written as $i(|2H\rangle_1|0\rangle_2 + |0\rangle_1|2H\rangle_2)/\sqrt{2}$, where $|2H\rangle$ is a state with two horizontally polarized photons.

Suppose we now place a rotator in, say, the idler beam, such that its polarization is rotated by the angle θ to the horizontal, as indicated in Fig. 9.9. This transforms the idler polarization state, to

$$|\theta\rangle_i = |H\rangle_i \cos\theta + |V\rangle_i \sin\theta. \tag{9.17}$$

The polarization state vector for a photon with polarization orthogonal to this is given by

$$|\theta^\perp\rangle_i = -|H\rangle_i \sin\theta + |V\rangle_i \cos\theta. \tag{9.18}$$

The input state to the beam splitter is now

$$|H\rangle_s|\theta\rangle_i = \cos\theta|H\rangle_s|H\rangle_i + \sin\theta|H\rangle_s|V\rangle_i. \tag{9.19}$$

The output state will be

$$\begin{aligned}|\psi_{out}(\theta)\rangle = &\frac{i}{\sqrt{2}}\cos\theta(|2H\rangle_1|0\rangle_2 + |0\rangle_1|2H\rangle_2)\\ &+\frac{1}{\sqrt{2}}\sin\theta(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2),\end{aligned} \tag{9.20}$$

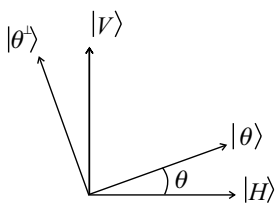


Fig. 9.9. The relationship between the $|\theta\rangle$, $|\theta^\perp\rangle$ and $|H\rangle$, $|V\rangle$ polarization vectors.

where the terms containing the single-photon states no longer cancel each other as the photons have different polarizations.* If the polarization of the idler beam is rotated all the way to the vertical, i.e. to $\theta = \pi/2$, then the output state will be

$$|\psi_{out}(\pi/2)\rangle = |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2). \tag{9.21}$$

* Remember that to include the polarization state of a photon we really must expand the Hilbert space to write $|H\rangle_1 = |1\rangle_{H_1}|0\rangle_{V_1}$, etc., so that the last term will read $(|1\rangle_{H_1}|0\rangle_{V_1}|0\rangle_{H_2}|1\rangle_{V_2} - |0\rangle_{H_1}|1\rangle_{V_1}|1\rangle_{H_2}|0\rangle_{V_2})/\sqrt{2}$, which perhaps makes it more apparent why the single-photon states don't cancel as before.

There will now be *only* coincident counts in the detectors and neither detector, assuming 100% efficiency, will fire alone. The "marking" of one of the beams by the rotation of its polarization by 90° has the effect of removing photon indistinguishability and thus it is *possible* to determine the path taken by each of the photons before the beam splitter. Interestingly, in the scenario discussed, the photon polarization is never measured. The experimenter need not know the polarization of the photons; only the counts are measured. Evidently, the interference can be destroyed by the mere potential of obtaining which-path information, even if that information is never known to the experimenter. Now in the case of one-photon input to a beam splitter, one can observe the particle-like nature of the photon by placing detectors at the beam splitter outputs, or one can observe the wave-like nature, with interference effects, by using a Mach–Zehnder interferometer, as we have discussed in Chapter 6, providing an example of Bohr complementarity. Usually, it is the loss of coherence engendered by the availability of which-path information that is ascribed to the disappearance of interference. This "decoherence" supposedly renders the state vector into a statistical mixture where only probabilities, and not probability amplitudes, appear. But in the case of the two-photon interferometry experiment discussed here, there is no time, up to the point when the photons are detected, where we do not have a pure state.

Suppose now that a linear polarizer is placed at an angle θ_1 to the horizontal in front of the detector in output beam 1. The placing of a polarizer followed by photon detection constitutes a von Neumann projection (see Appendix B) onto the state vector

$$|\theta_1\rangle_1 = |H\rangle_1 \cos\theta_1 + |V\rangle_1 \sin\theta_1. \tag{9.22}$$

That is, only photons with polarization state $|\theta_1\rangle_1$ will be registered by the detector. The state of Eq. (9.21) is reduced to the pure state

$$\begin{aligned} |\psi, \theta_1\rangle &= \frac{|\theta_1\rangle\langle\theta_1 \mid \psi_{\text{out}}(\pi/2)\rangle}{\langle\psi_{\text{out}}(\pi/2) \mid \theta_1\rangle\langle\theta_1 \mid \psi_{\text{out}}(\pi/2)\rangle^{1/2}} \\ &= |\theta_1\rangle(|V\rangle_2 \cos\theta_1 - |H\rangle_2 \sin\theta_1) \end{aligned} \tag{9.23}$$

where the photon of mode 2 has a polarization orthogonal to that of the detected photon (see Eq. (9.18)). Suppose we similarly place a polarizer at the angle θ_2 to the horizontal in front of the detector for output beam 2. The probability that there will be coincident detections of a photon polarized at angle θ_1 and another at θ_2 is then

$$\begin{aligned} P_{\text{coin}} &= |\langle\theta_1|\langle\theta_2 \mid \psi_{\text{out}}(\pi/2)\rangle|^2 \\ &= \frac{1}{2}\sin^2(\theta_2 - \theta_1). \end{aligned} \tag{9.24}$$

From this result we see that the dip in the coincident count rate can be revived, depending on the relative angle of the polarizers. The effect of the polarizers placed just before the detectors is to erase the information encoded onto one

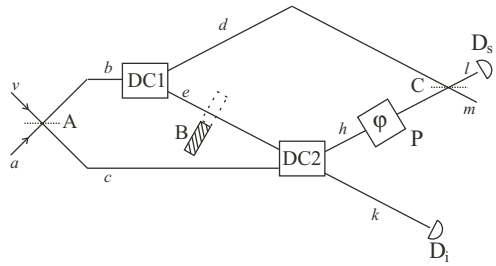


Fig. 9.10. Schematic of the Zou, Wang and Mandel experiment. The mode labeled *v* contains a vacuum state. B is a beam stop that can be moved in and out of the beam *e*.

of the beams by the rotator. The rotator creates which-path (*Welcher-Weg*, as it is often called) information and the polarizers erase it. Notice that the erasure takes place *outside* the HOM interferometer, that is, after the beam splitter. The optical quantum erasure effect has been demonstrated experimentally by Kwiat, Steinberg and Chiao [7].

9.4 Induced coherence

We now describe and explain another experiment, one that has been described, quite properly, as "mind boggling"[8], where, again, which-path information destroys interference but now where the mechanism of obtaining such information is not even in the interfering pathways. The experiment was performed by Zou, Wang and Mandel [9]. A sketch of the experiment is given in Fig. 9.10. Notice that without beam stop B blocking the idler of the down-converter DC1, the idler modes of the two down-converters are aligned.

Qualitatively, here's what happens. A pump photon from an argon-ion laser enters beam splitter A and goes to either DC1 or DC2, both assumed to be of type I. The down-converters, at a very low conversion rate of about $\gamma = 10^{-6}$, produce pairs of photons. With the path lengths adjusted appropriately, and with beam stop B removed, the detector D_i monitoring the aligned idler beams cannot determine in which crystal down-conversion has occurred (the crystals are transparent). As a result, interference can be seen in the detector D_s monitoring one of the outputs of beam splitter C upon which fall the two signal beams. The interference can be observed by adjusting the path length, or equivalently, the phase φ as indicated. The interference occurs whether the aligned idler beams are monitored or not. However, if the idler beam of DC1 is blocked, a detection of a photon by D_i would indicate that that photon had to have been generated in DC2, which, in turn, means that the initial input photon had taken the lower path to DC2. Thus which-path information is available and the interference at the other detector disappears. The distinctive feature of this experiment is that idler beams are not even in the paths leading to interference.

We now try to understand this experiment quantitatively. Using the labeling of the modes as given in Fig. 9.10, the initial state of the system is $|1\rangle_a|0\rangle_v$ with all

other modes in vacuum states. The first beam splitter causes the transformation

$$|1\rangle_a\,|0\rangle_v \to \frac{1}{\sqrt{2}}\,(|1\rangle_b\,|0\rangle_c + i\,|0\rangle_b\,|1\rangle_c)\,. \tag{9.25}$$

Now, at DC1 we have $|1\rangle_b|0\rangle_d|0\rangle_e \to \gamma|0\rangle_b|1\rangle_d|1\rangle_e$ and similarly at DC2 we have $|1\rangle_c|0\rangle_h|0\rangle_k \to \gamma|0\rangle_c|1\rangle_h|1\rangle_k$ so that just after the down-converters we have

$$|\psi_{\mathrm{DC1,2}}\rangle \sim \frac{\gamma}{\sqrt{2}}|0\rangle_b|0\rangle_c[|1\rangle_d|1\rangle_e|0\rangle_h|0\rangle_k + i|0\rangle_d|0\rangle_e|1\rangle_h|1\rangle_k]. \tag{9.26}$$

The phase shift in the h beam gives us

$$|\psi_{\mathrm{PS}}\rangle \sim \frac{\gamma}{\sqrt{2}}\,|0\rangle_b\,|0\rangle_c\,[|1\rangle_d\,|1\rangle_e\,|0\rangle_h\,|0\rangle_k + ie^{i\varphi}\,|0\rangle_d\,|0\rangle_e\,|1\rangle_h\,|1\rangle_k]. \tag{9.27}$$

The beam splitter at C performs the transformations

$$\begin{aligned} |1\rangle_d\,|0\rangle_h &\to \frac{1}{\sqrt{2}}\,(|1\rangle_m\,|0\rangle_l + i\,|0\rangle_m\,|1\rangle_l)\,, \\ |0\rangle_d\,|1\rangle_h &\to \frac{1}{\sqrt{2}}\,(|0\rangle_m\,|1\rangle_l + i\,|1\rangle_m\,|0\rangle_l)\,, \end{aligned} \tag{9.28}$$

so finally we have

$$\begin{aligned} |\psi_{\mathrm{out}}\rangle \sim \frac{\gamma}{2}\,&|0\rangle_b\,|0\rangle_c\,[|1\rangle_e\,|0\rangle_k\,(|1\rangle_m\,|0\rangle_l + i\,|0\rangle_m\,|1\rangle_l) \\ &+ie^{i\varphi}\,|0\rangle_e\,|1\rangle_k\,(|0\rangle_m\,|1\rangle_l + i\,|1\rangle_m\,|0\rangle_l)]. \end{aligned} \tag{9.29}$$

Now, if the idler beams are perfectly aligned, modes e and k are identical, and that mode contains one photon. That is, we can write $(|0\rangle_e|1\rangle_k, |1\rangle_e|0\rangle_k) \to |1\rangle_k$ reflecting the fact that the detector can't tell which down-converter produced the photon. Thus, Eq. (9.29) can be written as

$$|\psi_{\mathrm{out}}\rangle \sim \frac{\gamma}{2}|0\rangle_b|0\rangle_c|1\rangle_k[(1-e^{i\varphi})|1\rangle_m|0\rangle_l + i(1+e^{i\varphi})|0\rangle_m|1\rangle_l]. \tag{9.30}$$

The factorization of the state $|1\rangle_k$ containing the photon from the aligned idler beams is crucial to predicting interference in the single photon counts. The probability amplitude for $\mathrm{D_s}$ and $\mathrm{D_i}$ detecting single photons jointly in modes l and k, respectively, is then $i\gamma(1+e^{i\varphi})/2$ and the corresponding probability is

$$P\,(1_l, 1_k) \sim \frac{\gamma^2}{2}\,(1+\cos\varphi)\,, \tag{9.31}$$

which clearly contains information on the phase and thus exhibits interference effects. Actually, because the state $|1\rangle_k$ is factored out in Eq. (9.30), again the result of the alignment of the e and k beams, we really need not do any detection at all on the idler beams; all we need do is to perform detection with $\mathrm{D_s}$ where the detection probability for a single photon is

$$P\,(1_l) = P\,(1_l, 1_k) \sim \frac{\gamma^2}{2}\,(1+\cos\varphi)\,. \tag{9.32}$$

But now suppose we insert the beam stop B, which, in fact, could be just a mirror reflecting the photon in the e mode away from any detector. In any case, the idler beams e and k are no longer aligned and thus it is no longer possible to

factor out the states of these modes in Eq. (9.29). The probability of obtaining coincident counts by D_s and D_i is now

$$P(1_l, 1_k) \sim \left| \frac{ie^{i\varphi}\gamma}{2} \right|^2 = \frac{\gamma^2}{4}, \tag{9.33}$$

which contains no interference term. The detection of a photon by D_i necessarily means that the photon was generated by DC2 and thus yields path information. Other detection combinations are possible, but all the probabilities are the same:

$$P(1_l, 1_k) = P(0_l, 0_k) = P(1_l, 0_k) = P(0_l, 1_k) \sim \frac{\gamma^2}{4}. \tag{9.34}$$

Note that nondetection of a photon by D_i also yields which-path information.

Suppose we don't do *any* measurement on the k beam: we just remove detector D_i. We shall now obtain no which-path information. Will the quantum interference revive in the single photon counts by D_s? Detection of a single photon state in the l mode can be represented by the projection operator $|1\rangle_{ll}\langle 1|$. From Eq. (9.29), the probability that a single photon is detected in mode l is

$$P(1_l) = \langle \psi_{\text{out}} \mid 1\rangle_{l\,l}\langle 1 \mid \psi_{\text{out}}\rangle \sim \frac{\gamma^2}{2}, \tag{9.35}$$

as can easily be checked. But this contains no phase information and hence there is *still no interference* in the single-photon counting at detector D_s. Again, the conventional wisdom is that lack of which-path information is a condition for quantum interference. It must not be sufficient. Here we have no which-path information but we have no interference either. This is another example, perhaps more dramatic than the previous one, demonstrating that it is not necessarily the *possession* of which-path information that destroys quantum interference but rather it is the mere *potential* of possessing such information.

9.5 Superluminal tunneling of photons

An interesting application of the HOM interferometer is to the measurement of tunneling times of photons through a barrier. Tunneling, one of the most striking predictions of quantum mechanics, is, of course, a strictly quantum-mechanical phenomenon having no classical analog. A perennial question has been: what is the time, the tunneling time, for a particle to tunnel through a barrier? The answers to this question have been various over the years and we do not wish to review the literature on that topic here (but see the reviews in Reference [10]). But we shall point out that the HOM interferometer can be used to measure the time for tunneling photons in a manner that would simply not be possible with ordinary electronic laboratory instruments, these devices being too slow. The key, again, is that the signal and idler photons are produced simultaneously.

So consider once again the HOM interferometer but with a tunneling barrier placed in the signal arm as pictured in Fig. 9.11. In the experiment of Steinberg,

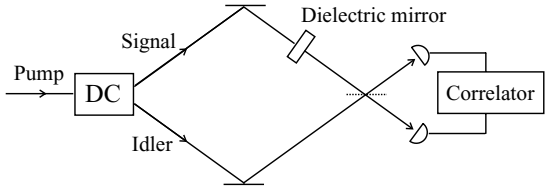


Fig. 9.11. Experiment on superluminal tunneling of photons. With the barrier placed in the signal beam, the path length of that beam must be increased in order to bring back the cancellation of coincident photon counts after the beam splitter.

Kwiat and Chiao [11], the barrier was a one-dimensional photonic band-gap material consisting of a multilayered dielectric mirror with a thickness of $d = 1.1\mu$m. For a particle traveling at the speed of light c, the traversal time across the barrier would be $d/c = 3.6$ fs. Taking into account the index of refraction, of course, leads to longer traversal time.

Suppose that the HOM interferometer, without the barrier, is balanced such that there are no coincident counts by the detectors: both signal and idler photons reach the beam splitter simultaneously. Now we insert the barrier. Naïvely, we might expect that the photon from the idler mode will arrive first at the beam splitter. But this is another example where naïve intuition fails. If the idler *were* to arrive at the beam splitter first, we would expect that we could regain balance (no coincident counts) by lengthening its path. However, experiments show that it is the path of the *signal* that needs to be lengthened indicating that the photon passing through that barrier arrives at the beam splitter first. In fact, tunneling photons seem to be traveling at a speed greater than the speed of light, at about $1.7c$!

Needless to say, the interpretation of this result has led to some controversy. The conventional explanation of this effect is as follows. The emission time of a photon (or the photon pair) is not precisely known. Thus its location is not precisely known either though one can picture its position as being described by a Gaussian probability distribution – a Gaussian wave packet. Because of the correlations between the photons, both will have identical Gaussian wave packets. But the packet striking the barrier becomes split, the bulk of it being reflected with a very small part transmitted. Both of these packets describe the same photon. Most of the time the photon is reflected. The part representing the tunneling of the photon through the barrier is thought to be, as a result of pulse reshaping, composed mostly of the front part of the original wave packet. Thus that section of the wave packet, because of the uncertainty in the time of the production of the photons, reaches the beam splitter before the centroid of the packet in the other beam. At no time is any photon moving faster than the speed of light, and one certainly cannot use this phenomenon for faster-than-light signaling. However, this explanation has not satisfied everyone and controversy remains [12]. In Section 9.7 we examine another consequence of the uncertainty in the time of generation of a pair of photons.

9.6 Optical test of local realistic theories and Bell's theorem

We have already shown how the Bell state $|\Psi^-\rangle$ can be produced from type I down-conversion by placing a $\pi/2$ rotator in one of the beams. The other Bell states can be generated as well via down-conversion in some form. For example, one can stack two type I down-converters whose optical axes are 90° to each other and pump with photons polarized at 45° to the optical axes of the crystals [13]. This source turns out to be a rather bright source of entangled light. Such an arrangement can produce the Bell states $|\Phi^\pm\rangle$. Type II down-conversion, as described by the Hamiltonian of Eq. (9.6), can produce any of the four Bell states [14].

Let us now suppose, to make a definite choice, that we are able to routinely produce the Bell state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 - |V\rangle_1|H\rangle_2) \tag{9.36}$$

through a down-conversion process. Further, we define, in each of the modes, polarization states along and orthogonal to the directions θ (for mode 1) and ϕ (for mode 2):

$$\begin{aligned} |\theta\rangle_1 &= \cos\theta|H\rangle_1 + \sin\theta|V\rangle_1, \\ |\theta^\perp\rangle_1 &= -\sin\theta|H\rangle_1 + \cos\theta|V\rangle_1, \\ |\phi\rangle_2 &= \cos\phi|H\rangle_2 + \sin\phi|V\rangle_2, \\ |\phi^\perp\rangle_2 &= -\sin\phi|H\rangle_2 + \cos\phi|V\rangle_2. \end{aligned} \tag{9.37}$$

In terms of these states we can write

$$\begin{aligned} |\Psi^-\rangle = \frac{1}{\sqrt{2}}\{&(\cos\theta\sin\phi - \sin\theta\cos\phi)|\theta\rangle_1|\phi\rangle_2 \\ &+(\cos\theta\cos\phi + \sin\theta\sin\phi)|\theta\rangle_1|\phi^\perp\rangle_2 \\ &-(\sin\theta\sin\phi + \cos\theta\cos\phi)|\theta^\perp\rangle_1|\phi\rangle_2 \\ &-(\sin\theta\cos\phi - \cos\theta\sin\phi)|\theta^\perp\rangle_1|\phi^\perp\rangle_2\}. \end{aligned} \tag{9.38}$$

For the choice $\phi = \theta$ this result reduces to

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle_1|\theta^\perp\rangle_2 - |\theta^\perp\rangle_1|\theta\rangle_2), \tag{9.39}$$

demonstrating the rotational invariance of $|\Psi^-\rangle$. The state $|\Psi^-\rangle$ is mathematically equivalent to a singlet spin state, a spin-zero state composed of two spin-one-half particles. The other Bell states are *not* rotationally invariant. This property, or the lack of it, has no bearing on what follows.

We are now in a position to discuss tests of quantum mechanics against the predictions of local realistic theories through a theorem enunciated 40 years ago by John Bell [1]. But first we should describe how photon polarization states can be measured. It is really very simple and requires only a piece of birefringent (double-refracting) material such as calcite, familiar in classical optics, and photon detectors. The action of a calcite crystal on an unpolarized light beam is to

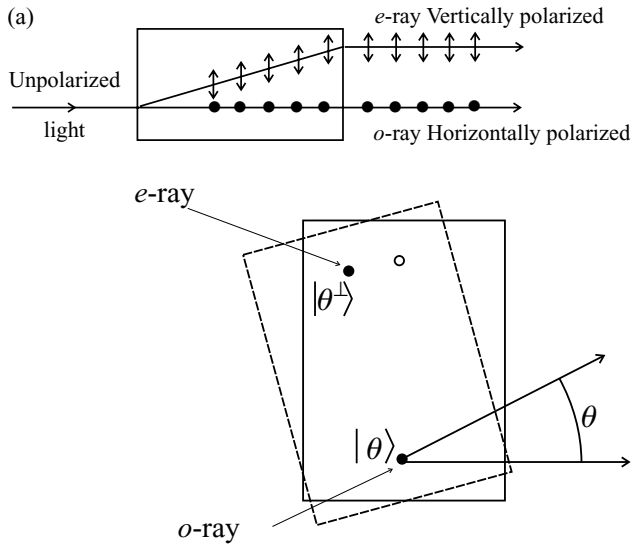


Fig. 9.12. (a) A birefringent material such as a calcite crystal splits an unpolarized beam into two beams of orthogonal polarizations. The *o*-ray is undeflected while the *e*-ray, whose polarization is orthogonal to that of the *o*-ray, is refracted. (b) In rotating the crystal about an axis through the *o*-ray, the polarizations of the two beams are still orthogonal.

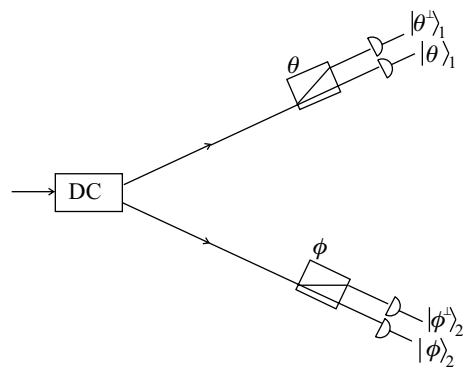


Fig. 9.13. Schematic of a test of Bell's theorem using down-converted photons. The down-converter could be of type II or of type I with a polarization rotator placed in one of the beams.

separate the beam into horizontally and vertically polarized beams where the former, called the ordinary (*o*) beam (or ray) is along the path of the input beam and the latter, called the extraordinary (*e*) beam (ray) emerges along a different path as illustrated in Fig. 9.12(a). Operationally, the polarization state of a single photon is determined by observing which beam emits the photon: $|H\rangle$ if in the *o*-ray and $|V\rangle$ if in the *e*-ray. If we rotate the calcite crystal about the incident beam, the *o*-ray remains unchanged while the *e*-ray follows the rotation as indicated in Fig. 9.12(b). For a crystal rotated by an angle θ, single photons emerging in the *o*-ray are in the state $|\theta\rangle$ and those in the *e*-ray are in the state $|\theta^{\perp}\rangle$.

Suppose we have prepared a state of the form of Eq. (9.39) and that we now place calcite crystals in each of the beams in order to do polarization measurements, as indicated in Fig. 9.13. We have assumed that the beams are propagating in directions such that the calcite measuring devices are widely separated. Ideally, we need to have the separations great enough so that the measurement events have

a space-like space-time interval. Assuming this to be the case, and assuming both calcite crystals are set at an angle θ to the horizontal, here is what will happen. If the photon in mode 1 is found to be in the polarization state $|\theta\rangle_1$, then the other photon will be detected in the state $|\theta^\perp\rangle_2$, and vice versa. This is, of course, owing to the strong correlations that exist in the state of Eq. (9.39). But there is a mystery that goes to the heart of the interpretation of the quantum-mechanical state vector. Does the state described by Eq. (9.39) mean that, for a given pair of photons, the polarization of each within the pair is determined at the time of formation as either $|\theta\rangle_1|\theta^\perp\rangle_2$ or $|\theta^\perp\rangle_1|\theta\rangle_2$ and the measurements merely reveal which; or is it the measurement itself that randomly collapses one of the photons onto a particular polarization state and somehow forces the other, very distant, photon into the state with orthogonal polarization? If the latter, it would seem that some sort of violation of locality must be involved. If the measurements are performed with a space-like interval, how do the distant photons "know" what states the detectors must see in order to maintain the correlations apparent in their two-particle entangled state $|\Psi^-\rangle$? But there is more mystery. To illustrate that photon polarization is isomorphic to spin one-half, we may construct the following operators

$$\begin{aligned}\hat{\sum}_1 &= |\theta\rangle\langle\theta^\perp| + |\theta^\perp\rangle\langle\theta|,\\ \hat{\sum}_2 &= -i(|\theta\rangle\langle\theta^\perp| - |\theta^\perp\rangle\langle\theta|),\\ \hat{\sum}_3 &= |\theta\rangle\langle\theta| - |\theta^\perp\rangle\langle\theta^\perp|,\end{aligned} \tag{9.40}$$

which satisfy the algebra $[\hat{\sum}_i, \hat{\sum}_j] = 2i\varepsilon_{i,j,k}\hat{\sum}_k$, the same algebra as satisfied by the Pauli spin operators. Thus the Heisenberg uncertainty principle comes into play with regard to the components of the polarization. If, for example, the photon is known to be in the state $|\theta\rangle$ giving $\hat{\sum}_3$ the value of 1, then the values of $\hat{\sum}_1$ and $\hat{\sum}_2$ are undetermined as these operators do not commute with each other or with $\hat{\sum}_3$. But in the case of the Bell state of Eq. (9.39), we have the following situation. If mode 1 is detected in, say, state $|\theta\rangle_1$, then it follows that we can say with certainty that *if* a measurement *were* performed on beam 2 with the calcite again along θ, the result would be $|\theta^\perp\rangle_2$. But if we do no such measurement, can we conclude that the photon nevertheless must be in the state $|\theta^\perp\rangle_2$? If we draw such a conclusion, i.e. that the photon in mode 2 is in the eigenstate $|\theta^\perp\rangle_2$ of that mode's "spin" operator $\hat{\sum}_3^{(2)}$, then we are free to perform a measurement of, say, the operator $\hat{\sum}_1^{(2)}$, and find the photon in one of the eigenstates of that operator, which in turn implies a violation of the uncertainty relation as these operators do not commute.

The conundrum presented by the above discussion was first addressed, in a slightly different form, by Einstein, Podolsky and Rosen (EPR) in 1935 [15] and later by Bohm [16] in the context of a spin singlet state, mathematically identical to polarization states above. We are faced with two problems: the locality problem

and the problem of assigning definite values to observables at all times, the problem of "realism", what EPR referred to as the "elements of reality". The lack of determinate elements of reality led EPR to the conclusion that quantum mechanics must be an incomplete theory. Regarding the locality issue, we would expect a reasonable physical theory to be local. That is, it should not predict effects having a physical consequence, such as the possibility of superluminal signaling. As for the latter, classical-like reasoning might suggest the existence of "hidden variables" which supplement, and complete, quantum mechanics. In 1952, Bohm [2] proposed a hidden variable theory that reproduced all the results of quantum mechanics. But it was nonlocal. Bell, in 1964 [1], was able to show that *local* hidden variable theories make predictions that differ from those of quantum mechanics and thus provide a way to falsify either quantum mechanics or local hidden variable theories.

To see how this comes about, we follow Bell and define the correlation function

$$C(\theta, \phi) = \text{Average}\left[A(\theta)\, B(\phi)\right], \tag{9.41}$$

where the average is taken over a large number of experimental runs. The quantity where $A(\theta) = +1$ if photon 1 comes out in the o-beam (i.e. it is in the state $|\theta\rangle_1$) and -1 if it comes out in the e-beam (the state $|\theta^\perp\rangle_1$). A similar definition holds for $B(\phi)$ with regard to photon 2. As for the product $A(\theta)B(\phi)$, it has the value $+1$ for the outcome pairs $|\theta\rangle_1|\phi\rangle_2$ and $|\theta^\perp\rangle_1|\phi^\perp\rangle_2$ and -1 for the pairs $|\theta\rangle_1|\phi^\perp\rangle_2$ and $|\theta^\perp\rangle_1|\phi\rangle_2$. The average then must be

$$\begin{aligned} C(\theta, \phi) = {} & \Pr(|\theta\rangle_1|\phi\rangle_2) + \Pr(|\theta^\perp\rangle_1|\phi^\perp\rangle_2) \\ & - \Pr(|\theta\rangle_1|\phi^\perp\rangle_2) - \Pr(|\theta^\perp\rangle_1|\phi\rangle_2), \end{aligned} \tag{9.42}$$

where $\Pr(|\theta\rangle_1|\phi\rangle_2)$ is the probability of getting the outcomes $|\theta\rangle_1|\phi\rangle_2$, etc. For the state $|\Psi^-\rangle$ of Eq. (9.38) we can read off the probability amplitudes and then square them to obtain, after a bit of manipulation,

$$C(\theta, \phi) = -\cos[2(\theta - \phi)]. \tag{9.43}$$

The same result is obtained from the expectation value

$$C(\theta, \phi) = \langle\Psi^-|\hat{\sum}_3^{(1)}(\theta)\hat{\sum}_3^{(2)}(\phi)|\Psi^-\rangle = -\cos[2(\theta - \phi)], \tag{9.44}$$

where

$$\begin{aligned} \hat{\sum}_3^{(1)}(\theta) &= |\theta\rangle_{11}\langle\theta| - |\theta^\perp\rangle_{11}\langle\theta^\perp|, \\ \hat{\sum}_3^{(2)}(\phi) &= |\phi\rangle_{22}\langle\phi| - |\phi^\perp\rangle_{22}\langle\phi^\perp|, \end{aligned} \tag{9.45}$$

are the "spin" projections along a fictitious "3" axis. Clearly these operators have as eigenvalues the numbers specified above for the values of $A(\theta)$ and $B(\phi)$. Thus the correlation function of Eq. (9.43) is a straightforward quantum mechanical prediction.

To incorporate locality, we assume the existence of hidden variables, denoted λ, whose role is to determine the correlation between the photons in a pair when

they are created. The variables can vary randomly from one run of the experiment to the next but the values of $A(\theta)$ and $B(\phi)$ must depend on the hidden variables for a given run. Thus we should write these now as $A(\theta, \lambda)$ and $B(\phi, \lambda)$, where we still have $A(\theta, \lambda) = \pm 1$ and $B(\phi, \lambda) = \pm 1$. According to the principle of locality, the result of a measurement on one photon cannot depend on the angle setting of the calcite for the measurement of the other, and vice versa. For this reason we exclude functions of the form $A(\theta, \phi, \lambda)$ and $B(\theta, \phi, \lambda)$. We assume that for a large number of runs of the experiment, $\rho(\lambda)$ represents the probability distribution of the hidden variables λ such that

$$\int \rho(\lambda)\, d\lambda = 1. \tag{9.46}$$

Thus for the local hidden-variable theory the correlation function will be

$$C_{\mathrm{HV}}(\theta, \phi) = \int A(\theta, \lambda) B(\phi, \lambda)\, \rho(\lambda)\, d\lambda. \tag{9.47}$$

We must determine whether or not this expression is consistent with the quantum-mechanical expression in Eq. (9.43). To this end, we need the following simple mathematical result: if $X_1, X_1', X_2, X_2' = \pm 1$, then we have

$$\begin{aligned} S &= X_1 X_2 + X_1 X_2' + X_1' X_2 - X_1' X_2' \\ &= X_1(X_2 + X_2') + X_1'(X_2 - X_2') = \pm 2, \end{aligned} \tag{9.48}$$

as can be easily checked. If we set $X_1 = A(\theta, \lambda)$, $X_1' = A(\theta', \lambda)$, $X_2 = B(\phi, \lambda)$, and $X_2' = B(\phi', \lambda)$, multiply by $\rho(\lambda)$ and integrate over λ, we obtain

$$-2 \leq C_{\mathrm{HV}}(\theta, \phi) + C_{\mathrm{HV}}(\theta', \phi) + C_{\mathrm{HV}}(\theta', \phi) - C_{\mathrm{HV}}(\theta', \phi') \leq +2. \tag{9.49}$$

This is a Bell inequality in the form given by Clauser, Horne, Shimony and Holt [17], the CHSH inequality for short. If we set $\theta = 0, \theta' = \pi/4$, $\phi = \pi/8$, and $\phi' = -\pi/8$, we find from the quantum mechanical correlation function of Eq. (9.43) that

$$S = C(\theta, \phi) + C(\theta', \phi) + C(\theta', \phi) - C(\theta', \phi') = 2\sqrt{2}, \tag{9.50}$$

a result clearly outside the bounds placed by the inequality of Eq. (9.49) obtained from local hidden-variable theory. This violation of the CHSH inequality means that the quantum-mechanical result cannot be explained by any local hidden-variable theory. Of course, experiment must be the final arbiter in this matter.

Experimental tests of inequalities of the Bell type with polarized light generated from atomic $J = 0 \rightarrow 1 \rightarrow 0$ cascade transitions in calcium began in the early 1970s with the experiment of Freedman and Clauser [18], culminating in the experiments of Aspect *et al*. in 1982 [19,20]. In the Aspect experiments, the polarization analyzers were rotated while the photons were in flight. Subsequently, and independently, Ou and Mandel [21] and Shih and Alley [22] performed the first test using down-conversion as a source of polarization-entangled photons.

Most of the experiments exhibit violations of the inequalities tested, but not without certain supplementary assumptions being used in the data analysis. In any real experiment done with photons, one must take into account the fact that the detector efficiency, which we represent by $\eta_{\rm det}$, is never at the ideal, $\eta_{\rm det} = 1$, although we assumed the contrary in deriving our results. For the nonideal case, Eq. (9.43) should be replaced by

$$C(\theta, \phi) = -\eta_{\rm det}^2 \cos[2(\theta - \phi)]. \tag{9.51}$$

This comes about because each term in Eq. (9.42) is multiplied by $\eta_{\rm det}^2$ which, in turn, is the result of the fact that each term is a joint probability of two measurements where a factor of $\eta_{\rm det}$ appears from each of the measurements. Thus, from Eq. (9.51), we must now have $S = \eta_{\rm det}^2 2\sqrt{2}$. Recalling that Bell's inequality is violated whenever $S > 2$, it is evident that we must place the constraint on the detector efficiency as $\eta_{\rm det} > 1/\sqrt[4]{2} \approx 0.84$ in order to obtain a violation.

With inefficient detectors, not all photon pairs will register counts. In a practical experiment it is necessary to introduce an untestable supplementary assumption. The assumption is that the detectors perform a fair sampling of the ensemble of all events. That is, that those events for which both photons are detected are representative of the entire ensemble. We thus redefine the correlation function according to

$$C(\theta, \phi) = \frac{\text{Average}[A(\theta)\, B(\phi)]}{\text{Average}[N_1 N_2]} \tag{9.52}$$

where N_1 is the total number of photons detected in beam 1 (in any given run, $N_1 = 0$ or 1) with a similar definition for N_2, where

$$\begin{aligned}\text{Average}[N_1 N_2] = {} & \Pr(\theta, \phi) + \Pr(\theta^\perp, \phi^\perp) \\ & + \Pr(\theta, \phi^\perp) + \Pr(\theta^\perp, \phi).\end{aligned} \tag{9.53}$$

Because both numerator and denominator are proportional to the factor $\eta_{\rm det}^2$, the correlation function will be independent of it. With this fair sampling assumption, the CHSH for of Bell's inequality will be violated according to quantum mechanics for detectors of any efficiency.

Our description of these experiments is, of course, considerably oversimplified. Nevertheless, the experiments, or experiments like them, are not too complicated to be done, at modest expense, even in an undergraduate laboratory setting. Dehlinger and Mitchell [23] have described an apparatus for the generation of polarization-entangled states that may be used to demonstrate the violation of Bell's inequalities in the undergraduate laboratory. Their approach is to use two face-to-face type I BBO down-converters, where the second is rotated by 90° around the normal to the first. The crystals are pumped by a beam of 45° polarized light obtained from a diode laser. This arrangement produces an entangled state of the form

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1 |H\rangle_2 + |V\rangle_1 |V\rangle_2). \tag{9.54}$$

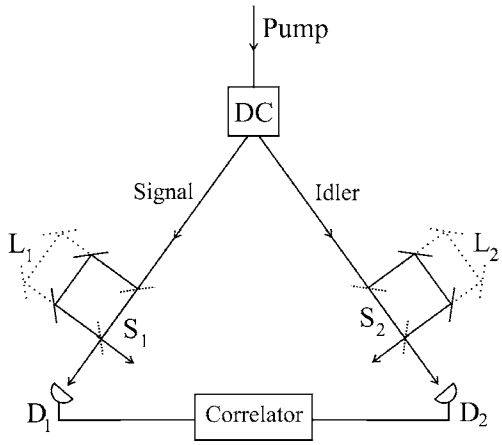


Fig. 9.14. Schematic for Franson's experiment. The path lengths of the interferometer are adjustable as indicated.

The two-crystal scheme described is an efficient method of generating polarization states and was first used by Kwiat *et al*. [24]. Their results violated Bell's inequality by 21 standard deviations.

Not all experiments demonstrating a violation of a Bell-type inequality using down-converted light have involved polarization entanglement. Rarity and Tapster [25] have done experiments with momentum entanglement of the beams. Another, proposed by Franson [26], involves time–energy uncertainty. We now discuss this experiment.

9.7 Franson's experiment

Consider the experimental setup pictured in Fig. 9.14. A pair of photons, signal and idler, is simultaneously produced by the down-converter and directed toward the interferometers. There is no single-photon interference inside a single interferometer as the photon coherence lengths are much shorter than the differences between path lengths of the arms of the interferometers. However, there is two-photon interference in the coincidence detection between the detectors D_1 and D_2. To see how this comes about, consider the following. To reach the detectors, the photons can both take the short paths (S, S), both take the long paths (L, L), the short–long (S, L) or the long–short (L, S) paths. The first two cases are indistinguishable as we do not know when the photons are created. Both detectors will tend to fire simultaneously, assuming identical path lengths in the two interferometers. The last two cases are distinguishable because of the delay between the "clicking" of the two detectors, the detector of the photon taking the short path will "click" first. So, the last two cases are distinguishable from each other *and* from the first two cases in which the detectors "click" simultaneously. Under experimental conditions, the fast electronics of the correlator can be set for a sufficiently narrow timing window such that counts from the

two distinguishable processes are rejected. This has the effect of post-selecting (reducing) the two-photon output state to be

$$|\psi\rangle = \frac{1}{2}(|S\rangle_1|S\rangle_2 + e^{i\Phi}|L_1\rangle_1|L_2\rangle_2), \tag{9.55}$$

evidently another form of Bell state, where we have assumed that $S_1 = S_2 = S$ but that L_1 and L_2 can be adjusted as indicated in Fig. 9.14. The phase Φ, the relative phase between the (S, S) and (L, L) processes, is the sum of the relative phases acquired by the individual photons:

$$\begin{aligned}\Phi &= \omega_s \Delta L_1/c + \omega_i \Delta L_2/c \\ &= \frac{\omega_s + \omega_i}{2}(\Delta L_1 + \Delta L_2) + \frac{\omega_s - \omega_i}{2}(\Delta L_1 - \Delta L_2) \\ &\approx \frac{\omega_p}{2}(\Delta L_1 + \Delta L_2)\end{aligned} \tag{9.56}$$

valid as $\Delta L_1 - \Delta L_2$ is taken to be small compared to the inverse bandwidth of the signal and idler frequencies ω_s and ω_i. The frequency $\omega_p = \omega_s + \omega_i$ is, of course, the frequency of the pump field. If we now use the Feynman dictum of adding amplitudes associated with indistinguishable processes, the probability for coincident two-photon detection is

$$\begin{aligned}P_{\text{coin}} &= \frac{1}{4}|1 + e^{i\Phi}|^2 = \frac{1}{2}[1 + \cos\Phi] \\ &= \frac{1}{2}\left[1 + \cos\left(\frac{\omega_p}{2}(\Delta L_1 + \Delta L_2)\right)\right].\end{aligned} \tag{9.57}$$

This result exhibits 100% visibility, meaning that the minimum of the probability of coincident detection is zero. This in turn means that the photons entering the two interferometers become *anti-correlated*: if one takes the short path the other takes the long path, and vice versa. The meaning of the maximum is that the photons become *correlated* in the interferometers: either both take the short path or both take the long. No classical or hidden variable model can predict a visibility greater that 50%. Fringes of visibility greater than 70.7% violate a Bell-type inequality. The experiments of Kwiat *et al*. [27] exhibit a visibility of 80.4±0.6%. Other experimental realizations of the Franson experiment are those of Ou *et al*. [28], Brendel *et al*. [29], and Shih *et al*. [30].

9.8 Applications of down-converted light to metrology without absolute standards

Finally in this chapter, we briefly discuss the application of down-converted light to problems of practical interest. These applications are possible because of the tight correlations of the photons that are spontaneously emitted during the down-conversion process as was originally shown by Burnham and Weinberg [3]. The virtue of using down-converted light is, as we shall see, that measurements can be performed to yield absolute results without the need of a calibrated standard.

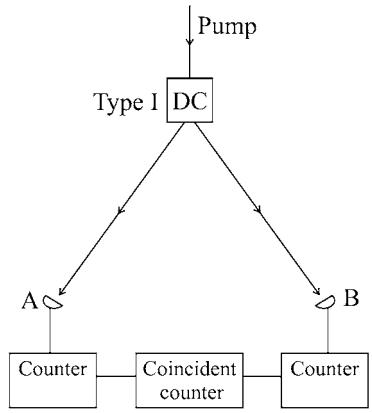


Fig. 9.15. Schematic for absolute calibration of a photon detector.

We first discuss the determination of the absolute calibration of a photon detector, essentially measuring its absolute quantum efficiency. A schematic of the measurement is given in Fig. 9.15. Two identical photon detectors are placed in the outputs of a type I down-converter. We regard detector A as the detector to be calibrated and detector B as the trigger, though which is which is quite arbitrary. Because the down-converter creates photons in pairs, coincident photons would be detected in both arms for ideal detectors. The measurement of the quantum efficiency is quite simple. Suppose that N is the number of pairs of photons produced by the down-converter. Whenever detector B fires, the experimenter checks to see if detector A also fires. Then the fraction of detections by B for which there is a coincident detection by A is the measured quantum efficiency of detector A. The method does not require knowledge of the quantum efficiency of detector B. If there is no detection by B, the experimenter simply disregards detector A altogether. More quantitatively, if we let the quantum efficiencies of detectors A and B be represented respectively by η_A and η_B, then the number of photons detected by each is $N_A = \eta_A N$ and $N_B = \eta_B N$. The number of coincident counts will be given by $N_C = \eta_A \eta_B N = \eta_A N_B$. Thus $\eta_A = N_B/N_C$, which is obviously independent of the quantum efficiency of detector B. The idea behind this sort of measurement is implicit in the work of Burnham and Weinberg [3]. Migdall *et al*. [31] at NIST performed a test of the method in 1995, comparing the method with a more-conventional method involving a standard, and found good agreement.

Another metrologic application, the only other one we discuss, is to the measurement of absolute radiance. This application was proposed by Klyshko [32] and was first demonstrated in the laboratory in 1979 [33]. It is a variation of the process of spontaneous parametric down-conversion where a light beam whose radiance is to be measured (the unknown) is aligned to overlap with one of the output beams, say, the signal beam, of the down-converter as indicated in Fig. 9.16. The setup constitutes a parametric amplifier where the presence of the unknown input beam stimulates the production of photon pairs (they must always be

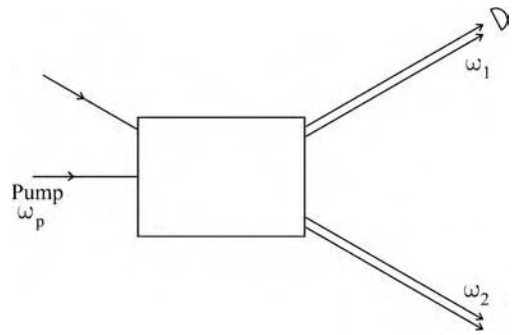


Fig. 9.16. Schematic for the measurement of absolute radiance.

produced in pairs) into the output beams. This is indicated in Fig. 9.16 by the double lines at the outputs. The idea is to monitor the output of the idler beam and to determine the increase in strength of the radiation in going from spontaneous down-conversion, where the unknown beam is blocked, to the case of stimulated down-conversion, when the unknown input beam is unblocked. The ratio of the strengths of the stimulated and spontaneous output signals yields the radiance of the unknown beam. No standard is required in such a measurement.

The applications discussed here have not involved entanglement. However, applications that do involve entanglement, and thus require type II down-conversion have been proposed. A specific example is the application to ellipsometry [34]. For reviews with many references, see the papers listed in the bibliography.

Problems

1. In the perturbation expansion of Eq. (9.8), in first order the output of a type I down-converter is the pair state $|1\rangle_s|1\rangle_i$. (a) Show that, in the second order, the output is the pair state $|2\rangle_s|2\rangle_i$. (b) Supposing that the two photons within each of the states, and both pairs of the photons, simultaneously fall onto the inputs of a 50:50 beam splitter, what will be the output state? (c) Show that your result does not agree with that which would be obtained classically using a simple binomial distribution. (d) Devise an experiment able to discriminate between the quantum mechanical and classical predictions for the oputput of the beam splitter. (*After* working through this problem you may wish to consult a paper by Ou, Rhee and Wang [35].)
2. Derive a CHSH–Bell inequality starting from the Bell state

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 + |V\rangle_1|H\rangle_2).$$

3. As mentioned in the text above and in References [23] and [24], by using back-to-back type I down-converters, where one is rotated by 90° around the normal to the other and where both are pumped by light polarized at 45°, the Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|H\rangle_2 + |V\rangle_1|V\rangle_2)$$

can be produced. Analyse the possible violations of Bell's inequalities (in the form of CHSH) with respect to this state.

4. For the state $|\Psi^-\rangle$, consider the quantity

$$S_{\text{Bell}} \equiv |C(\theta, \phi) - C(\theta, \phi')| + |C(\theta', \phi') + C(\theta', \phi)|.$$

Show, for a local realistic theory, where we replace $C(\theta, \phi)$ by $C_{\text{HV}}(\theta, \phi)$, that $S_{\text{Bell}} \leq 2$. Find an example where quantum mechanics violates the inequality. (This is the inequality originally discussed by Bell.)

References

[1] J. S. Bell, *Physics* (*NY*), **1** (1964), 195. See also J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press, 1987).

[2] D. Bohm, *Phys. Rev*. **85** (1952), 166, 180.

[3] D. C. Burnham and D. L. Weinberg, *Phys. Rev. Lett.*, **25** (1970), 84.

[4] A. Joobeur, B. Saleh and M. Teich, *Phys. Rev. A*, **50** (1994), 3349.

[5] See, for example, R. W. Boyd, *Nonlinear Optics*, 2nd edition (San Diego: Academic Press, 2003).

[6] C. K. Hong, Z. Y. Ou and L. Mandel, *Phys. Rev. Lett.*, **59** (1987), 2044. A related experiment was later performed by J. G. Rarity, P. R. Tapster, E. Jakeman, T. Larchuk, R. A. Campos, M. C. Teich and B. E. A. Saleh, *Phys. Rev. Lett.*, **65** (1990), 1348.

[7] P. G. Kwiat, A. M. Steinberg and R. Y. Chiao, *Phys. Rev. A*, **45** (1992), 7729.

[8] D. M. Greenberger, M. A. Horne and A. Zeilinger, *Physics Today* (August 1993), p. 22.

[9] X. Y. Zou, L. J. Wang and L. Mandel, *Phys. Rev. Lett.*, **67** (1991), 318.

[10] R. Y. Chiao and A. M. Steinberg, *Progress in Optics XXXVII*, E. Wolf (editor) (Amsterdam: Elsevier, 1997), p. 345; G. Nimtz and W. Heitmann, *Prog. Quant. Electr.*, **21** (1997), 81.

[11] A. M. Steinberg, P. G. Kwiat and R. Y. Chiao, *Phys. Rev. Lett*., **71** (1993), 708.

[12] See H. Winful, *Phys. Rev. Lett.*, **90** (2003), 023901; M. Büttiker and S. Washburn, *Nature*, **422** (2003), 271; and the exchange between Winful and Büttiker and Washburn, *Nature* **424** (2003), 638.

[13] See any good optics text such as F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*, 2nd edition (Englewood Cliffs: Prentice Hall, 1993). For a discussion of polarized single-photon states and the use of calcite crystals to detect them, along with a discussion of their use in elucidating a number of puzzling features in quantum mechanics, see A. P. French and E. F. Taylor, *An Introduction to Quantum Physics* (New York: W. W. Norton, 1978), chapter 6.

[14] See W. Tittel and G. Weihs, *Quant. Inf. Comp.*, **1** (2001), 3; Y. Shih, *Rep. Prog. Phys.*, **66** (2003), 1009.

[15] A. Einstein, B. Podolsky and N. Rosen, *Phys. Rev.*, **47** (1935), 777.

[16] D. Bohm, *Quantum Theory* (New York: Prentice-Hall, 1951), chapter 22.

[17] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, *Phys. Rev. Lett.*, **23** (1969), 880.

[18] S. J. Freedman and J. F. Clauser, *Phys. Rev. Lett.*, **28** (1972), 938.

[19] A. Aspect, P. Grangier and G. Roger, *Phys. Rev. Lett.*, **49** (1982), 91.

[20] A. Aspect, J. Dalibard and G. Roger, *Phys. Rev. Lett.*, **49** (1982), 1804.

[21] Z. Y. Ou and L. Mandel, *Phys. Rev. Lett.*, **61** (1988), 50.
[22] Y. H. Shih and C. O. Alley, *Phys. Rev. Lett.*, **61** (1988), 2921.
[23] D. Dehlinger and M. W. Mitchell, *Am. J. Phys.*, **70** (2002), 898, 903.
[24] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko and Y. H. Shih, *Phys. Rev. Lett.*, **75** (1995), 4337; P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum and P. H. Eberhard, *Phys. Rev. A*, **60** (1999), R773.
[25] J. G. Rarity and P. R. Tapster, *Phys. Rev. Lett.*, **64** (1990), 2495.
[26] J. D. Franson, *Phys. Rev. Lett.*, **62** (1989), 2205.
[27] P. G. Kwiat, A. M. Steinberg and R. Y. Chiao, *Phys. Rev. A*, **47** (1993), R2472.
[28] Z. Y. Ou, X. Y. Zou, L. J. Wang and L. Mandel, *Phys. Rev. Lett.*, **65** (1990), 321.
[29] J. Brendel, E. Mohler and W. Martienssen, *Euro. Phys. Lett.*, **20** (1992), 575.
[30] Y. H. Shih, A. V. Sergienko and M. H. Rubin, *Phys. Rev. A*, **47** (1993), 1288.
[31] A. Migdall, R. Datla, A. Sergienko, J. S. Orszak and Y. H. Shih, *Metrologica*, **32** (1995/6), 479.
[32] D. N. Klyshko, *Sov. J. Quantum Electron.*, **7** (1977), 591.
[33] G. Kitaeva, A. N. Penin, V. V. Fadeev and Yu. A. Yanait, *Sov. Phys. Dokl.*, **24** (1979), 564.
[34] A. F. Abouraddy, K. C. Toussaint, A. V. Sergienko, B. E. A. Saleh and M. C. Teich, *J. Opt. Soc. Am. B*, **19** (2002), 656.
[35] Z. Y. Ou, J.-K. Rhee and L. J. Wang, *Phys. Rev. Lett.*, **83** (1999), 959.

Bibliography

D. N. Klyshko, *Photons and Nonlinear Optics* (New York: Gordon and Breach, 1988).
R. Y. Chiao and A. M. Steinberg, "Tunneling times and superluminality", in *Progress in Optics XXXVII*, E. Wolf (editor) (Amsterdam: Elsevier, 1997), p. 345.
R. Y. Chiao, P. G. Kwiat and A. M. Steinberg, "Faster than light?", *Scientific American* (August 1993), p. 52.
P. Hariharan and B. C. Sanders, "Quantum phenomena in optical interferometry", *Progress in Optics XXXVI*, E. Wolf (editor) (Amsterdam: Elsevier, 1996).
L. Mandel, "Quantum effects in one-photon and two-photon interference", *Rev. Mod. Phys.*, **71** (1999), S274.

The discussion of the Bell inequalities above follows closely that of the following.
L. Hardy, "Spooky action at a distance in quantum mechanics", *Contemp. Phys.*, **39** (1998), 419.

The early experiments using atomic cascades to generate polarization entangled photons are reviewed in the following.
J. F. Clauser and A. Shimony, "Bell's theorem: experimental tests and implications", *Rep. Prog. Phys.*, **41** (1978), 1881.

Two articles that review applications of down-converted light to metrology are:
A. Migdall, "Correlated-photon metrology without absolute standards", *Physics Today* (January 1999), p. 41.
A. V. Sergienko and G. S. Sauer, "Quantum information processing and precise optical measurements with entangled photon pairs", *Contemp. Phys.*, **44** (2003), 341.

Chapter 10

Experiments in cavity QED and with trapped ions

In this chapter, we discuss two more experimental realizations of quantum optical phenomena, namely the interaction of an effective two-level atom with a quantized electromagnetic field in a high Q microwave cavity, the subject usually referred to as cavity QED, or sometimes CQED, and in the quantized motion of a trapped ion. Strictly speaking, these experiments are not optical, but they do realize interactions of exactly the type that are of interest in quantum optics, namely the Jaynes–Cummings interaction between a two-level system (an atom) and a bosonic degree of freedom, a single-mode cavity field in the case of a microwave cavity, and a vibrational mode of the center-of-mass motion of a trapped ion, the quanta being phonons in this case. We shall begin with a description of the useful properties of the so-called Rydberg atoms that are used in the microwave CQED experiments, proceed to discuss some general considerations of the radiative behavior of atoms in cavities, the CQED realization of the Jaynes–Cummings model, and then discuss the use of the dispersive, highly off-resonant, version of the model to generate superpositions of coherent states, i.e. the Schrödinger cat states of the type discussed in Chapters 7 and 9 for traveling wave optical fields but this time for a microwave cavity field. Finally, we discuss the realization of the Jaynes–Cummings interaction in the vibrational motion of a trapped ion.

10.1 Rydberg atoms

A Rydberg atom is an ordinary atom where one of its electrons, usually the valence electron in an alkali atom, is excited to a state of very high principal quantum number [1], a Rydberg state. In the experiments to be discussed, the Rydberg atoms used are those of rubidium with the principal quantum number n (not to be confused with the photon number of a field state as should be clear by context) of the valence electron in the vicinity of $n \sim 50$. The electronic binding energy of a Rydberg state is given by

$$E_{nl} = -\frac{R}{(n-\delta_l)^2} = -\frac{R}{n^{*2}} \tag{10.1}$$

where $n^* = n - \delta_l$ and where δ_l is the "quantum defect" due to the hydrogenic "core" which corrects for the deviation of the binding potential from a purely hydrogenic situation. The parameter R is the Rydberg constant having the approximate value $R = 13.6\,\text{eV}$. The quantum defect does depend on the angular momentum quantum number l and has a value of unity for low l states, decreasing for higher l. As we are interested mainly in the high l where δ_l is small, we may characterize our Rydberg states with the usual principal quantum number n. States for which l takes on the highest value allowed for a given n, $l = n - 1$, and with $|m| = n - 1$, where m is the magnetic quantum number, are known as circular Rydberg states as in the classical limit they describe an electron in a circular orbit [2]. It is the circular Rydberg states that are used in the CQED experiments. Nussenzweig *et al*. [3] have described how such states may be prepared in the case of rubidium.

So what are the special properties of circular Rydberg atoms that make them suitable for CQED experiments? They are numerous. To begin, only one electric dipole transition is allowed, $n \leftrightarrow n - 1$, $|m| \leftrightarrow |m| - 1$, so that such states closely approximate a two-level system. The "classical" radius of a Rydberg atom scales as $n^2 a_0$, where $a_0 = 0.5\,\text{Å}$ is the Bohr radius. For $n \sim 100$ the atom is about the size of a virus. (Rydberg atoms making transitions from states with principal quantum numbers as high as $n \sim 733$ have been observed astrophysically.) The matrix element of the electric dipole operator between states between two circular Rydberg states of principal quantum numbers n and n' for $\Delta n = n - n' = 1$ is

$$d = \langle n|\hat{d}|n'\rangle \sim q n^2 a_0 \tag{10.2}$$

where $\hat{d} = q\,r$. For $n = 50$ we obtain $d \sim 1390$ A.U., where $q\,a_0 = 1$ A.U., a value for the dipole moment close to 300 times those of typical optical transitions for low-lying hydrogenic states. The large dipole moments of the circular Rydberg atoms make them particularly attractive for CQED experiments as their coupling to a single-mode cavity field can be quite large. For a Rydberg atom in large n state, the frequency of the radiation emitted in a transition for which $\Delta n = 1$ is close to the Bohr orbital frequency itself:

$$\omega_0 = \frac{E_{nl} - E_{n'l'}}{\hbar} \simeq \frac{2R}{\hbar n^3}. \tag{10.3}$$

For $n \sim 50$, the frequency of the emitted radiation is $\nu_0 = \omega_0/2\pi \sim 36\,\text{GHz}$ which corresponds to a wavelength of $\lambda_0 = c/\nu_0 \sim 8$ mm. This last number gives a sense of the dimensions required of a cavity to support a standing microwave field.

The rate of spontaneous emission is given by

$$\Gamma = \frac{d^2\omega_0^3}{3\pi\varepsilon_0\hbar c^3}\ . \tag{10.4}$$

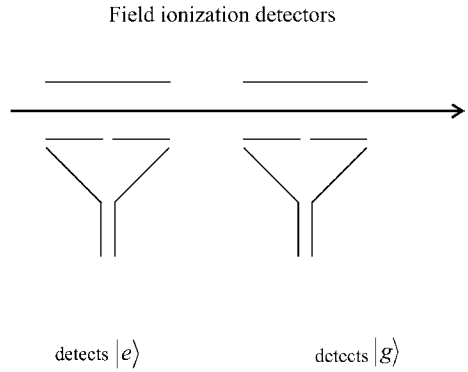


Fig. 10.1. Field ionization of an atomic beam of highly excited Rydberg atoms allows for selective state detection. An electric field is applied that ramps up so that the first detector monitors the excited-state population whereas the stronger field downstream monitors the more strongly bound ground state.

Using the previous results for circular Rydberg transitions, $d^2 \sim n^4$ and $\omega_0^3 \sim n^{-9}$, we find that the decay rate for these transitions is

$$\Gamma \sim \Gamma_0 n^{-5}, \tag{10.5}$$

where $\Gamma_0 = c\alpha^4/a_0, \alpha$ here being the fine-structure constant $\alpha = e^2/\hbar c \approx 1/137$. The rate Γ_0 is that for the usual strongly allowed lower level spontaneous emissions (we normally deal with optical transitions) and typically has an order of magnitude $\Gamma_0 \sim 10^9$ s^{-1} with corresponding excited state lifetimes $\tau_0 = 1/\Gamma_0 \sim 10^{-9}$s. The circular Rydberg transitions will have a lower spontaneous emission rate and longer excited state lifetimes given by $\tau = 1/\Gamma = \tau_0 n^5$. For $n \sim 50$ we obtain thus $\tau \sim 10^{-1}$ s, a very long lifetime indeed compared with the "ordinary" spontaneous emission lifetimes of around $\tau_0 \sim 10^{-9}$ s.

Finally, we discuss the issue of atomic-state detection. For a Rydberg atom, the electron is typically far from the hydrogenic core and its binding energy is rather small, so this allows relatively easy ionization by an external applied field. The electric field normally felt by the electron, with no applied field, from Coulomb's law is

$$E_{\text{Coulomb}} \approx \frac{e}{4\pi\varepsilon_0 (n^2 a_0)^2}, \tag{10.6}$$

and thus the ionization rate with an applied constant uniform field $E \sim E_{\text{Coulomb}}$ goes as n^{-4}. The ionization rate of the Rydberg state with principal quantum number n is smaller than a state with $n-1$ by a factor of $(1+4/n)$. Though the difference in the rates for adjacent Rydberg levels is small in comparison with the case for low-lying levels, it is sufficiently large for state-selective measurements to be performed by field ionization. The idea is quite simple: one allows the atom to pass through a set of two ionization detectors, each having different field strengths (see Fig. 10.1). The atom must first encounter the one with the field set to detect the higher of two states, say the state with quantum number n. The second detector has a stronger applied field to detect the state with quantum number $n-1$. The strength of the applied fields must be carefully set so that the first detector will ionize the atom only if the electron is in the upper level while

that of the second must be strong enough to ionize the lower. Of course, it is the detection of the field-ionized electron in one of the detectors that constitutes selective atomic-state detection.

In summary, circular atomic Rydberg states have the following desirable properties: (1) the transitions are restricted to other circular states thus setting the stage for them to act as "two-level" atoms, (2) the dipole moments of the allowed transitions are large, (3) the states have long lifetimes, and (4) they may be selectively ionized by applied fields to achieve selective atomic state detection.

10.2 Rydberg atom interacting with a cavity field

As we have shown, the wavelengths of the radiation emitted as the result of transitions between neighboring circular Rydberg states are in the range of a few millimeters. This is microwave radiation. A Rydberg atom in free space then should undergo, in addition to spontaneous emissions, stimulated emissions and absorptions as a result of ambient black-body radiation in the microwave part of the spectrum. We label an adjacent pair of Rydberg states $|e\rangle$ and $|g\rangle$, which typically could correspond to states of principal quantum numbers $n = 50$ and $n = 49$, respectively. If $U_{\rm BB}(\omega)$ is the black-body energy density then the probability rate of stimulated absorption to take the atom from $|g\rangle$ to $|e\rangle$ is given by [4]

$$W_{ge}^{\rm BB} = \frac{\pi d^2}{\varepsilon_0 \hbar^2} \overline{(\mathbf{e}_{\rm a} \cdot \mathbf{e})^2} U_{\rm BB}(\omega_0), \tag{10.7}$$

where $\mathbf{e}_{\rm a}$ and $\mathbf{e}$ are the polarization unit vectors of the atomic dipole and the radiation field, respectively, and the bar represents an average over all possible polarization and propagation directions. One can show that $\overline{(\mathbf{e}_{\rm a} \cdot \mathbf{e})^2} = 1/3$. Further, for a thermal field in free space, we can write the energy density as $U_{\rm BB}(\omega_0) = \rho_{\rm fs}(\omega_0)\bar{n}(\omega_0)$, where $\rho_{\rm fs}(\omega_0)$ is the mode density of free space, $\rho_{\rm fs}(\omega_0) = \omega_0^2/(\pi^2 c^3)$ (see Eq. (2.75)), and

$$\bar{n}(\omega_0) = [\exp(\hbar\omega_0/k_{\rm B}T) - 1]^{-1}. \tag{10.8}$$

Thus we can write

$$W_{ge}^{\rm BB} = B\,\bar{n}, \tag{10.9}$$

where

$$B = \frac{d^2\omega_0^3}{3\pi\varepsilon_0\hbar c^3}. \tag{10.10}$$

The probability rate of stimulated emission from the $|e\rangle$ to $|g\rangle$ transition is given by

$$W_{eg}^{\rm BB} = B\,\bar{n}\,(\omega_0) = W_{ge}^{\rm BB}. \tag{10.11}$$

Spontaneous emission from state $|e\rangle$ can be taken into account by writing

$$W_{eg}^{\mathrm{BB+SpE}} = B[\bar{n}(\omega_0) + 1]. \tag{10.12}$$

From Eq. (10.4) we notice that $B = \Gamma$. From this equivalence we can write

$$\Gamma = \frac{\pi d^2}{3\varepsilon_0 \hbar^2} \rho_{\mathrm{fs}}(\omega_0)\, \hbar\omega_0. \tag{10.13}$$

If $\bar{n} = 0$, the rate of spontaneous emission appears to be related to the vacuum fluctuations of all the modes around the atom. For a Rydberg transition with wavelengths in the millimeter range, $\hbar\omega_0 \ll k_{\mathrm{B}}T$ even at room temperatures, so we may use the Rayleigh–Jeans limit of the Planck radiation law for $\bar{n} \approx k_{\mathrm{B}}T/\hbar\omega_0 \gg 1$ to obtain

$$W_{eg}^{\mathrm{BB}} \approx \frac{d^2\omega_0^2 k_{\mathrm{B}}T}{3\pi\varepsilon_0\hbar^2c^3}. \tag{10.14}$$

Because for circular Rydberg states $d^2 \sim n^4$ and $\omega_0^2 \sim n^{-6}$, we have $W_{eg}^{\mathrm{BB}} \sim n^{-2}$. This needs to be compared with the *total* spontaneous emission rate of the state $|e\rangle$ which is $\Gamma \sim n^{-5}$. Because $W_{eg}^{\mathrm{BB}}/\Gamma \sim n^3$, the stimulated emission rate due to the black-body radiation surrounding the atom is much greater than the rate of spontaneous emission.

The rates W_{eg}^{BB} and Γ are essentially irreversible rates: the atom radiates a photon into a continuum of modes, the photon never to be reabsorbed by the atom. Our goal has been to obtain the conditions under which the atom–field dynamics will be reversible. We can do this by placing the atom in a cavity of appropriate dimensions such that the density of modes is much, much, less than in free space. Ideally we would wish to have a cavity constructed to support just one mode with a frequency close to the resonance with the atom transition frequency ω_0. This will reduce the rate of irreversible *spontaneous* emissions as only one cavity mode is available into which the atom can radiate. At the same time, we wish to suppress the transitions due to the presence of black-body radiation of wavelength corresponding to the atomic transition frequency. For a cavity with walls of low enough temperature such that $\hbar\omega_0 \gg k_{\mathrm{B}}T$, we have $\bar{n} \approx \exp(-\hbar\omega_0/k_{\mathrm{B}}T) \ll 1$, and thus the stimulated transition rate due to thermal photons will be small.

As a simple example of an atom in such a cavity, we consider that case of an atom in the excited state $|e\rangle$ placed in a cavity supporting a single-mode field at resonance with the atomic transition frequency ω_0. We assume that the field is initially in the vacuum so that our initial state of the atom–field system is $|1\rangle = |e\rangle|0\rangle$. When the atom undergoes a spontaneous emission the system goes to the state $|2\rangle = |g\rangle|1\rangle$. These atom–field states have identical total energies: $E_1 = E_2 = \hbar\omega_0/2$ (at resonance). If there are other interactions involved, the system would simply oscillate periodically between the states $|1\rangle$ and $|2\rangle$ at the vacuum Rabi frequency $\Omega_0 = \Omega(0) = 2\lambda$ (see Section 4.5). But there is one other possibility: the photon could leak out of the cavity before the atom is able to reabsorb it. No matter how low the temperature of the cavity walls, they interact

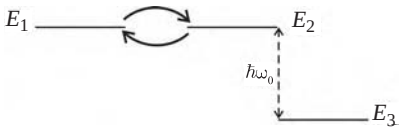


Fig. 10.2. Energy level diagram for the atom–field system. The states $|1\rangle = |e\rangle|0\rangle$ and $|2\rangle = |g\rangle|1\rangle$ have the same energy and without the loss of a photon the system oscillates between these two states. But if the photon is lost, the system state becomes $|3\rangle = |g\rangle|0\rangle$ with energy lower by the amount $\hbar\omega_0$.

with the atom–field system and have two important effects: the energy field decays as the result of the loss of photons, and there is a decoherence of the quantum dynamics of the system. So there is a third state that we need to consider in our description of the interaction of an excited atom with a single-mode cavity field: the state $|3\rangle = |g\rangle|0\rangle$ is decoupled from the reversible dynamics but appears as the result of the irreversible dynamics associated with losses. State $|3\rangle$ has the energy $E_3 = -\hbar\omega_0/2$, lower than E_1 and E_2 by the amount $\hbar\omega_0$ as indicated in Fig. 10.2. This energy goes into the excitation of the reservoir. The issue of incoherent losses is the subject of Chapter 8. Here, as there, we will introduce the appropriate master equation to model the losses.

If we let the interaction Hamiltonian be represented by

$$\hat{H}_{\rm I} = \hbar\lambda(\hat{a}\,\hat{\sigma}_+ + \hat{a}^\dagger\,\hat{\sigma}_-) \tag{10.15}$$

as we did in Chapter 3, then we can write the master equation for the evolution of the density operator, derived and discussed in Chapter 8, as

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}_{\rm I}, \hat{\rho}] - \frac{\kappa}{2}(\hat{a}^\dagger\hat{a}\hat{\rho} + \hat{\rho}\hat{a}^\dagger\hat{a}) + \kappa\hat{a}\hat{\rho}\hat{a}^\dagger, \tag{10.16}$$

where we shall take $\kappa = \omega_0/Q$ and where Q is a characteristic quality factor parameter of the cavity describing the rate of loss. The time $T_{\rm r} = 1/\kappa$ is the cavity relaxation time. For Q very large, the rate of losses will be (desirably) small. With the states $|i\rangle$, $i = 1, 2, 3$, such that $\rho_{ij} = \langle i|\hat{\rho}|j\rangle$, Eq. (10.16) is equivalent to the set of coupled first-order differential equations

$$\begin{aligned}
\frac{d\rho_{11}}{dt} &= \frac{i}{2}\Omega_0\,(\rho_{12} - \rho_{21}) \equiv \frac{i}{2}\Omega_0 V,\\
\frac{d\rho_{22}}{dt} &= -\frac{\omega_0}{Q}\rho_{22} - \frac{i}{2}\Omega_0 V,\\
\frac{dV}{dt} &= i\Omega_0 W - \frac{1}{2}\left(\frac{\omega_0}{Q}\right)V,\\
\frac{d\rho_{33}}{dt} &= \frac{\omega_0}{Q}\rho_{22},
\end{aligned} \tag{10.17}$$

where we have set $V = \rho_{12} - \rho_{21}$ and $W = \rho_{11} - \rho_{22}$. The first three of these equations are precisely of the form of the Bloch equations familiar from magnetic resonance [5] and of optical resonance with a classical

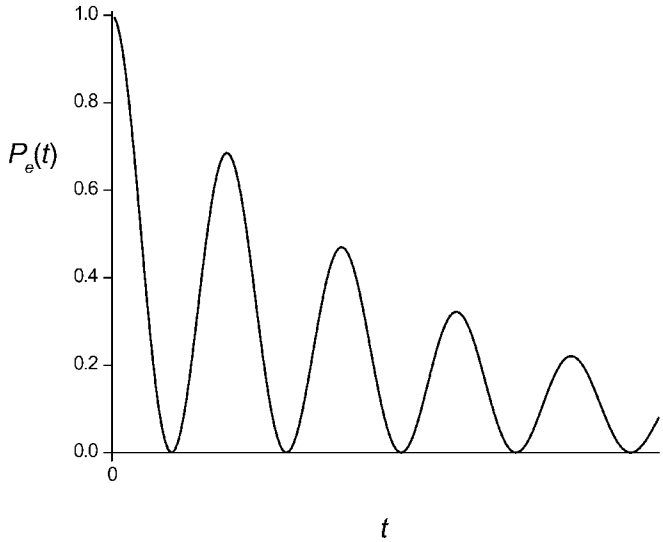


Fig. 10.3. Damped excited state probability for an atom in a high-Q cavity, where many vacuum Rabi oscillations are visible within a cavity damping time.

field [6]. We refer to Eqs. (10.17) as "one-photon" Bloch equations although Ω_0 is the vacuum Rabi frequency. In matrix form, the first three of these read

$$\frac{d}{dt}\begin{pmatrix}\rho_{11}\\ \rho_{22}\\ V\end{pmatrix} = \begin{pmatrix}0 & 0 & i\Omega_0/2\\ 0 & -\omega_0/Q & -i\Omega_0/2\\ i\Omega_0 & -i\Omega_0 & -\omega_0/2Q\end{pmatrix}\begin{pmatrix}\rho_{11}\\ \rho_{22}\\ V\end{pmatrix}. \tag{10.18}$$

The initial conditions are $\rho_{11}(0) = 1$, $\rho_{22}(0) = 0$, and $\rho_{12}(0) = 0$. Of course, we also have $\rho_{33}(0) = 0$. The eigenvalue problem of the matrix in Eq. (10.18) leads to the equation

$$\left(\Lambda + \frac{\omega_0}{2Q}\right)\left(\Lambda^2 + \frac{\omega_0}{Q}\Lambda + \Omega_0^2\right) = 0 \tag{10.19}$$

whose solutions are

$$\begin{aligned}\Lambda_0 &= -\frac{\omega_0}{2Q},\\ \Lambda_\pm &= -\frac{\omega_0}{2Q} \pm \frac{\omega_0}{2Q}\left(1 - \frac{4\Omega_0^2 Q^2}{\omega_0^2}\right)^{1/2}.\end{aligned} \tag{10.20}$$

If the cavity decay rate is weak such that $\omega_0/Q < 2\Omega_0$ then the eigenvalues $\Lambda_\pm$ will be complex and we obtain damped oscillations of frequency Ω_0 in the probability of finding the atom in the excited state $P_e(t) = \rho_{11}(t)$ as given in Fig. 10.3. The oscillations, which here are the vacuum Rabi oscillations, reflect the fact that the spontaneous emission is reversible for weak field decay. The oscillations are damped at the rate $\omega_0/2Q$. But if there is strong cavity damping such that $\omega_0/Q > 2\Omega_0$, the eigenvalues $\Lambda_\pm$ are real and so there will now be irreversible spontaneous emission as indicated in Fig. 10.4. The largest time constant, and

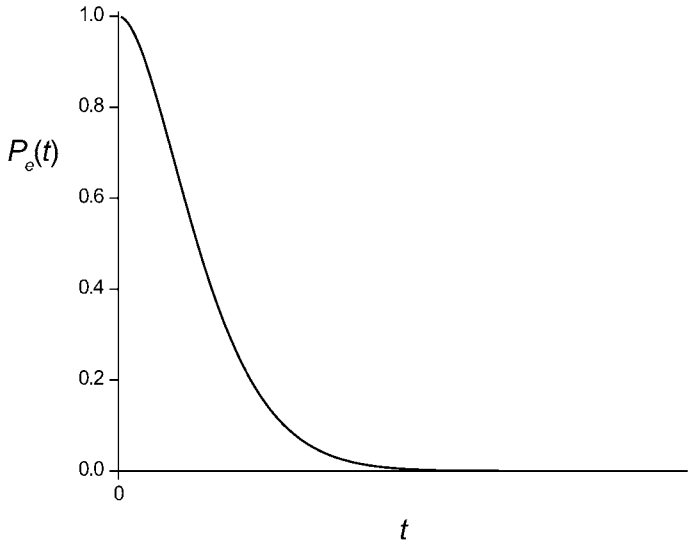


Fig. 10.4. Decay of excited state of an atom in a low-Q cavity, with no visible Rabi oscillations. The atom decays at a rate that depends on the cavity Q.

thus the smallest eigenvalue in this case is Λ_+, which we may approximate as

$$\Lambda_+ \simeq -\frac{\Omega_0^2 Q}{\omega_0} = -\frac{4d^2 Q}{\hbar \varepsilon_0 V}, \tag{10.21}$$

where the right-hand side has been obtained by using the relations $\Omega_0 = 2\lambda$, $\lambda = dg/\hbar$ and where g is given by Eq. (4.87), V being the effective mode volume of the cavity. We have assumed that the atom is sitting near an anti-node of the standing-wave cavity field such that $\sin^2(kz) \approx 1$. Therefore we may define the rate of irreversible decay as

$$\Gamma_{\text{cav}} = \frac{4d^2 Q}{\hbar \varepsilon_0 V}. \tag{10.22}$$

Note that the decay rate increases with increasing Q and thus we have a cavity-enhanced rate of spontaneous emission. If we write the decay rate after the form in Eq. (10.13) as

$$\Gamma_{\text{cav}} = \frac{\pi d^2}{\varepsilon_0 \hbar^2} \rho_{\text{cav}}(\omega_0) \hbar \omega_0 \tag{10.23}$$

then we find that the number of modes per unit volume of a cavity is

$$\rho_{\text{cav}}(\omega) = \frac{4Q}{\pi \varepsilon_0 V} \tag{10.24}$$

and thus we have

$$\frac{\Gamma_{\text{cav}}}{\Gamma} = \frac{3\rho_{\text{cav}}(\omega)}{\rho_{\text{fs}}(\omega)} = \frac{4Q\lambda_0^3}{3\pi \varepsilon_0 V}. \tag{10.25}$$

This result, predicting enhanced spontaneous emission in a cavity, was derived by Purcell [7] in the context of magnetic resonance. Of course, the result assumes that the cavity supports a mode of the proper wavelength into which the atom

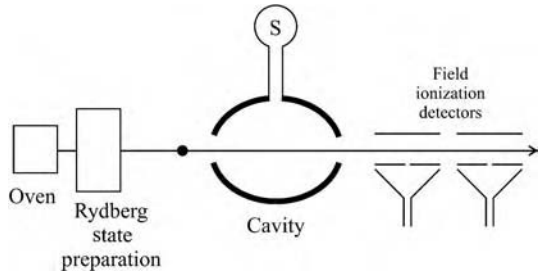


Fig. 10.5. Experimental setup for the realization of the Jaynes–Cummings model. The cavity field is prepared in a coherent state obtained via a waveguide from the source S, which itself is driven by classical currents. The atoms are prepared in the excited state, velocity selected, and then directed into the cavity. The selective ionization detectors yield the statistics for the inversion as a function of interaction time, which itself is controlled by the velocity selection.

can radiate. But if the cavity is so small that the wavelength λ_0 cannot be supported, spontaneous emission will be inhibited [8], and this has been observed experimentally [9].

10.3 Experimental realization of the Jaynes–Cummings model

The experimental realization of cavity-enhanced and reversible transition dynamics have been pioneered by groups at the Max Planck Institute for Quantum Optics and at the Ecole Normale Supérieure.

In the experiments performed by Haroche and collaborators at l'Ecole Normale Supérieure (ENS) in Paris, the cavity used consists of a Fabry–Perot resonator constructed out of superconducting niobium spherical mirrors of diameter 50 mm and radius of curvature 40 mm separated by 27 mm. The measured Q for the cavity is $Q = 3 \times 10^8$ though higher values are not out of the question. The storage time for photons for such a cavity is about 1 ms, shorter than the atomic decay time for a circular Rydberg atom with $n \sim 50$ as estimated above, but much longer that the atom–field interaction time which, at the speeds atoms are injected into the cavity, is on the order of a few microseconds. The cavity walls are cooled to about 1 K, leaving an average of 0.7 microwave photons in the cavity. But this number can be further reduced by injecting a sequence of ground-state atoms through the cavity in order to absorb the photons, effectively "cooling" the cavity further. The procedure reduces the average photon number to 0.1. In Fig. 10.5 we sketch the setup for a typical cavity QED experiment. The microwave cavity, which is open on the sides to allow the passage of atoms, has attached to it a source of classical microwaves (a klystron) whose output can be piped into the cavity to form a coherent state of low amplitude. In addition there are other

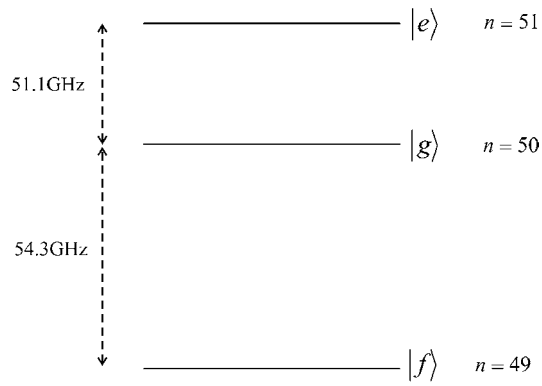


Fig. 10.6. The circular Rydberg states and their corresponding transition frequencies used for the cavity QED experiments performed by the Paris group.

sources of classical resonant microwaves that can be used to manipulate the atomic states before and after the atom leaves the cavity. These fields essentially perform Ramsey interferometry [10], as will be demonstrated below, and are thus referred to as Ramsey fields R_1 and R_2. The fields have very rapid decay (cavities of low Q are used with small photon numbers) and thus may be treated classically as no entanglement is generated [11]. Finally, selective ionization detectors are placed after the second Ramsey field.

As for the circular Rydberg atoms, the states used are those of principal quantum numbers $n = 49, 50,$ and 51 which we denote respectively as $|f\rangle, |g\rangle,$ and $|e\rangle$. The energy-level diagram for the atomic states is given in Fig. 10.6. The transition frequency between the $|e\rangle \leftrightarrow |g\rangle$ transition is 51.1 GHz and that of the $|g\rangle \leftrightarrow |f\rangle$ transition is 54.3 GHz. The latter transition frequency is far enough off-resonance with the cavity field mode such that there are no transitions between these levels. However, that does not mean that the cavity has no effect, as we shall see in the next section. The atomic states can be tuned in and out of resonance with the cavity field by the application of a small static electric field (a Stark shift) if necessary.

An early experiment designed to observe collapse and revival phenomena in atom–field interactions in CQED was reported in 1987 by Rempe, Walther and Klein [12]. This involved a thermal field in a cavity and, indeed, an oscillation was observed in the atomic-state statistics, but the experiment was not able to manifest a wide enough range of interaction times to explore collapse and revival. In 1996, Haroche and collaborators [13] were able to perform an experiment with the requisite control over the interaction times and, furthermore, were able to inject quantized coherent cavity fields of up to $\bar{n} = 1.77(\pm 0.15)$ photons. In this particular experiment, the cavity used had a relaxation time of $T_r = 220\,\mu s$ corresponding to a $Q = 7 \times 10^7$. The relaxation time is longer than the atom–field interaction times t which range from $40 < t < 90\,\mu s$. The corresponding range of the atomic velocities v is $110 < v < 250\,m/s$. The results of the experiment

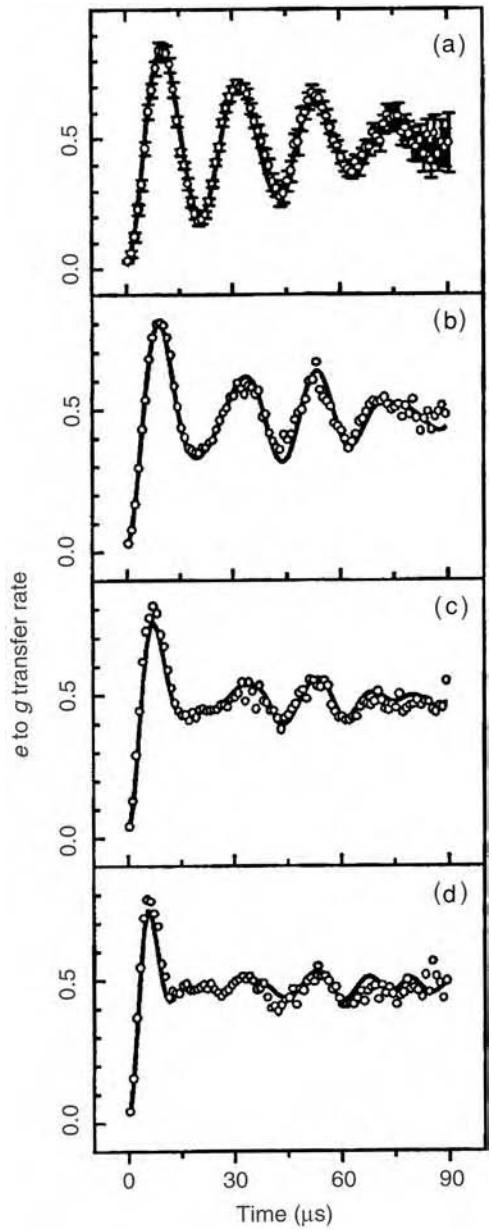


Fig. 10.7. The $|e\rangle$ to $|g\rangle$ transition rates for various initial cavity field states. Part(a) is for no injected field with an average of 0.06($\pm$0.01) thermal photons, whereas (b), (c) and (d) are for injected coherent states with average photon numbers 0.040($\pm$0.02), 0.85($\pm$0.04), and 1.77($\pm$0.15) respectively. (Redrawn from Reference [13].)

are given in Fig. 10.7, reprinted from Reference [13]. The first (a) shows the oscillations obtained with no injected field but with a thermal field containing $\bar{n} = 0.06(\pm 0.01)$ photons. Several oscillations are clearly present. For the coherent fields of increasing amplitude, collapse followed by a revival is evident. The largest coherent field used had an average photon number of $\bar{n} = 1.77(\pm 0.15)$. (Reprinted with permission.)

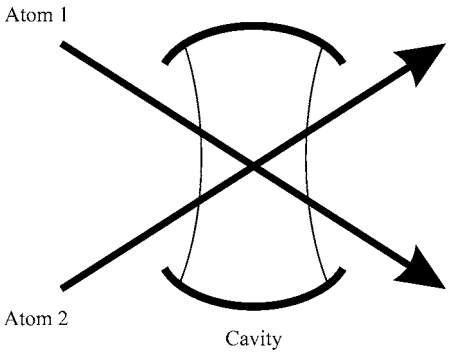


Fig. 10.8. Proposed method for creating an entanglement of two atoms. The first atom, prepared in the excited state, is velocity selected such that it emits a photon upon passing through the cavity. Subsequently, the second atom, prepared in the ground state, is velocity selected to absorb the photon. As a result, both atoms, though distant, become entangled.

10.4 Creating entangled atoms in CQED

The experimental arrangements used for the CQED experiments described above, with some modification, can be used to generate entanglement between successive *atoms* injected through the cavity. In Fig. 10.8 we show a setup with two atoms crossing the cavity, where atom 1 is prepared in the excited state with atom 2 initially in the ground state. The cavity field is assumed to be initially in the vacuum. While the atom is in the cavity, the field and atom evolve according to

$$|e\rangle_1 |0\rangle_{\text{cav}} \rightarrow \cos(\lambda t_1) |e\rangle_1 |0\rangle - i \sin(\lambda t_1) |g\rangle_1 |1\rangle . \tag{10.26}$$

If the speed of the atom is chosen such that $\lambda t_1 = \pi/4$ then we have

$$|e\rangle_1 |0\rangle_{\text{cav}} \rightarrow \frac{1}{\sqrt{2}} (|e\rangle_1 |0\rangle - i |g\rangle_1 |1\rangle) . \tag{10.27}$$

Sending in the second atom we have

$$|e\rangle_1 |g\rangle_2 |0\rangle_{\text{cav}} \rightarrow \frac{1}{\sqrt{2}} (|e\rangle_1 |g\rangle_2 |0\rangle - i |g\rangle_1 |g\rangle_2 |1\rangle) , \tag{10.28}$$

and using that

$$|g\rangle_2 |1\rangle \rightarrow \cos(\lambda t_2) |g\rangle_2 |1\rangle - i \sin(\lambda t_2) |e\rangle_2 |0\rangle, \tag{10.29}$$

and with the velocity of the second atom chosen so that $\lambda t_2 = \pi/2$, we obtain finally

$$|e\rangle_1 |g\rangle_2 |0\rangle_{\text{cav}} \rightarrow \left|\Psi^+\right\rangle |0\rangle , \tag{10.30}$$

where

$$\left|\Psi^+\right\rangle = \frac{1}{\sqrt{2}} (|e\rangle_1 |g\rangle_2 + |g\rangle_1 |e\rangle_2) \tag{10.31}$$

is just one of the maximally correlated Bell states, this time given in terms of the states of a two-level atom. Using classical resonant fields and selective ionization detection, one can use such a state to test local realist theories by seeking violations of Bell's inequalities. However, as the atom state measurements are not likely to have space-like separations, this type of test retains a loophole in regard to the criterion of locality. The generation of entangled atoms via the method discussed here was proposed by Cirac and Zoller [14]. Experimental realization was achieved by the Paris group [15].

10.5 Formation of Schrödinger cat states with dispersive atom–field interactions and decoherence from the quantum to the classical

In Section 4.8 we discussed the Jaynes–Cummings model in the case of large detuning. The interaction in that case becomes, as shown in Appendix C, represented by the effective Hamiltonian

$$\hat{H}_{\text{eff}} = \hbar\chi[\hat{\sigma}_+\hat{\sigma}_- + \hat{a}^\dagger\hat{a}\,\hat{\sigma}_3] \tag{10.32}$$

where $\hat{\sigma}_+\hat{\sigma}_- = |e\rangle\langle g \mid g\rangle\langle e| = |e\rangle\langle e|$ and $\hat{\sigma}_3 = |e\rangle\langle e| - |g\rangle\langle g|$. The large detuning is between the cavity-field frequency and the frequency for the $|e\rangle \rightarrow |g\rangle$ transition. The lower-level state $|f\rangle$ is, of course, still dynamically disconnected as it is very far from resonance with the cavity field. If the atomic state $|e\rangle$ is not to become populated, the effective interaction Hamiltonian can be written as

$$\hat{H}_{\text{eff}} = -\hbar\chi\hat{a}^\dagger\hat{a}\,|g\rangle\,\langle g|\,. \tag{10.33}$$

Let us suppose now that the setup pictured in Fig. 10.9 is used where the Ramsey zone R_1, whose field we now take to be resonant with the $|f\rangle \leftrightarrow |g\rangle$ transition (as is also assumed to be the case for R_2), is used to prepare an atom in a superposition of the form

$$|\psi_{\text{atom}}\rangle = \frac{1}{\sqrt{2}}\,(|g\rangle + |f\rangle)\,. \tag{10.34}$$

The atom now enters the cavity, which itself has been prepared in a coherent state $|\alpha\rangle$. Thus the initial state of the atom–field system is

$$|\Psi\,(0)\rangle = |\psi_{\text{atom}}\rangle\,|\alpha\rangle = \frac{1}{\sqrt{2}}\,(|g\rangle + |f\rangle)\,|\alpha\rangle\,. \tag{10.35}$$

Using the effective Hamiltonian given by 10.33, this state evolves to

$$|\Psi\,(t)\rangle = \exp[-i\hat{H}_{\text{eff}}t/h]|\Psi(0)\rangle = \frac{1}{\sqrt{2}}[|g\rangle|\alpha e^{i\chi t}\rangle + |f\rangle|\alpha\rangle]. \tag{10.36}$$

We now suppose that in the second Ramsey zone, R_2, a classical resonant field performs the transformations

$$\begin{aligned} |g\rangle &\rightarrow \cos(\theta/2)|g\rangle + \sin(\theta/2)|f\rangle, \\ |f\rangle &\rightarrow \cos(\theta/2)|f\rangle - \sin(\theta/2)|g\rangle, \end{aligned} \tag{10.37}$$

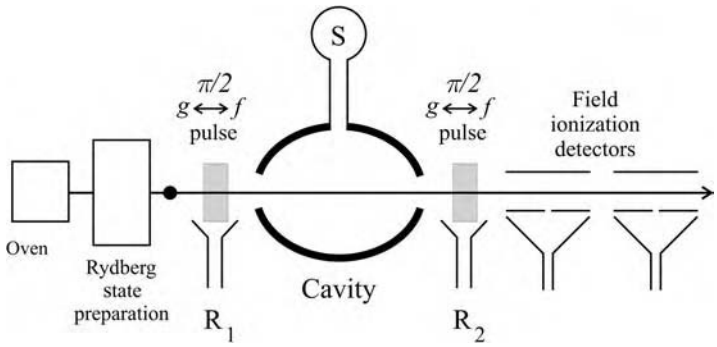


Fig. 10.9. A sketch of the experiment designed to generate the Schrödinger cat states, as discussed in the text. R_1 and R_2 are the Ramsey zones that effect $\pi/2$ pulses on the $|g\rangle \leftrightarrow |f\rangle$ transition.

where the Rabi angle $\theta = \Omega_R T$ and where Ω_R is the semiclassical Rabi frequency and T is the interaction time. Equation (10.36) is transformed into

$$\begin{aligned} |\Psi(t),\theta\rangle &= \frac{1}{\sqrt{2}}[(\cos(\theta/2)|g\rangle + \sin(\theta/2)|f\rangle)|\alpha e^{i\chi t}\rangle \\ &\quad + (\cos(\theta/2)|f\rangle - \sin(\theta/2)|g\rangle)|\alpha\rangle], \\ &= |g\rangle\frac{1}{\sqrt{2}}[\cos(\theta/2)|\alpha e^{i\chi t}\rangle - \sin(\theta/2)|\alpha\rangle] \\ &\quad + |f\rangle\frac{1}{\sqrt{2}}[\sin(\theta/2)|\alpha e^{i\chi t}\rangle + \cos(\theta/2)|\alpha\rangle]. \end{aligned} \tag{10.38}$$

With the interaction time chosen such that $\chi t = \pi$, and with the Rabi angle θ chosen as $\theta = \pi/2$, we have

$$|\Psi(\pi/\chi),\pi/2\rangle = \frac{1}{2}[|g\rangle(|-\alpha\rangle - |\alpha\rangle) + |f\rangle(|-\alpha\rangle + |\alpha\rangle)]. \tag{10.39}$$

With the ionization detectors set to detect the states $|g\rangle$ and $|f\rangle$, a detection of $|f\rangle$ projects the cavity field into an even Schrödinger cat state

$$|\psi_e\rangle = \mathcal{N}_e(|\alpha\rangle + |-\alpha\rangle) \tag{10.40}$$

while if $|g\rangle$ is detected the cavity field is projected into the odd cat state

$$|\psi_o\rangle = \mathcal{N}_o(|\alpha\rangle - |-\alpha\rangle), \tag{10.41}$$

states first given in Eqs. (7.115)–(7.117). The atomic velocity required to achieve the requisite phase shift of $\chi t = \pi$ for generating the even and odd cat states is about 100 m/s.

The generation of the even and odd cat states, as just discussed, can – in fact, must – be seen as a consequence of interference. We illustrate this in Fig. 10.10.The Ramsey zones act as beam splitters, the second effectively creating ambiguity in the choice of two "paths".

One of the motivations for creating the Schrödinger cats states was to provide an opportunity to attempt monitoring the decoherence of a mesoscopic superposition state into a statistical mixture. If we assume that the atom–field interaction time is short enough so that dissipative effects can be ignored while the atom is

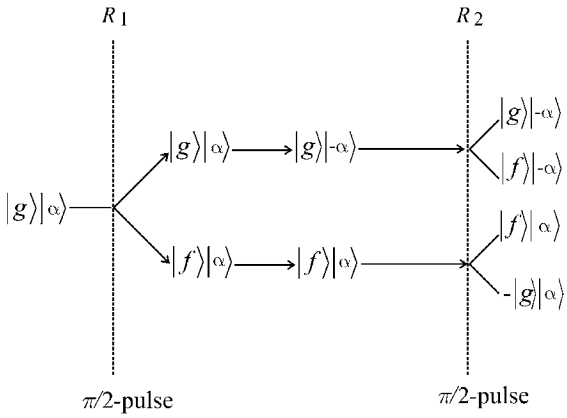


Fig. 10.10. A graphical representation of the interference involved in the generation of the Schrödinger cat states. The Ramsey zones act like beam splitters to create "path" ambiguity. On the far right, the field states associated with like atomic states become superposed upon selective atomic state ionization.

in the cavity, then after the atom exits and is ionized we can take the *field* state, now unentangled with the atom, to evolve according to the master equation

$$\frac{d\hat{\rho}_{\rm F}}{dt} = -\frac{\kappa}{2}(\hat{a}^\dagger\hat{a}\hat{\rho}_{\rm F} + \hat{\rho}_{\rm F}\hat{a}^\dagger\hat{a}) + \kappa\hat{a}\hat{\rho}_{\rm F}\hat{a}^\dagger, \tag{10.42}$$

where we have $\hat{H}_{\rm I} = 0$ as the atom and the field are no longer interacting and where we have made the notational replacement $\gamma \to \kappa$. As we showed in Chapter 8, the solution to the master equation for the field initially in an even cat state is

$$\hat{\rho}_{\rm F}(t) = \mathcal{N}_{\rm e}^2\Big\{|\alpha\, e^{-\kappa t/2}\rangle\langle\alpha\, e^{-\kappa t/2}| + |-\alpha\, e^{-\kappa t/2}\rangle\langle-\alpha\, e^{-\kappa t/2}| + e^{-2|\alpha|^2(1-e^{-\kappa t})}[|\alpha\, e^{-\kappa t/2}\rangle\langle-\alpha\, e^{-\kappa t/2}| + |-\alpha\, e^{-\kappa t/2}\rangle\langle\alpha\, e^{-\kappa t/2}|]\Big\}. \tag{10.43}$$

Notice, again, that as $t \to \infty$, the last two terms, the "coherences", decay faster than energy itself decays provided that $|\alpha|^2$ is not too small. Under this condition we find that the initial cat state, a quantum mechanical superposition state, decoheres into a statistical mixture of coherent states

$$\hat{\rho}_{\rm F}(t) \xrightarrow[t \text{ large}]{} \frac{1}{2}[|\beta(t)\rangle\langle\beta(t)| + |-\beta(t)\rangle\langle-\beta(t)|] \tag{10.44}$$

where $\beta(t) = \alpha\, e^{-\kappa t/2}$. Of course, as $t \to \infty$, we obtain the (pure) vacuum state: $\hat{\rho}_{\rm F}(t) \underset{t\to\infty}{\longrightarrow} |0\rangle\langle 0|$.

How can this decoherence be monitored? Davidovich *et al.* [16] proposed the following method: a second atom (2) prepared in the state $(|g\rangle_2 + |f\rangle_2)/\sqrt{2}$ is injected into the cavity a time T after the first atom and acts as a quantum "mouse" to detect the presence of the cat. The new "initial" state, i.e. just before the second atom is injected is, for the atom–field system,

$$\hat{\rho}_{\rm AF}(T) = \frac{1}{2}(|g\rangle_2 + |f\rangle_2)({}_2\langle g| + {}_2\langle f|)\hat{\rho}_{\rm F}(T) \tag{10.45}$$

where $\hat{\rho}_{\rm F}(T)$ is given by Eq. (10.43). Assuming the velocity of atom 2 is adjusted so that once again we have $\chi t_2 = \pi$, where t_2 is the interaction time of the second

atom with the cavity field, then the atom–field state just after atom 2 exits the cavity is given by the density operator

$$\hat{\rho}_{\rm AF} = \frac{1}{2}[|g\rangle_{22}\langle g|\, e^{i\pi\hat{a}^\dagger\hat{a}}\hat{\rho}_{\rm F}(T)\, e^{-i\pi\hat{a}^\dagger\hat{a}} + |f\rangle_{22}\langle f|\,\hat{\rho}_{\rm F}(T) + |g\rangle_{22}\langle f|\, e^{i\pi\hat{a}^\dagger\hat{a}}\hat{\rho}_{\rm F}(T) + |f\rangle_{22}\langle g|\,\hat{\rho}_{\rm F}(T)\, e^{-i\pi\hat{a}^\dagger\hat{a}}]. \tag{10.46}$$

The second Ramsey zone performs a $\pi/2$-pulse on the atom so that just before the atom encounters the ionization detectors the density operator takes the form

$$\begin{aligned}\hat{\rho}_{\rm AF}(T) = \frac{1}{4}[&(|g\rangle_2 + |f\rangle_2)(_2\langle g| + {}_2\langle f|)e^{i\pi\hat{a}^\dagger\hat{a}}\hat{\rho}_{\rm F}(T)e^{-i\pi\hat{a}^\dagger\hat{a}} \\ &+(-|g\rangle_2 + |f\rangle_2)(-{}_2\langle g| + {}_2\langle f|)\hat{\rho}_{\rm F}(T) \\ &+(|g\rangle_2 + |f\rangle_2)(_2 - \langle g| + {}_2\langle f|)e^{i\pi\hat{a}^\dagger\hat{a}}\hat{\rho}_{\rm F}(T) \\ &+(-|g\rangle_2 + |f\rangle_2)(_2\langle g| + {}_2\langle f|)\hat{\rho}_{\rm F}(T)e^{-i\pi\hat{a}^\dagger\hat{a}}].\end{aligned} \tag{10.47}$$

The reduced density operator for the atom is found by tracing over the field according to

$$\hat{\rho}_{\rm A}(T) = {\rm Tr}_{\rm F}\hat{\rho}_{\rm AF}(T), \tag{10.48}$$

from which we find (the details are left as an exercise) that the probability of detecting the state $|g\rangle_2$ is

$$P_{g_2}(T) = {}_2\langle g|\hat{\rho}_{\rm A}(T)|g\rangle_2 = \frac{1}{2}\left[1 - \frac{e^{-2|\alpha|^2e^{-\kappa T}} + e^{-2|\alpha|^2\left(1-e^{-\kappa T}\right)}}{1+e^{-2|\alpha|^2}}\right], \tag{10.49}$$

and the probability of detecting $|f\rangle_2$ is

$$P_{f_2}(T) = {}_2\langle f|\hat{\rho}_{\rm A}(T)|f\rangle_2 = \frac{1}{2}\left[1 + \frac{e^{-2|\alpha|^2e^{-\kappa T}} + e^{-2|\alpha|^2\left(1-e^{-\kappa T}\right)}}{1+e^{-2|\alpha|^2}}\right], \tag{10.50}$$

where we must have $P_{f_2}(T) + P_{g_2}(T) = 1$. Recall that the initial state of the field, the even cat state, was projected by the detection of the *first* atom to be in the state $|f\rangle$. Well, for $T = 0$ we have that $P_{f_2}(0) = 1$, meaning that if there has not yet been any dissipation of the cavity field, the second atom must also be detected in the state $|f\rangle$. But for increasing T, the probability of detecting the atom in the state $|f\rangle$ starts to slide toward $1/2$. Thus by increasing the delay time between the first and second atoms, one can monitor the decoherence of the initial even cat state into a statistical mixture.

In the preceding, the Ramsey zones were taken to be resonant with the $|g\rangle \leftrightarrow |f\rangle$ but where the state $|f\rangle$ was too far from resonance with the cavity field to give rise to any shift or the phase of the field. But suppose now that the Ramsey zones are tuned resonant with the $|e\rangle \leftrightarrow |g\rangle$ transition and that the first zone produces the superposition

$$|\psi_{\rm atom}\rangle = \frac{1}{\sqrt{2}}(|e\rangle + |g\rangle). \tag{10.51}$$

In this case, we must use the full effective interaction Hamiltonian of Eq. (10.32). With the same initial field state we will have at time t the state

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}[e^{-i\chi t}|e\rangle|\alpha\, e^{-i\chi t}\rangle + |g\rangle|\alpha\, e^{i\chi t}\rangle] \tag{10.52}$$

where we notice that the coherent state components are counter-rotating in phase space. Maximal entanglement will occur at the maximal phase-space separation, which occurs when $\chi t = \pi$, giving

$$|\Psi(\pi/\chi)\rangle = \frac{1}{\sqrt{2}}[-i|e\rangle|-i\,\alpha\rangle + |g\rangle|i\,\alpha\rangle]. \tag{10.53}$$

With the second Ramsey zone effecting the transformations

$$\begin{aligned} |e\rangle &\rightarrow \frac{1}{\sqrt{2}}(|e\rangle + |g\rangle), \\ |g\rangle &\rightarrow \frac{1}{\sqrt{2}}(|g\rangle - |e\rangle), \end{aligned} \tag{10.54}$$

we obtain for arbitrary interaction time t

$$\begin{aligned} |\Psi(t), \pi/2\rangle = \frac{1}{2}[&|e\rangle(e^{-i\chi t}|e^{-i\chi t}\alpha\rangle - |e^{i\chi t}\alpha\rangle) \\ &+ |g\rangle(e^{-i\chi t}|e^{-i\chi t}\alpha\rangle + |e^{i\chi t}\alpha\rangle)]. \end{aligned} \tag{10.55}$$

With the detectors set to distinguish between $|e\rangle$ and $|\mathrm{g}\rangle$, state reduction produces the cavity-field states

$$|\psi_{\pm}\rangle \sim e^{-i\chi t}|\alpha\, e^{-i\chi t}\rangle \pm |\alpha\, e^{i\chi t}\rangle. \tag{10.56}$$

States of exactly this type have, in fact, been generated experimentally and their decoherence has been observed [17]. However, the atomic velocities were rather high in this experiment so that phase shifts of $\chi t = \pi$ were not achieved. In fact, the greatest phase shift achieved was with $\chi t < \pi/4$. Nevertheless, this, along with the fact that for the initial coherent state injected into the cavity had an average photon number $\bar{n} \simeq 10$, afforded a large-enough initial separation of the component coherent states so that observation of the progressive decoherence of the field state was possible.

10.6 Quantum nondemolition measurement of photon number

In the standard approach to measurement as discussed in Appendix B, the quantity being measured is typically destroyed as a result of the measurement. For example, in the context of quantum optics, if one wishes to measure the photon number of a quantized field state, the photo-detectors used will absorb photons thus removing them from the field state. Here we show, in the context of CQED, how photon number can be measured without the absorption of any photons. This will provide an example of what is known as a quantum nondemolition (QND) measurement [18].

Suppose the cavity field contains a definite, but unknown, number of photons n. A Rydberg atom is then prepared in a superposition state $(|e\rangle + |g\rangle)/\sqrt{2}$ and injected into the cavity. The initial state is thus

$$|\Psi(0)\rangle = \frac{1}{\sqrt{2}}(|e\rangle + |g\rangle)|n\rangle = \frac{1}{\sqrt{2}}(|e\rangle|n\rangle + |g\rangle|n\rangle). \tag{10.57}$$

After a time t, which is the time the atom exits the cavity, the state vector is

$$\begin{aligned}|\Psi(t)\rangle &= \exp[-i\hat{H}_{\text{eff}}t/\hbar]|\Psi(0)\rangle \\ &= \frac{1}{\sqrt{2}}[e^{-i\chi t(n+1)}|e\rangle + e^{i\chi t n}|g\rangle]|n\rangle.\end{aligned} \tag{10.58}$$

If a Ramsey field subjects the atom to a resonant $\pi/2$ pulse, thus effecting the transformations of Eq. (10.54), then the state vector becomes

$$|\Psi(t), \pi/2\rangle = \frac{1}{2}\Big[|e\rangle\Big(e^{-i\chi t(n+1)} - e^{+i\chi t n}\Big) + |g\rangle\Big(e^{-i\chi t(n+1)} + e^{+i\chi t n}\Big)|n\Big]. \tag{10.59}$$

The probabilities of detecting the atom in the excited and ground states respectively are then

$$\begin{aligned}P_e(t) &= \frac{1}{2}\{1 + \cos[(2n+1)t]\}, \\ P_g(t) &= \frac{1}{2}\{1 - \cos[(2n+1)t]\},\end{aligned} \tag{10.60}$$

which clearly exhibit oscillations (Ramsey fringes) with time. The frequency of the oscillations yields the photon number n. Note that there will be oscillations even when $n = 0$. These Ramsey fringes have been observed by Brune *et al.* [19] for the cases $n = 0$ and $n = 1$. The cavity is prepared in a single photon state by injecting into the cavity an excited atom of appropriate speed with the cavity tuned into resonance.

10.7 Realization of the Jaynes–Cummings interaction in the motion of a trapped ion

We conclude this chapter with a discussion of what amounts to a realization of the Jaynes–Cummings interaction but where the quantized cavity field is replaced by the vibrational motion of the center-of-mass of a trapped ion. In other words, the *photons* are replaced by *phonons*, the quanta of mechanical vibrations. The realization of such an interaction is the result of significant advances in the laser cooling of ions (and neutral atoms) over the past decade. We do not discuss methods of laser cooling here and we refer the reader to the articles listed in the bibliography.

We assume that an ion of mass M has been laser cooled and is contained in some form of electromagnetic trap. Two types of traps are in general use. One is the Penning trap which uses a combination of a uniform magnetic and quadrupole electric fields. The other is the Paul trap, also known as the r.f. trap, which uses radio-frequency electric fields to trap ions. It is, of course, not possible to trap ions

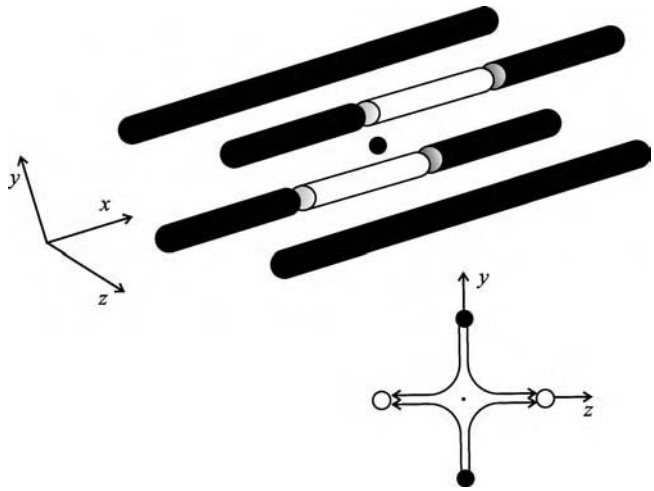


Fig. 10.11. Sketch of a linear r.f. Paul trap. The *x*-axis is taken to be along the axis of the trap and the black dot represents a single trapped ion. The smaller figure represents an end view. Then a common r.f. field is applied to the dark electrodes while the others are held at constant electric potentials.

with static electric fields on account of Earnshaw's theorem [20]. We shall not go into the details of the traps (but see Horvath *et al.* [21]). In a typical experiment of the type we have in mind involving a single trapped ion, it is the Paul trap that is used. A sketch of such a device is given in Fig.10.11.

We take as our starting point a two-level ion, with internal electronic energy levels, usually hyperfine levels, denoted $|e\rangle$ and $|\mathrm{g}\rangle$, contained within the trap. A laser of frequency ω_{L}, is directed along the axis, denoted the x-axis, of the trap. This laser field is assumed to be tunable. The trap itself is approximated as a harmonic oscillator of frequency ν. The value of this frequency is determined by the various trap parameters. The position of the center-of-mass of the ion is given by the operator $\hat{x}$, which we shall write below in terms of annihilation and creation operators, respectively $\hat{a}$, $\hat{a}^\dagger$, which are now taken to be operators that annihilate and create quanta of the vibration motion of the center-of-mass of the ion, not those of a quantized field. (There are *no* quantized radiation fields involved with this problem.) The Hamiltonian for our system is given by

$$\hat{H} = \hat{H}_0 + [\mathcal{D}E^{(-)}(\hat{x}, t)\,\hat{\sigma}_- + H.c.], \tag{10.61}$$

where

$$\hat{H}_0 = \frac{\hbar\omega_0}{2}\hat{\sigma}_3 + \hbar\nu\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{10.62}$$

is the interaction-free Hamiltonian and where $\mathcal{D}$ is the dipole moment of the $|g\rangle \leftrightarrow |e\rangle$ transition, and $E^{(-)}(\hat{x}, t)$ is the negative frequency part of the laser field given by

$$E^{(-)}(\hat{x}, t) = E_0 \exp[i(\omega_{\mathrm{L}}t - k_{\mathrm{L}}\hat{x} + \phi)], \tag{10.63}$$

E_0 is the amplitude of the laser field, $k_{\rm L} = 2\pi/\lambda_{\rm L}$ the wave vector of the field, and ϕ is just some phase. We write the position operator in the usual way as

$$\hat{x} = \sqrt{\frac{\hbar}{2\nu M}}\left(\hat{a} + \hat{a}^\dagger\right). \tag{10.64}$$

Substitution into Eq. (10.63) yields

$$E^{(-)}(\hat{x}, t) = E_0 e^{i(\phi+\omega_{\rm L}t)} e^{-i\,\eta\left(\hat{a}+\hat{a}^\dagger\right)}, \tag{10.65}$$

where $\eta \equiv k_{\rm L}(\hbar/2\nu M)^{1/2}$ is the so-called Lamb–Dicke parameter. This parameter is typically small, $\eta \ll 1$, this range being known as the Lamb–Dicke regime. In this regime, the vibrational amplitude of the ion is much smaller than the wavelength of the laser light field. With $\hat{U} = \exp(-i\,\hat{H}_0 t/\hbar)$ we now transform to the interaction picture obtaining the interaction Hamiltonian

$$\begin{aligned}\hat{H}_{\rm I} &= \hat{U}^\dagger \hat{H} \hat{U} + i\,\hbar \tfrac{d\,\hat{U}^\dagger}{d\,t}\hat{U},\\ &= \mathcal{D}E_0 e^{i\,\phi} e^{i\,\omega_{\rm L} t}\exp[-i\,\eta(\hat{a}\,e^{i\,\nu t} + \hat{a}^\dagger\,e^{-i\,\nu t})]\hat{\sigma}_- e^{-i\,\omega_0 t} + H.c.\end{aligned} \tag{10.66}$$

as the reader can verify. As η is small, we expand to first order

$$\exp[-i\,\eta(\hat{a}\,e^{i\,\nu t} + \hat{a}^\dagger\,e^{-i\,\nu t})] \approx 1 - i\,\eta(\hat{a}\,e^{i\,\nu t} + \hat{a}^\dagger\,e^{-i\,\nu t}), \tag{10.67}$$

and thus we have to this order

$$\begin{aligned}\hat{H}_{\rm I} \approx \mathcal{D}E_0 e^{i\,\phi}[e^{i(\omega_{\rm L}-\omega_0)\,t} - i\,\eta(\hat{a}\,e^{i(\omega_L-\omega_0+\nu)\,t} + \hat{a}^\dagger\,e^{i(\omega_L-\omega_0-\nu)\,t})]\hat{\sigma}_-\\ + H.c.\end{aligned} \tag{10.68}$$

Suppose now that the laser is tuned such that $\omega_{\rm L} = \omega_0 + \nu$. We shall then have

$$\hat{H}_{\rm I} \approx \mathcal{D}E_0 e^{i\,\phi}[e^{i\nu\,t} - i\,\eta(\hat{a}\,e^{i2\nu\,t} + \hat{a}^\dagger)]\hat{\sigma}_- + H.c. \tag{10.69}$$

The terms rotating at frequencies ν and 2ν oscillate very rapidly compared with the remaining term and average to zero. Dropping these terms (essentially making a rotating wave approximation) we obtain

$$\hat{H}_{\rm I} \approx -i\,\hbar\eta\,\Omega\,e^{i\,\phi}\hat{a}^\dagger\hat{\sigma}_- + H.c., \tag{10.70}$$

which is the Jaynes–Cummings interaction, where $\Omega = \mathcal{D}E_0/\hbar$ is the Rabi frequency associated with the semiclassical laser–atom interaction. But with regard to the Jaynes–Cummings interaction, it is important to note that the internal states of the ion are coupled with the vibrational motion of the its center-of-mass. Further, note that this interaction is not the only one possible. If the laser is tuned according to $\omega_{\rm L} = \omega_0 - \nu$ we obtain the interaction

$$\hat{H}_{\rm I} \approx -i\,\hbar\eta\,\Omega\,e^{i\,\phi}\hat{a}\,\hat{\sigma}_- + H.c., \tag{10.71}$$

which contains the non-energy-conserving interaction of the form $\hat{a}\,\hat{\sigma}_-$. This interaction is sometimes known as the anti-Jaynes–Cummings model, an interaction that simply cannot be realized in the case of an atom interacting with a quantized field.

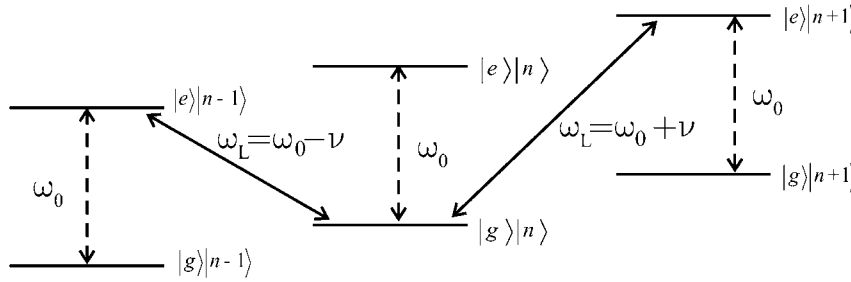


Fig. 10.12. Energy-level diagram for a trapped ion showing the side-band transitions.

By keeping more terms in the expansion (10.67), other higher-order Jaynes–Cummings type interactions can be generated. For example, suppose we retain terms to second order and set $\omega_L = \omega_0 \pm 2\nu$. It should be easy to see that the interactions obtained will be of the form

$$\begin{aligned} \hat{H}_I &\sim \eta^2 \hat{a}^{\dagger 2}\hat{\sigma}_- + H.c., \quad \text{for} \quad \omega_L = \omega_0 + 2\nu, \\ \hat{H}_I &\sim \eta^2 \hat{a}^2\hat{\sigma}_- + H.c., \quad \text{for} \quad \omega_L = \omega_0 - 2\nu. \end{aligned} \tag{10.72}$$

With the laser tuned to the frequencies $\omega_L = \omega_0 \pm l\,\nu$, $l = 0, 1, 2\ldots$, it is evident that interactions of the form

$$\begin{aligned} \hat{H}_I &\sim \eta^l \hat{a}^{\dagger l}\hat{\sigma}_- + H.c., \quad \text{for} \quad \omega_L = \omega_0 + l\,\nu, \\ \hat{H}_I &\sim \eta^l \hat{a}^l\hat{\sigma}_- + H.c., \quad \text{for} \quad \omega_L = \omega_0 - l\,\nu, \end{aligned} \tag{10.73}$$

are possible, though η^l for $l > 2$ may be too small in practice. The frequencies $\omega_L = \omega_0 \pm l\,\nu$ are known as the "side-band" frequencies and are illustrated in Fig. 10.12 for the cases $l = 0$ and $l = 1$.

Various kinds of nonclassical motional states of a single trapped ion, and various types of ion–phonon couplings, have been studied experimentally. Meekhof *et al.* [22], in 1996, were able to create thermal, Fock, coherent, and squeezed states in the motion of a trapped $^9Be^+$ ion initially laser cooled to the zero point energy. They were able to engineer the Jaynes–Cummings interaction and to observe the collapse and revival phenomena in this system. In fact, the collapse and revivals were more pronounced in this system then in the CQED experiment realizing the original model of a two-level atom interacting with a single-mode field [13]. In addition, they were able to construct the anti-Jaynes–Cummings and the "two-phonon" interactions. In later work [23], this group created Schrödinger cat-like states consisting of a single trapped ion where the ion was in a superposition of two Gaussian wave packets separated by a few nanometers – not exactly a macroscopic separation but still large on an atomic scale, a mesoscopic state.

10.8 Concluding remarks

This chapter has mainly been devoted to cavity quantum electrodynamic realizations of the interaction of a quantized radiation field with an effective two-level atom. However, the cavity fields have been taken as microwave fields, not optical fields. This is because as a practical matter it is easier to realize the two-level

approximation for the atom by using circular Rydberg states whose transitions are in the microwave spectrum. The main source of losses in the microwave experiments is the absorption of photons by the walls of the cavity. But cavity QED within the optical spectrum is also an area of active research. An essential difference between microwave and optical cavity QED is that in the latter, one must take into account, assuming the cavity is open, losses from spontaneous emission to modes external to the cavity. This requires and extra dissipative term in the master equation where the field operators $\hat{a}$ and $\hat{a}^\dagger$ are replaced by the atomic transition operators $\hat{\sigma}_-$ and $\hat{\sigma}_+$ respectively. The effects of these additional losses on the collapse and revival phenomenon have been studied by Quang, Knight and Bužek [24].

Finally, another related problem we have ignored is the case when there is more than one atom in the cavity. Dicke [25] studied the collective behavior of a collection of two-level atoms all within a volume $V < \lambda^3$. The atoms do not directly interact with each other but they can act collectively though their coupling to a common field mode. The phenomenon of "superradiance" was predicted whereby, with a properly prepared state of a collection of N atoms, the atoms collectively radiate with an intensity proportional to N^2, whereas the intensity would go only as N if the atoms radiate independently. The extension of the model to the case of a quantized field was first considered by Tavis and Cummings [26]. The mathematics of the operators describing a collection of N identical two-level atoms is, as shown originally by Dicke [25], identical to that of the angular momentum algebra and the collective atomic states may be mapped onto angular momentum states $|J, M\rangle$ where $J = N/2$. The state $|J, -J\rangle = |g\rangle_1|g\rangle_2 \ldots |g\rangle_N$, just the product state of all the atoms with each in the ground state. Similar, $|J, J\rangle = |e\rangle_1|e\rangle_2 \ldots |e\rangle_N$. The angular momentum states for other values of M, however, are *not* product states [27]. Upon this basis it is possible to construct a set of *atomic coherent states* [28], also known as the *spin coherent states* [29], analogous in some ways to the field coherent states. Consideration of multi-atom problems is beyond the scope of this book but the interested reader should consult the references cited and the bibliography.

Problems

1. Justify the claim in Eq. (10.2) that, for adjacent circular Rydberg states of principal quantum numbers n and $n - 1$, the dipole moment goes as $\sim n^2$.
2. Carry through the calculation of the eigenvalues of the matrix in Eq. (10.18).
3. Suppose that an atom prepared in the superposition state

$$|\psi_{\text{atom}}\rangle = \frac{1}{\sqrt{2}}(|e\rangle + e^{i\varphi}|g\rangle), \tag{10.74}$$

is injected into a cavity whose field is initially in a vacuum state. Obtain the time evolution for the excited state population assuming both high and low Q cavities. What is the effect, if any, upon the relative phase φ?

4. For a cavity field initially in a coherent state of average photon number $\bar{n} = 5$ and with the atom initially in the excited state, numerically integrate the set of differential equations associated with the Jaynes–Cummings model master equation of Eq. (10.16) for the cases where (a) $\kappa = 0$, (b) $\kappa = 0.01\lambda$, and (c) $\kappa = 0.03\lambda$. For each case, plot the atomic inversion as a function of the scaled time λt. Also plot the average photon number as a function of the scaled time. Compare your results with those in References [30].
5. Consider the cat-like superposition state

$$|\text{sup}\rangle \sim |\alpha\, e^{i\phi}\rangle + |\alpha\, e^{-i\phi}\rangle. \tag{10.75}$$

   Normalize the state. How does the initial, short time, decoherence rate of this state vary with the angle ϕ?
6. Investigate the possibility of performing a QND measurement of photon number for optical fields. Use the cavity QED method described above as a guide and keep in mind the cross-Kerr interaction described in Chapter 7.
7. For a single trapped ion, obtain the interaction for the case when the laser is tuned to resonance with the internal states of the ion, the $|g\rangle \leftrightarrow |e\rangle$ transition such that $\omega_L = \omega_0$, and where terms of second order in η are retained. Further, if the internal state of the ion is initially the ground state $|g\rangle$, and if the center-of-mass motion is prepared in a coherent state, investigate the evolution of the system. Do the internal and vibrational degrees of freedom become entangled? (Hint: obtain the dressed states of the effective interaction Hamiltonian.)

References

[1] S. Haroche, in *New Trends in Atomic Physics*, G. Grynberg and R. Stora (editors) (Amsterdam: Elsevier, 1984), p. 193.
[2] R. G. Hulet and D. Kleppner, *Phys. Rev. Lett.*, **51** (1983), 1430.
[3] P. Nussenzweig, F. Bernardot, M. Brune, J. Hare, J. M. Raimond, S. Haroche and W. Gawlik, *Phys. Rev. A*, **48** (1993), 3991.
[4] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum Mechanics*, volume 2 (New York: Wiley Interscience, 1977).
[5] F. Bloch, *Phys. Rev.*, **70** (1946), 460.
[6] See L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (New York: Wiley Interscience, 1975) and (Mineola: Dover, 1987), chapter 2.
[7] E. M. Purcell, *Phys. Rev.*, **69** (1946), 681.
[8] D. Kleppner, *Phys. Rev. Lett.*, **47** (1981), 233.
[9] R. G. Hulet, E. S. Hilfer and D. Kleppner, *Phys. Rev. Lett.*, **55** (1985), 2137.
[10] N. F. Ramsey, *Rev. Mod. Phys.*, **62** (1990), 541; *Molecular Beams* (New York: Oxford University Press, 1985).
[11] I. I. Kim, K. M. Fonseca Romero, A. M. Horiguti, L. Davidovich, M. C. Nemes and A. F. R. de Toledo Piza, *Phys. Rev. Lett.*, **82** (1999), 4737.
[12] G. Rempe, H. Walther and N. Klein, *Phys. Rev. Lett.*, **58** (1987), 353.

[13] M. Brune, F. Schmidt-Kaler, A. Maaili, J. Dreyer, E. Hagley, J. M. Raimond and S. Haroche, *Phys. Rev. Lett.*, **76** (1996), 1800.
[14] J. I. Cirac and P. Zoller, *Phys. Rev. A*, **50** (1994), R2799. See also S. J. D. Phoenix and S. M. Barnett, *J. Mod. Opt.*, **40** (1993), 979; I. K. Kudryavtsev and P. L. Knight, *J. Mod. Opt.*, **40** (1993), 1673; M. Freyberger, P. K. Aravind, M. A. Horne and A. Shimony, *Phys. Rev. A*, **53** (1996), 1232; A. Beige, W. J. Munro and P. L. Knight, *Phys. Rev. A*, **62** (2000), 052102.
[15] E. Hagley, X. Maître, G. Nogues, C. Wunderlich, M. Brune, J. M. Raimond and S. Haroche, *Phys. Rev. Lett.*, **79** (1997), 1.
[16] L. Davidovich, M. Brune, J. M. Raimond and S. Haroche, *Phys. Rev. A*, **53** (1996), 1295.
[17] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond and S. Haroche, *Phys. Rev. Lett*., **77** (1996), 4887.
[18] See the review articles by P. Grangier, J. A. Levensen and J.-P. Poizat, *Nature*, **396** (1998), 537; V. B. Braginsky, V. B. Vorontsov and K. S. Thorne, *Science*, **209** (1980), 547. An example of an optical application can be found, for example, in F. X. Kärtner and H. A. Haus, *Phys. Rev. A*, **47** (1993), 4585.
[19] M. Brune, P. Nussenzveig, F. Schmidt-Kaler, F. Bernardot, J. M. Raimond and S. Haroche, *Phys. Rev. Lett.*, **72** (1994), 3339.
[20] See P. Lorrain, D. R. Corson and F. Lorrain, *Fundamentals of Electromagnetic Phenomena* (New York: W. H. Freeman, 2000), p. 57.
[21] G. Sz. K. Horvath, R. C. Thompson and P. L. Knight, *Contemp. Phys.*, **38** (1997), 25.
[22] D. M. Meekhoff, C. Monroe, B. E. King, W. M. Itano and D. J. Wineland, *Phys. Rev. Lett.*, **76** (1996), 1796.
[23] C. Monroe, D. M. Meekhoff, B. E. King and D. J. Wineland, *Science*, **272** (1996), 1131.
[24] T. Quang, P. L. Knight and V. Bužek, *Phys. Rev. A*, **44** (1991), 6092.
[25] R. H. Dicke, *Phys. Rev.*, **93** (1954), 99. See also N. E. Rehler and J. E. Eberly, *Phys. Rev. A*, **3** (1971), 1735, and J. H. Eberly, *Am. J. Phys.*, **40** (1972), 1374.
[26] M. Tavis and F. W. Cummings, *Phys. Rev.*, **170** (1968), 379.
[27] For an elementary account of the mapping of the atomic states onto angular momentum states see M. Sargent, III, M. O. Scully and W. E. Lamb, Jr., *Laser Physics* (Reading: Addison-Wesley, 1974), appendix G.
[28] F. T. Arrechi, E. Courtens, R. Gilmore and H. Thomas, *Phys. Rev. A*, **6** (1972), 2211.
[29] J. M. Radcliffe, *J. Phys. A: Gen. Phys.*, **4** (1971), 313.
[30] S. M. Barnett and P. L. Knight, *Phys. Rev. A*, **33** (1986), 2444; R. R. Puri and G. S. Agarwal, *Phys. Rev. A*, **33** (1986), R3610 and *Phys. Rev. A*, **35** (1977), 3433.

Bibliography

Two books on Rydberg atoms are:

T. F. Gallagher, *Rydberg Atoms* (Cambridge: Cambridge University Press, 1994).

J. P. Connerade, *Highly Excited Atoms* (Cambridge: Cambridge University Press, 1998).

Some review articles on cavity QED with emphasis toward microwave cavities are:

E. A. Hinds, "Cavity Quantum Electrodynamics", in *Advances in Atomic, Molecular, and Optical Physics*, **28** (1991), 237.

D. Meschede, "Radiating atoms in confined spaces: from spontaneous emission to micromasers", *Phys. Rep.*, **211** (1992), 201.

P. Meystre, "Cavity quantum optics and the quantum measurement process", in *Progress in Optics XXX*, edited by E. Wolf (Amsterdam: Elsevier, 1992).
J. M. Raimond, M. Brune and S. Haroche, "Manipulating quantum entanglement with atoms and photons", *Rev. Mod. Phys.*, **73** (2003), 565.
S. Haroche, "Entanglement experiments in cavity QED", *Fortschr. Phys.*, **51** (2003), 388.

A collection of articles covering most aspects of cavity QED is:
P. Berman (editor) *Cavity Quantum Electrodynamics* (New York: Academic Press, 1994).

Some references on ion traps and the physics of trapped ions:
P. K. Ghosh, *Ion Traps* (Oxford: Oxford University Press, 1995).
J. I. Cirac, A. S. Parkins, R. Blatt and P. Zoller, "Nonclassical states of motion in trapped ions", in *Advances in Atomic, Molecular, and Optical Physics*, **37** (1996), 237.
D. Liebfried, R. Blatt, C. Monroe and D. Wineland, "Quantum dynamics of single trapped ions", *Rev. Mod. Phys.*, **75** (2003), 281.

A review covering all aspects of cavity QED, *including optical cavities and collective behavior in multi-atom systems, is*:
S. Haroche, "Cavity Quantum Electrodynamics", in *Fundamental Systems in Quantum Optics, Les Houches Session LIII*, edited by J. Dalibard, J. M. Raimond and J. Zinn-Justin (Amsterdam: Elsevier, 1992).

For an extensive discussion of superradiance, see:
M. G. Benedict, A. M. Ermolaev, V. A. Malyshev, I. V. Sokolov and E. D. Trifonov, *Superradiance* (Bristol: Institute of Physics, 1996).

Chapter 11

Applications of entanglement: Heisenberg-limited interferometry and quantum information processing

"All information is physical", the slogan advocated over many years by Rolf Landauer of IBM, has recently led to some remarkable changes in the way we view communications, computing and cryptography. By employing quantum physics, several objectives that were thought impossible in a classical world have now proven to be possible. Quantum communications links, for example, become impossible to eavesdrop without detection. Quantum computers (were they to be realized) could turn some algorithms that are labelled "difficult" for a classical machine, no matter how powerful, into ones that become "simple". The details of what constitutes "difficult" and what "easy" are the subject of mathematical complexity theory, but an example here will illustrate the point and the impact that quantum information processors will have on all of us. The security of many forms of encryption is predicated on the difficulty of factoring large numbers. Finding the factors of a 1024-digit number would take longer than the age of the universe on a computer designed according to the laws of classical physics, and yet can be done in the blink of an eye on a quantum computer were it to have a comparable clock speed. But only if we can build one, and that's the challenge! No one has yet realized a quantum register of the necessary size, or quantum gates with the prerequisite accuracy. Yet it is worth the chase, as a quantum computer with a modest-sized register could out-perform any classical machine. And in addition, quantum mechanics will allow us to construct a number of novel nonclassical technologies, some requiring already realizable resources, as we will see.

Quantum information processing offers a qualitatively different way in which to think about manipulating information. We may well be forced to adopt this new way as we exhaust classical resources and push traditional information technology into the quantum regime. One of Moore's laws [1], formulated by the founder of INTEL, using historical data, states that the number of transistors per chip (i.e. the complexity of computers) grows exponentially with time; more precisely, it doubles every one and a half years (see Fig. 11.1).

This law has been obeyed almost precisely over the past 30 years. If this exponential growth is extrapolated into the near future we see that, at the rate suggested by Moore's law, a bit of information will be encoded into a single atom

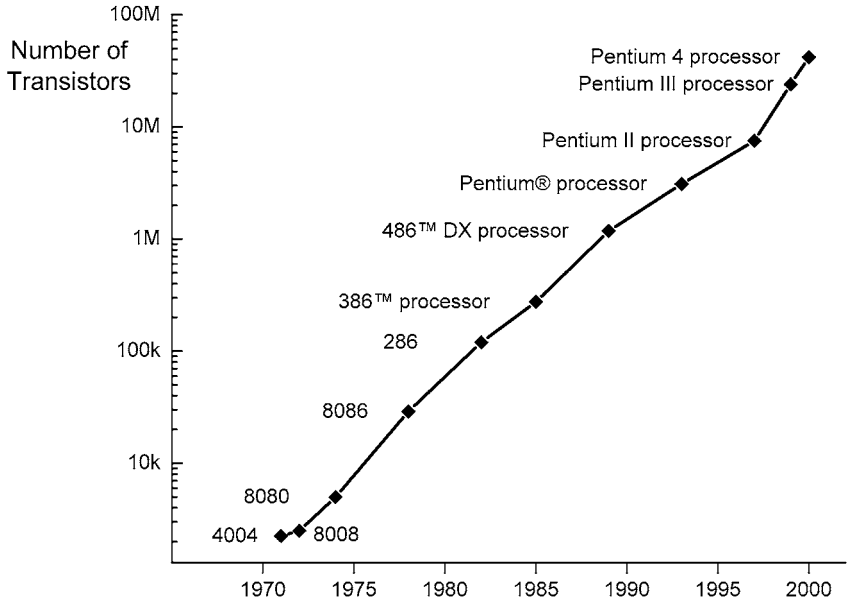


Fig. 11.1. Moore's law, plotting historical data of the number of transistors within a chip as a function of date, demonstrating expanded growth.

by the year 2017. In fact, even before that, by the year 2012, quantum effects will become very important, so much so that they will have a considerable effect upon computation, and should no longer be neglected. Not only our theoretical curiosity, but also technological progress, requires that we study and understand quantum information processing. But in this chapter we are interested in what quantum mechanics allows us to do, which is substantially different from what classical physics allows. The implication of Moore's law is that there will be about one electron per transistor by around the year 2017 and that this lone electron will be confined to a region small enough that it will act as a quantum-mechanical particle, and not as a classical charged billiard ball. Fortunately, at this point, quantum physics not only becomes more important in designing classical computers, but offers a whole new field of quantum computing with an entirely different mindset.

11.1 The entanglement advantage

We have already discussed in previous chapters, and in Appendix A, the nature of entanglement and some of its consequences such as the violation of Bell's inequalities. Entanglement can loosely be described as a kind of correlation. But it is quite a bit more than that and, besides, classical systems can also exhibit correlations. A correlation between two systems is simply the statement that if a measurement of one system yields the result A then a measurement on the second system will yield the result B with some probability. Perfect correlation occurs when the second result is certain, given the outcome of the first. Some readers

may be familiar with John Bell's paper "Bertlmann's Socks and the Nature of Reality" [2]. Eccentric Professor Bertlmann always wears socks of two different colours, one of which is always pink. If an observer sees only one of Bertlmann's feet and notices a pink sock, then he knows that the other is certainly not pink. On the other hand, if he observes, say, a green sock, then he knows that the other *is* pink. The correlations in this case are entirely of a classical nature and there is nothing mysterious in this example. Products of spin states of the form $|\uparrow\rangle_1|\downarrow\rangle_2$ and $|\downarrow\rangle_1|\uparrow\rangle_2$ have obvious correlations and, though the spin states of each particle are quantum mechanical, the correlations between the states are purely classical. For Bell states of the form

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1|\downarrow\rangle_2 \pm |\downarrow\rangle_1|\uparrow\rangle_2\right), \tag{11.1}$$

there are also classical-like correlations in the sense that a measurement of the z-component of the spin of one particle also determines the outcome of a measurement of the z-component of the spin of the other, although in the case of the Bell states there is an element of probability involved: 50% of the experimental runs give $|\uparrow\rangle_1|\downarrow\rangle_2$ and 50% $|\downarrow\rangle_1|\uparrow\rangle_2$ and thus outcomes of measurements of spin along the z-direction are always correlated. But when we consider measurements of spin along other directions, the product states exhibit no correlations between measurements on the different particles whereas the Bell states do exhibit certain kinds of correlations, correlations that lead to violations of Bell-type inequalities. Bell's inequality, as should be clear from the discussion in Chapter 9, can be violated only by states that are entangled.

As we have previously said, so far, most (if not all) of the experimental tests of Bell's inequalities support quantum mechanics. But entanglement is now viewed not merely as an esoteric feature of quantum mechanics suitable only for metaphysical contemplation and tests of local hidden variable theories, but rather it serves as the basis for a new technology. The strong nonclassical correlations possessed by entangled states are precisely what we may exploit in a number of quantum technologies that we now describe.

11.2 Entanglement and interferometric measurements

The key to harnessing the potential of entanglement turns out to be related to our ability to count the number of particles in a system. We can illustrate this by considering the problem of quantum-enhanced phase measurement. Let us say that you wish to determine the phase shift induced by some optical element on a beam. One way to do this is to put the element into one arm of an interferometer, and to illuminate the interferometer input. Then a measurement of the difference photocurrent from two detectors looking at the output ports of the interferometer will exhibit a sinusoidal fringe pattern as the phase is changed. A particular phase can be measured by comparing the difference photocurrent with the known

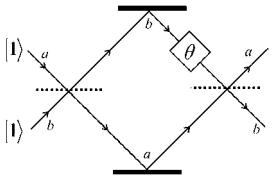


Fig. 11.2. Schematic of an interferometric measurement of a phase shift with twin single-photon states as inputs. Parity measurements must be performed on one of the out put beams.

maximum and minimum photocurrents. The accuracy of this phase measurement will be set by the photocurrent fluctuations. If the input light is classical, these will be at the shot noise limit. That is, if the input radiation has a mean photon number $\bar{n}$, the accuracy of the phase measurement will be $\Delta\theta = 1/\sqrt{\bar{n}}$ as we saw in Chapter 6.

Phase-measurement accuracy can be improved if nonclassical light is used to illuminate the interferometer. Recall that when light in a coherent state (classical light) is injected into an interferometer, the light inside is always describable by separable states; the states of light along the two beam paths never become entangled. But now consider the case when exactly one photon is put into each of the input ports a and b of the interferometer. The state of the system at the input can be denoted $|1\rangle_a|1\rangle_b$. After the beam splitter, the system is in an entangled state, $(|2\rangle_a|0\rangle_b + |0\rangle_a|2\rangle_b)/\sqrt{2}$ as given in Eq. (6.17), apart from an irrelevant overall phase. If the phase shifter is in the b-arm of the interferometer, as pictured in Fig.11.2, then just before the second beam splitter, the state is $(|2\rangle_a|0\rangle_b + e^{2i\theta}|0\rangle_a|2\rangle_b)/\sqrt{2}$, and after the second beam splitter the state is

$$|\text{out}\rangle = \frac{1}{2\sqrt{2}}(1 - e^{2i\theta})\,(|2\rangle_a\,|0\rangle_b - |0\rangle_a\,|2\rangle_b) + \frac{i}{2}(1 + e^{2i\theta})\,|1\rangle_a\,|1\rangle_b\,. \tag{11.2}$$

Now to measure the phase shift we might try to follow the procedure used in the case of the coherent state input, or even for the case when the input is $|1\rangle_a|0\rangle_b$, as we discussed in Chapter 6. In those cases we obtained an oscillatory function of the phase shift θ. But in the present situation there is a problem. If we calculate the average photon number in each of the output beams, that is, $\langle \hat{a}^\dagger_{\text{out}}\hat{a}_{\text{out}}\rangle$ and $\langle \hat{b}^\dagger_{\text{out}}\hat{b}_{\text{out}}\rangle$, we find that they both have a value of unity and do not depend on θ. Obviously their difference will vanish as well. (Note, though, that the sum of these averages is 2, the total number of photons involved, as expected.) So the standard measures to obtain θ won't work. However, we are free to take a different observable to measure and the one that seems appropriate is that of the parity operator (the evenness or oddness of the photon numbers) in just one of the output beams. Suppose we decide to measure the parity of the output b-mode [3]. The parity operator of this mode is

$$\hat{\prod}_b = (-1)^{\hat{b}^\dagger \hat{b}} = \exp(i\,\pi\;\hat{b}^\dagger \hat{b}). \tag{11.3}$$

The expectation value of this operator with respect to the state of Eq. (11.2) is easily found to be

$$\left\langle \hat{\prod}_b \right\rangle = \left\langle \text{out} \left| \hat{\prod}_b \right| \text{out} \right\rangle = \cos(2\theta) \tag{11.4}$$

which depends not just on θ but on 2θ. If we once again use the calculus of error propagation, we will obtain, rather generally, the uncertainty in the phase measurements as

$$\Delta\theta = \Delta\prod_b \Big/ \left| \frac{\partial\,\langle \hat{\prod}_b \rangle}{\partial\theta} \right|, \tag{11.5}$$

where $\Delta\prod_b = \sqrt{1-\langle\hat{\prod}_b\rangle^2}$ owing to the fact that $\hat{\prod}_b^2 = \hat{I}$, the identity operator. In our particular case, we obtain $\Delta\theta = 1/2$. If we compare this result with the optimal coherent state result for the *same average photon number* at the input, $\Delta\theta = 1/\sqrt{2}$ for $\bar{n} = 2$, we see an improvement in the phase sensitivity, the resolution, by a factor of $1/\sqrt{2}$. This improvement is the result of the entanglement of the state inside the interferometer. In fact the state, the state of Eq. (6.17), is a special type of state known as a maximally entangled state (MES). It is "maximal" in the sense that it is a superposition of states where all the photons are in one beam and where they are all in the other and maximizes the correlations between the modes. A more general form of such a state for a total of N photons is

$$|\psi_N\rangle_{\text{MES}} = \frac{1}{\sqrt{2}}(|N\rangle_a\,|0\rangle_a + e^{i\Phi_N}\,|0\rangle_a\,|N\rangle_a), \tag{11.6}$$

where Φ_N is some phase that may depend on N. *If* such states could be generated, and that's a big "if", one can show that phase resolutions of the exact form $\Delta\theta = 1/N$ will result, an improvement of the optimal coherent state result with light of equivalent average photon number $\bar{n} = N$. This form represents, in fact, the tightest measure allowed by quantum mechanics for the interferometric measurement of phase shifts [4], and is referred to as the Heisenberg limit (HL): $\Delta\theta_{\text{HL}} \equiv 1/N$.

Of course, the maximally entangled states, sometimes called the *NOON* states for what we hope are obvious reasons, must be realized experimentally somehow. For the case of $N = 2$, as we have seen, the twin single-photon states at the input of a 50:50 beam splitter are sufficient to generate the two-photon maximally entangled state. But for arbitrary $N > 2$, a beam splitter will not generate maximally entangled states. However, a number of schemes have been proposed for generating such states, some involving various kinds of state-reductive measurements [5] and others involving nonlinear media with giant susceptibilities [6]. However, the necessary parity measurements require photon detectors able to distinguish photon numbers at the level of a single photon. But detectors with such resolutions are becoming available as we write [7]. As another tack, instead of generating maximally entangled states, one could instead feed twin number states $|N\rangle_a|N\rangle_b$ into a Mach–Zehnder interferometer [8]. At the output, the differences of the average photon numbers vanish as in the case for a maximally entangled state, but, once again, parity measurements can be performed on one of the output beams. It has been shown that the Heisenberg limit in the form $\Delta\theta_{\text{HL}} = 1/(2N)$, the total number of photons being $2N$ in this case, is approached as N becomes large, and is substantially below the standard quantum limit $\Delta\theta_{\text{SQL}} = 1/\sqrt{2N}$ for all N, at least if the phase shift is small [9]. But leaving aside all the details of these considerations, which are beyond the scope of this book, the essential point of this discussion is that entanglement is not merely a curiosity, but does have practical applications; in this case the improvement of interferometric measurements.

11.3 Quantum teleportation

Another application of entanglement, this one much more in the context of quantum information processing, is quantum teleportation whereby an unknown quantum state is transferred from one point to another, these points possibly being widely separated. We wish to emphasize that it is the unknown quantum *state* that is to be teleported, not the particle or particles in such states. We shall restrict ourselves to a simple and rather basic description of teleportation, as originally given by Bennett *et al.* [10]. We shall use states denoted $|0\rangle$ and $|1\rangle$ as usual, but we will not necessarily think of them in terms of photon number states but rather as states of some sort of "two-level" system, such as the polarization states of single photons. Furthermore, in the quantum information business, the symbols $|0\rangle$ and $|1\rangle$ and superpositions of these states represent the quantum bits of information, the *qubits*. Qubits are the quantum information theoretic name for the states of any "two-level" system and a *single qubit* refers to one system whose states can be represented, in general, by superpositions of the states $|0\rangle$ and $|1\rangle$. Qubits could be realized also by two-level atoms. As this text is concerned in the main with photons, one might wish to interpret the states $|0\rangle$ and $|1\rangle$ as vertically and horizontally polarized states, $|V\rangle$ and $|H\rangle$ respectively, but the procedures described are rather general.

We require two participants in the teleportation procedure (or protocol) usually given the names Alice (A) and Bob (B). We suppose that Alice is given a photonic mode in the quantum state $|\Psi\rangle = c_0|0\rangle + c_1|1\rangle$ and that she wishes to teleport the state to Bob so that he can re-create it in a photonic mode in his possession. The state is unknown to Alice, i.e. she does not know the coefficients c_0 and c_1. In fact, if she did know, the state the protocol about to be described would be unnecessary as she could simply convey the information to Bob over a classical channel (e.g. via telephone). We further suppose that some source of light can produce an entangled state of the form

$$|\Psi_{\mathrm{AB}}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{\mathrm{A}}\,|0\rangle_{\mathrm{B}} + |1\rangle_{\mathrm{A}}\,|1\rangle_{\mathrm{B}}\right), \tag{11.7}$$

which is shared between Alice and Bob as indicated by the subscripts. So, Alice has in her possession the state to be teleported, $|\psi\rangle$,and part of the shared state $|\Psi_{\mathrm{AB}}\rangle$, whereas Bob has only his part of the shared state, so far. Thus we may write the total state for Alice and Bob, at this point, as

$$|\Phi_{\mathrm{AB}}\rangle := |\psi\rangle\,|\Psi_{\mathrm{AB}}\rangle = (c_0\,|0\rangle + c_1\,|1\rangle)\,(|0\rangle_{\mathrm{A}}\,|0\rangle_{\mathrm{B}} + |1\rangle_{\mathrm{A}}\,|1\rangle_{\mathrm{B}})/\sqrt{2}. \tag{11.8}$$

We can expand this as

$$\begin{aligned}|\Phi_{\mathrm{AB}}\rangle &= \frac{1}{\sqrt{2}}(c_0|0\rangle|0\rangle_{\mathrm{A}}|0\rangle_{\mathrm{B}} + c_0|0\rangle|1\rangle_{\mathrm{A}}|1\rangle_{\mathrm{B}} + c_1|1\rangle|0\rangle_{\mathrm{A}}|0\rangle_{\mathrm{B}} + c_1|1\rangle|1\rangle_{\mathrm{A}}|1\rangle_{\mathrm{B}})\\ &= \frac{1}{2}[|\Phi^+\rangle(c_0|0\rangle_{\mathrm{B}} + c_1|1\rangle_{\mathrm{B}}) + |\Phi^-\rangle(c_0|0\rangle_{\mathrm{B}} - c_1|1\rangle_{\mathrm{B}})\\ &\quad + |\Psi^+\rangle(c_0|1\rangle_{\mathrm{B}} + c_1|0\rangle_{\mathrm{B}}) + |\Psi^-\rangle(c_0|1\rangle_{\mathrm{B}} - c_1|0\rangle_{\mathrm{B}})],\end{aligned} \tag{11.9}$$

where we have introduced the Bell states

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle_A + |1\rangle|1\rangle_A), \tag{11.10}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle_A - |1\rangle|1\rangle_A), \tag{11.11}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle_A + |1\rangle|0\rangle_A), \tag{11.12}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle_A - |1\rangle|0\rangle_A), \tag{11.13}$$

which are mutually orthogonal and thus constitute a basis in a four-dimensional Hilbert space. These states are constructed from the basis states of the unknown state to be teleported and of that part of the entangled state shared by Alice. Each Bell state in Eq. (11.9) is correlated with a different superposition of the states in Bob's share of the entangled state provided. Perhaps it is important to point out that nothing physical has happened here yet: we have merely re-written the direct product of the state to be teleported and the provided, shared, entangled state. The next step in the protocol is for Alice to perform projective measurements onto the Bell basis. Each of the four Bell states $|\Phi^\pm\rangle$, $|\Psi^\pm\rangle$ will occur randomly with an equal probability of $\frac{1}{4}$. Suppose Alice obtains the state $|\Phi^+\rangle$ and that she *knows* that she has obtained that state. Bob's photonic system is then projected onto the state $c_0|0\rangle_B + c_1|1\rangle_B$. Over a classical channel Alice then tells Bob that she has detected the Bell state $|\Psi^+\rangle$ and thus both know that Bob already has in possession the teleported state and thus he needs to do nothing. Note that neither Alice nor Bob knows what that state is. On the other hand, if Alice reports that she has detected the state $|\Phi^-\rangle$, then Bob's system has been projected onto the state $c_0|0\rangle_B - c_1|1\rangle_B$,and he knows that his state differs from the original by the sign of the second term so that he can perform the transformation $|0\rangle_B \to |0\rangle_B$, $|1\rangle_B \to -|1\rangle_B$ to obtain the original state. If Alice detects $|\Psi^+\rangle$, Bob has $c_0|1\rangle_B + c_1|0\rangle_B$ and he needs to perform the "flip" operation $|1\rangle_B \to |0\rangle_B$, $|0\rangle_B \to |1\rangle_B$, the equivalent of a logical NOT operation. Finally, if Alice detects $|\Psi^-\rangle$ then Bob has $(c_0|1\rangle_B - c_1|0\rangle_B)$ and must perform the transformation $|1\rangle_B \to |0\rangle_B$, $|0\rangle_B \to -|1\rangle_B$. This completes the description of the teleportation protocol.

Within the two-dimensional basis $(|0\rangle_B, |1\rangle_B)$, the above operations can be represented by set of 2×2 matrices $\{I_2, \sigma_x, \sigma_y, \sigma_z\}$ where the first is the two-dimensional identity matrix and the rest are the familiar Pauli matrices, as the reader can easily check. Finally, notice that at no time did anyone know the original state being teleported. In the process of performing the Bell state measurements, this original state is itself destroyed.

Quantum teleportation may seem to have an element of magic to it and, if so, this is because the state shared by Alice and Bob is entangled. If we tried

teleporting a state by allowing Alice and Bob to share the statistical mixture

$$\hat{\rho}_{AB} = \frac{1}{2}\left(|0\rangle_A |0\rangle_B {}_A\langle 0|{}_B\langle 0| + |1\rangle_A |1\rangle_B {}_A\langle 1|{}_B\langle 1|\right), \tag{11.14}$$

teleportation would be impossible, as the reader can check. Notice, though, that the mixed state of Eq. (11.14) does, in fact, exhibit correlations; but these correlations are completely of a classical nature.

Quantum teleportation is not merely a theoretical curiosity but is an experimental reality. The initial experiments teleporting qubit states, where the qubits were represented by photon polarization states, were performed by groups in Rome [11] and Innsbruck [12]. Subsequently, a group at Caltech teleported a state with continuous variables [13], one made from squeezed states.

11.4 Cryptography

Secure communication is an area of great importance as we move into an information age. Such communication is built on crypto-systems, whose security is based on assumptions concerning the computational difficulty of solving various problems, especially of factoring. An alternative is quantum cryptography, which is based on a fundamental description of nature, namely quantum mechanics, and where security is ensured by the nature of measurements in quantum physics.

The rapid rise of electronic communication and commerce has led to increasing concern with the security and authentication of electronic messages. This is, of course, not a new problem. Since the earliest records of humanity there has been a need and requirement to pass secrets between different parties. For millennia communicating parties have devised schemes whereby messages can be authenticated (the signature) and secured from unauthorized access (cryptography). Modern methods for secure communication always involve the prior exchange of a random number or binary string, called the key. If the communicating parties share this number with each other and no one else, messages can be securely encrypted and decoded. The method, however, is vulnerable to a third party acquiring access to the key. Quantum mechanics enables two communicating parties to arrive at a shared secure key and yet be certain that no eavesdropper has acquired the key. Quantum cryptography will be reviewed in this section.

Quantum mechanics has the potential to revolutionize information processing in at least two almost complementary ways. Firstly, if a quantum computer can be realized, it could be utilized to factor large numbers beyond the capability of any classical computer. Since (as we will see), the difficulty of factoring underpins secure classical communication, such a quantum computer threatens all current cryptographic security. But while quantum mechanics can undermine cryptographic security, it can also save it. We will show that the basic laws of quantum mechanics can allow us to construct quantum cryptographic protocols secure against interception.

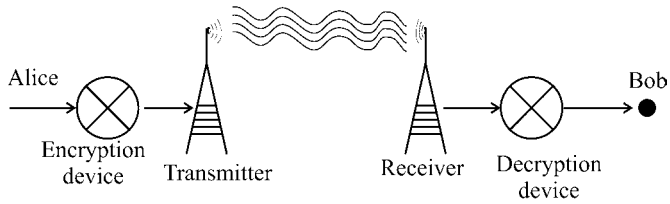


Fig. 11.3. Schematic of a crypto-system.

This section is structured as follows. We will begin by providing an introduction to classical private and public key crypto-systems. We will then consider several of the quantum protocols and their experimental realization in quantum optics. However, before we proceed further, let us define a few essential terms and concepts.

- *Alice and Bob*. Alice is an individual who wants to send a private message to Bob.
- *Eve* is an eavesdropper who wants to intercept the message sent by Alice to Bob and read it. We assume that Eve has full access to the communications channel used to transmit the message.
- *Public channel*: a communication channel through which information is transmitted from one endpoint to another. The eavesdropper has full access to all information passed through this channel.
- *Private channel*: a secure channel, to which the eavesdropper has no access. It is believed to be impossible to eavesdrop on information transmitted through this channel. This channel is often referred to as a trusted channel.
- *Key material*: a sequence of preferably random numbers known only to Alice and Bob.
- *Symmetric algorithm*: the symmetric key gets its name from the fact that the encryption key can be calculated from the decryption key and vice versa. In most cases the encryption key and decryption key are the same.
- *Public key algorithm*: the public key algorithms are asymmetric in nature and are designed such that the key used for encryption is different from the key used for decryption. Furthermore the decryption key cannot be calculated from the encryption key in a reasonable amount of time.

11.5 Private key crypto-systems

Until about 1970, all crypto-systems generally operated on a private key principle. In this situation, the two parties (Alice and Bob) wishing to communicate must have established a shared key beforehand. If Alice wishes to send a message to Bob, then Alice encrypts the message with an encoding key. This encrypted message is then transmitted to Bob in the open (via a public channel). Bob then uses his decoding key (which was established with Alice beforehand) to decrypt the transmitted signal and hence receive the original message from Alice, as indicated in Fig. 11.3.

One of the earliest examples is the so-called *Caesar crypto-system*, which involves translating our normal alphabet into numbers. One simplest mechanism is to translate our normal 26 character alphabet (A–Z) into the numerical equivalents 0 . . . 25 (for instance A = 0, . . . , Z = 25) and add a constant number, 10 say. To perform this operation one can define a function such that

$$f(x) = \begin{cases} x + 10, x < 15 \\ x - 15, x \geq 15 \end{cases} \tag{11.15}$$

or in the more compact form

$$f(x) = x + 10 \,(\mathrm{mod} 26)\,. \tag{11.16}$$

As an example, consider that Alice wants to send Bob the word "HELLO". She begins by translating the word "HELLO" into the numerical equivalents "7 5 11 11 14". She then applies the above function to get her encrypted message "17 15 21 21 24" which can be transmitted to Bob over an open channel. Having previously established the nature of the encoding, Bob knows the inverse function $f^{-1}(x)$ and hence decodes the encrypted message "17 15 21 21 24" to "7 5 11 11 14". The message is then simply converted back into the word "HELLO".

The above protocol requires that effectively both Alice and Bob know the form of the encoding function (this has to be established previously by some secure means). Alice uses an encoding function and Bob uses the decoding function that is the inverse of the encoding function. A more complicated version of the encoding could be given by

$$f(x) = ax + b \,(\mathrm{mod}\ N)\,, \tag{11.17}$$

where a and b are constants. The problem with this kind of encoding is that once an eavesdropper (Eve) knows the form of the function, then she can simply decode the message the same way Bob does. If the same encoding is used repeatedly, then Eve can attack the encoding and it can be broken.

A modification to the encoding scheme proposed above does allow a secure encryption scheme. A very well known private key crypto-system is the *Vernam cipher* (also called the *one-time pad*). This scheme works by Alice and Bob initially sharing a secret key string. This must be done privately and securely. When Alice and Bob want to pass a message, Alice encodes her message by adding the message and key string together. This is then transmitted over a public channel to Bob. Bob recovers the message by subtracting the same key string from the encoded message. The one-time pad is a well-used provably secure private key crypto-system. As its name indicates, the one-time pad is secure only if the key material is not reused. In fact, the key material must be at least as large as the size of the message for absolute security.

One of the major difficulties with the one-time pad is that it requires a huge amount of key material that must be distributed to the parties wishing to

communicate. This cannot be done over a public channel, which makes it impractical for general usage, but it is still used today in situations requiring provable security. Certain forms of banking still use the Vernam cipher. Today, most of the current crypto-systems use a public key arrangement.

11.6 Public key crypto-systems

The private key crypto-systems we have considered previously have relied on the distribution of key material. There is another approach that comes under the heading of public key crypto-systems, which does not require the prior secure communication of key material. Perhaps the most widely used one is the RSA (for Rivest, Shamir and Adleman [14]) system. The RSA crypto-system is based on the fact that in order to factorize a number with N digits, a classical computer needs to perform a number of steps that at least grows faster than polynomially with N. Consider the factorizing of the numbers 21 and 1073 by hand: the number 21 is easy but 1073 takes a little more time. Computers have the same problem. As N becomes large, factoring becomes a hard problem on a classical computer and, in fact, the number of steps required becomes exponential with N. This is how the security of the key arises. If these large numbers can be factored, then an eavesdropper can recover the message sent from Alice to Bob. Let us examine the RSA protocol which allows Bob to announce publicly a key (the public key) such that Alice can use the key, encrypt a message and send it publicly back to Bob. The protocol only allows Bob to decipher it. The RSA protocol is established as follows.

- It begins by Bob taking two very large prime numbers p and q (of size greater than 10^{1000}). He then calculates the quantities $N = pq$ (where obviously $N > 10^{2000}$ and $pq(N) = (p-1)(q-1)$). Bob then takes an integer $e < N$ such that the greatest common divisor between e and $pq(N)$ is one in which case e and $pq(N)$ are said to be *coprime* (i.e. they have no common factors other than 1). Once this is achieved, Bob computes $d = e^{-1} (\mathrm{mod}\, pq(N))$.
- Bob then sends the public key (e, N) to Alice via a public channel.
- Alice now encodes her plain text message m according to the rule $c = m^e (\mathrm{mod}\, N)$ and sends the result c to Bob.
- Once Bob receives the encrypted message c from Alice he computes $c^d (\mathrm{mod}\, N)$ and recovers the plain text message m.

It is interesting to observe that most of the work in this protocol is done by Bob (not Alice). Alice's main job is to encode the message m she wishes to send to Bob with the public key (e, N).

Given that all the communication is done only via a public channel, what is an effective mechanism for an eavesdropper to break this crypto-system? Eve can attack this crypto-system quite easily. To do this Eve must simply solve $d = e^{-1} (\mathrm{mod}\ pq(N))$ where she knows only e and N. Effectively this just requires

Eve to factor N. Once she factors N she obtains p and q and hence can calculate $pq(N)$. While this may sound like a relatively simple attack, it is not practical to factorize numbers of the size of $N > 10^{2000}$ on a classical computer with current technology in a reasonable time (in fact, to factor the numbers currently used in the RSA encryption system would take a very large supercomputer many, many, years). However, a recent proposal for a new kind of computer architecture, the quantum computer, opens the possibility for factoring these very large numbers in a short time.

The quantum computer, of course, has yet to be realized, but has the potential to revolutionize information processing through its ability to access a vastly larger state space than a classical machine. This, in turn changes the complexity classification of how difficult computational tasks are, so that some tasks which are known to be extremely difficult (that is, require exponential resources) on a classical machine become easier (need only polynomial resources) on a quantum machine. The gain comes from using two features of quantum physics: entanglement and quantum interference, allowing massively parallel processing capacity.

In 1985 David Deutsch [15] (with a generalization in 1992 co-authored with Richard Jozsa [16]) proved that in quantum mechanics the complexity level of some problems can change dramatically. This paved the way for a number of advances in quantum algorithms, in particular that of Peter Shor's for factoring large numbers. In 1994, Peter Shor [17] discovered a quantum algorithm that allows one to factor large numbers in polynomial time so that, provided one had a quantum computer, factoring would become essentially as easy as multiplying. Recently, a small-scale NMR quantum machine has factored the number 15 [18]. While this may be a small step forward, it shows the potential of the quantum computer as a demonstration of the implementation of a non-trivial algorithm. The central point here is that if quantum computing schemes can be scaled up to the point where very large numbers can be factored, RSA crypto-systems become vulnerable to deciphering. The NMR scheme is apparently not scalable to the required degree.

11.7 The quantum random number generator

Classical computers are frequently called upon to generate strings of random numbers. But the numbers generated are only pseudo-random as a deterministic algorithm is always involved. In many of the private key cryptography protocols it is necessary to have local access to a good source of true (and not pseudo-) random numbers. This is also true for the quantum cryptographic schemes we will describe below. While deterministically generated pseudo-random numbers can appear quite random, if one knows the algorithm and the seed, there is the possibility that the string could be reproduced by an eavesdropper. There are also a

number of hardware-based random generators based on the complexity of thermal noise fluctuations that produce a chaotic behaviour in the system. The difficulty in predicting the result of the chaotic process is assimilated to randomness. Under ideal conditions this produces a good source of random numbers. A potential problem, however, is that the thermal noise could be influenced and potentially controlled by someone who tampers with the external environment. In recent years a simple and elegant solution has arisen based on the principles of quantum mechanics (see Fig. 11.4).

This device described in Fig. 11.4 relies on the properties of a 50:50 beam splitter as discussed in Chapter 6. We take as inputs a single photon state in one port and the vacuum in the other. The detectors placed at the outputs yield which-path information. We assign a zero if the photon takes the reflected path and a one if it takes the transmitted path. As there is a 50% chance of it taking each path, we have a 50% chance of getting a zero or a one. This is our random bit. By performing this a number of times, a sequence of truly random bits can be generated.

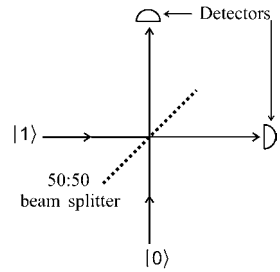


Fig. 11.4. Schematic setup of a quantum random number generator. The apparatus consists of a single photon state arriving at a 50:50 beam splitter. A vacuum enters the second port of the beam splitter. Assuming that the beam splitter is 50:50, the single photon has a 50% chance of being transmitted through the beam splitter and a 50% chance of being reflected. A random bit is obtained by measuring whether the photon was reflected or transmitted.

11.8 Quantum cryptography

Given that the quantum computer (a device yet to be built) is able to attack current public key crypto-systems (such as RSA), what means of absolute secure communication is possible? We know the private key crypto-system (one-time pad) is not vulnerable to the quantum computing architecture described above. However one-time pad classical schemes require a large amount of key material being shared between the two parties Alice and Bob. This has had to be done via private channels, to which the eavesdropper has no access. There are a number of practical solutions to this, such as Alice sending Bob a sealed can by courier with the key material within. Here Alice and Bob must trust the courier not to have read and copied the key material. Another solution based on the principles of quantum mechanics is possible, namely *quantum key distribution* (QKD) and it provides a mechanism for Alice and Bob to create a shared key over public channels. Once both parties have this key material, secure communication can take place. In all QKD proposals it is critical that one has a source of random numbers.

11.8.1 Quantum key distribution

Quantum key distribution offers the possibility for two remote parties – Alice and Bob – to exchange a secret key without physically meeting or requiring a trusted party to deliver the key material. The security of QKD is conditioned on the principles of quantum mechanics, a theory well tested and thought to be correct.

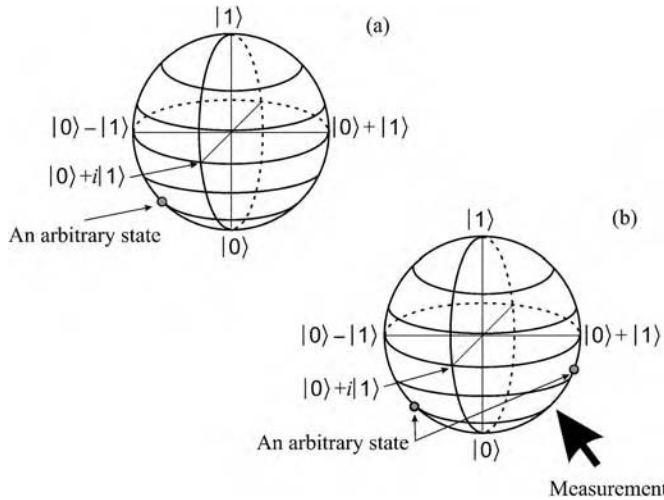


Fig. 11.5. Diagram representation of a qubit on the surface of a sphere (a). The ground state $|0\rangle$ is situated at the south pole of the sphere. The excited state $|1\rangle$ is situated at the north pole of the sphere. Superpositions of $|0\rangle$ and $|1\rangle$ lie on the surface of the sphere between the north and south poles. Equal superpositions of $|0\rangle$ and $|1\rangle$ lie on the equator. An arbitrary state can lie anywhere on the sphere. In (b) we show a measurement of a point on the sphere. On this sphere we have indicated two unknown states. These states are not orthogonal and so a single measurement cannot confirm which one of the two states we have. It is this key point we exploit in quantum key distribution.

The central idea behind quantum key distribution is that it is impossible for an eavesdropper to obtain all the information from the transmitted quantum state. What does this statement mean? Consider a single qubit in the superposition state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$. (We again use the symbols $|0\rangle$ and $|1\rangle$ to represent in a general manner the basis of our qubits. For a realization in terms of polarization states, we could again take $|0\rangle = |V\rangle$ and $|1\rangle = |H\rangle$.) For a single general measurement with c_0 and c_1 unknown, it is impossible to determine the state $|\psi\rangle$ precisely. For instance if a zero "0" is obtained on a single measurement, it is impossible to determine what the c_0 coefficient was (see Fig. 11.5). A single measurement yielding a result "0" does not distinguish between $|0\rangle$, $(|0\rangle + |1\rangle)/\sqrt{2}$, or $(|0\rangle + i|1\rangle)/\sqrt{2}$, or many other possible states for that matter. Over many runs with identical copies of the state, one can determine, via a number of measurements, the exact state. However, in QKD, the qubit is never reused and so it is impossible for an eavesdropper to determine the state completely if several non-orthogonal states are used. The key to QKD is the use of these non-orthogonal states.

There are a number of QKD protocols that we will examine in detail below. We will begin with the first one discovered, the BB84 protocol.

11.8.2 The BB84 protocol

The BB84 protocol (Bennett and Brassard [19]) begins by Alice creating and sending to Bob a random set of qubits, where the choice of qubits comes from the set

$$\begin{aligned} |\psi_0\rangle &= |0\rangle\,, \\ |\psi_1\rangle &= |1\rangle\,, \\ |\psi_+\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \\ |\psi_-\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right). \end{aligned} \tag{11.18}$$

Note that $\langle\psi_0 \mid \psi_1\rangle = 0$ and $\langle\psi_+ \mid \psi_-\rangle = 0$. The first two states form a basis and the second two states form another. In terms of single-photon polarization states, the second two states are oriented at $\pm 45^\circ$ to the first two. Clearly, the four states of Eq. (11.18) are not *all* orthogonal to each other: $\langle\psi_{0,1} \mid \psi_{+,-}\rangle \neq 0$. Hence there is no measurement protocol, which allows us with 100% certainty to determine the states Alice sends to Bob. Bob measures randomly in one of two bases, either "0" and "1" or "+" and "−" (here "+" and "−" represent measuring $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$). After the measurements are complete, Alice and Bob communicate via a public channel as to certain properties of the states sent and measured. More precisely Alice announces whether the qubit was sent in the 0,1 basis or +, – basis but not which exact state. Bob then announces which basis he measured in, but not the result. They then keep between them the results only if Alice's sending and Bob's measurement basis are the same. This gives a smaller subset of qubits (half the size) where Alice and Bob know what was sent and measured without having communicated exactly which state was sent. A key has then been established.

When Alice sends the qubits to Bob, it is possible that an eavesdropper (Eve) may intercept some or all of the transmitted qubits and interfere with them. Now what is the action of the eavesdropper? Basically the eavesdropper must perform some action between the time that Alice sends her quantum state and the time Bob receives it. This could be any form of measurement or no measurement. Let us consider the case when Eve performs a definite projective measurement. Eve does not know which state has been sent (and can not determine this at the time of her measurement), and so must choose whether to measure in the 0,1 basis or +,− basis. If Eve guesses the right basis (and she will not know this until after Alice and Bob have communicated), then she can transmit the right state to Bob (so that he does not know that it has been intercepted). If, however, Eve chooses the wrong basis to measure in, then, the state she transmits is wrong and not what Bob should get. If Bob performs some form of error checking, then he will see that the wrong state has been sent and hence that Eve is present.

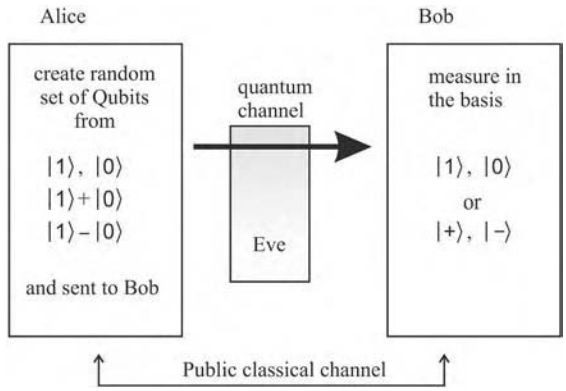


Fig. 11.6. Summary of the procedure to implement the BB84 protocol. After Bob's measurement, Alice and Bob communicate whether they sent and measured a state in $|0\rangle$, $|1\rangle$ or $|+\rangle$, $|-\rangle$ basis. They keep the results only when the sent and measured bases agree. Then they use a subset of this basis-agreed data to check for the presence of Eve.

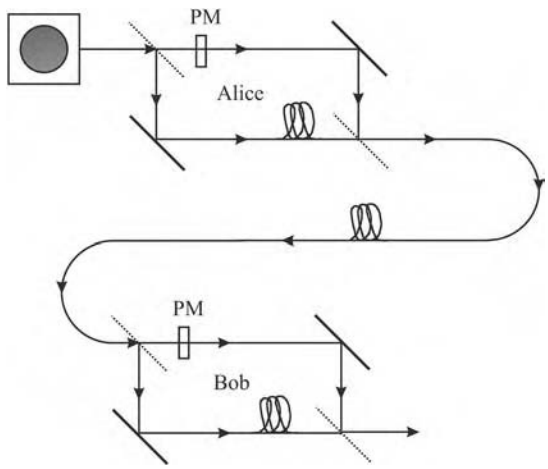


Fig. 11.7. Schematic representation of a realisation of the BB84 protocol using the interferometer. In this realization, the shorter and longer paths through the interferometer define the 0 and 1 states (instead of using polarization, which is hard to preserve in commercial fiber). Phase modulators (PM) are positioned within the upper arms of both Bob's and Alice's interferometer.

Alice and Bob in the second half of the protocol publicly disclose the basis of the state sent and the measurement basis. From this they generate an initial amount of shared key material. However, as we mentioned above Eve, may have interfered with this and hence know some of the key. Also because there is a 50% chance that she measured in the wrong basis and hence transmitted the wrong state to Bob, there will not be agreement between all the bits in Alice's and Bob's shared key. Alice and Bob can, however, see if an eavesdropper has been present by selecting a subset of the key bits and publicly telling each other what they are. By checking the error rate (cases where their bits disagree) they can see the presence of the eavesdropper. If this error rate is not too high, Alice and Bob can perform classical privacy amplification on the remaining bits to create a secure shared key. We summarise this protocol in Fig. 11.6 and describe an implementation in Fig. 11.7.

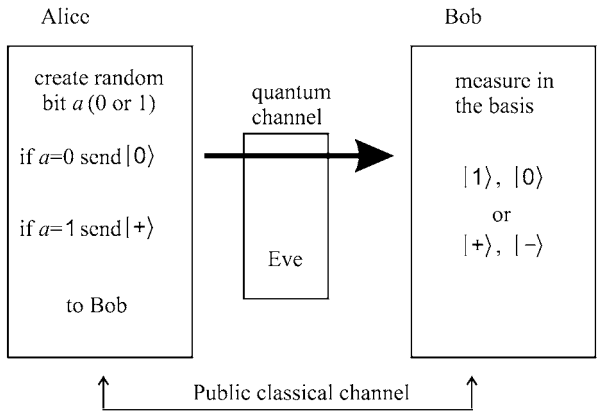


Fig. 11.8. Summary of the B92 protocol, which uses non-orthogonal bases.

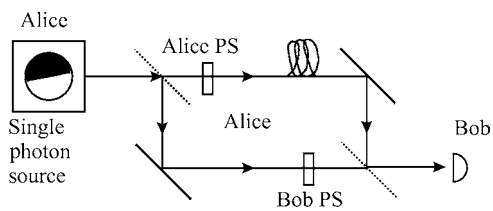


Fig. 11.9. Schematic representation of an optical realization of the B92 protocol: here the whole transmission between Bob and Alice takes place within the arms of an interferometer.

11.8.3 The B92 protocol

The BB84 protocol has been generalised to use other bases and states. One of the most well known is the B92 protocol [20], which uses two non-orthogonal states rather than 4. This protocol is established as follows. Alice creates a random bit a (either 0 or 1). She then transmits to Bob the qubit

$$|\psi\rangle = \begin{cases} |0\rangle\,, \text{if } a = 0, \\ |+\rangle\,, \text{if } a = 1. \end{cases} \tag{11.19}$$

Now Bob generates a random bit a' (again either 0 or 1). If the bit is zero, then Bob measures in the 0,1 basis. If the bit is 1, he measures in the $+$, $-$ basis. From Bob's measurement b he obtains either a zero or one, which he then announces over a public channel to Alice. Both Alice and Bob keep their random bits a and a' secret. Their initial key material is formed from the subset of results where the measurement result b was one. This occurs only if Alice had $a = 0$ and Bob $a' = 1 - a$ or vice versa with a probability of one half. The key is a for Alice and a' for Bob. Now because they wish to estimate the effects of the potential eavesdropper, they can use some of the key material to check for errors and then perform appropriate privacy amplification if the error rate is not too high. This then establishes secure shared key material. We summarize this protocol in Fig. 11.8 and show an implementation in Fig. 11.9.

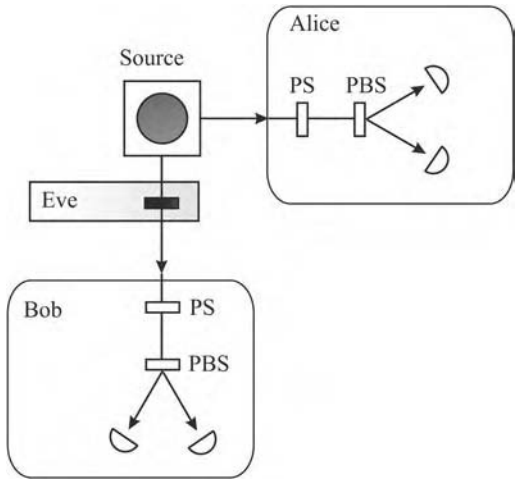


Fig. 11.10. Schematic representation of the Ekert protocol: the source emits correlated EPR pairs, one of which is transmitted to Alice and the other to Bob. Polarizing beam splitters (PBS) route the correlated photons to detectors.

11.8.4 The Ekert protocol

In the protocols we have discussed so far, we have relied on Alice sending single photons prepared in polarization or interferometric bases. Artur Ekert in 1991 [21] proposed a new quantum cryptographic protocol, which exploits quantum entanglement and Bell's inequality.

Suppose during a period of time, a quantum source emits two photons in an entangled state: the photons can be direction or polarization entangled, for example. We can use these in a secure communication system as we shall now describe. The basic idea of the Ekert protocol over the BB84 or B92 protocol consists of replacing the single quantum channel carrying qubits from Alice to Bob with a channel that carries two entangled qubits (one to Alice and one to Bob) from a central common source. We will describe the use of this novel quantum resource next.

Consider that Alice and Bob share a set of n entangled qubit pairs, where each pair is of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle\right). \tag{11.20}$$

The state on the left belongs to Alice and that on the right to Bob. Let us examine the establishment of a shared bit. Alice and Bob share the entangled pair (one qubit each); each now creates a random number (with values 0,1) independently, called, say, a (by Alice) and a' (by Bob). If Alice's random number is zero, she measures her qubit in the 0,1 basis, otherwise she uses the $+,-$ basis. She then records the measurement result b. Similarly Bob does the same thing with his qubit, and if Bob's random number was zero, he measures in the 0,1 basis, otherwise he uses the $+,-$ basis. He records the result b'. Alice and Bob now communicate their results (b, b') publicly. They keep the results only if $b = b'$ which establishes the key bit (a, a'). Performing this on the n entangled pairs establishes the required key. In Fig. 11.10, we sketch the layout of the Ekert protocol for a quantum cryptography system.

The experimental realizations of these schemes are reviewed at length by Gisin *et al*. [22].

11.9 Future prospects for quantum communication

Quantum physics, as we have described, offers the world of communications both opportunities and threats. Opportunities, from the fundamental nature of indivisible quanta and our inability to copy them, and threats because of the enormous potential of quantum computing provided the decoherence obstacle can be overcome so that large-scale quantum registers can be manipulated. We have concentrated here on secure key distribution for cryptography as it represents the most developed quantum technology. In QKD, classical information is transmitted using a quantum resource. However, quantum information itself can also be transmitted: for example, the quantum state of a system can be transferred from Alice to Bob if they share the quantum resource of entanglement. This is the basis of quantum teleportation as we have already discussed. This area has developed into one of the major growth fields of modern physics and is likely to generate new insights, new applications and new technologies over the coming decade.

11.10 Gates for quantum computation

A full discussion of quantum computing lies way beyond the scope of this book. But it does seem worthwhile nevertheless to discuss some aspects of quantum computing, in particular, the notion of quantum registers and quantum gates and their quantum optical realizations. This we shall do in what follows.

We begin with a discussion of quantum registers. The qubits we again represent in the generic basis $|0\rangle$ and $|1\rangle$, also known as the computational basis. The most general pure state for single qubit is the superposition state

$$|\psi_1\rangle = c_0\,|0\rangle + c_1\,|1\rangle\,, \quad |c_0|^2 + |c_1|^2 = 1. \tag{11.21}$$

A quantum register is a collection of qubits, say, N qubits. For example, the 3-qubit register in the state

$$|1\rangle \otimes |0\rangle \otimes |1\rangle \equiv |101\rangle = |5\rangle \tag{11.22}$$

provides a binary representation of the decimal number 5, $(5)_{10} = (101)_2$, as indicated on the right-hand side. But now suppose that the first qubit of the register is in the balanced superposition state $(|0\rangle + |1\rangle)/\sqrt{2}$. The state of the register is then

$$\begin{aligned} \frac{1}{\sqrt{2}}\,(|0\rangle + |1\rangle) \otimes |0\rangle \otimes |1\rangle \;&=\; \frac{1}{\sqrt{2}}\,(|001\rangle + |101\rangle) \\ &=\; \frac{1}{\sqrt{2}}\,(|1\rangle + |5\rangle)\,. \end{aligned} \tag{11.23}$$

This means that the 3-qubit quantum register *simultaneously* represents the decimal numbers 1 and 5. If all three qubits are in balanced superposition states, then the register is in the state

$$\begin{aligned}&\frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right)\otimes\frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right)\otimes\frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right)\\&\quad=\frac{1}{2^{3/2}}\left(|000\rangle+|001\rangle+|010\rangle+|011\rangle+|100\rangle+|101\rangle+|110\rangle+|111\rangle\right)\\&\quad=\frac{1}{2^{3/2}}\left(|0\rangle+|1\rangle+|2\rangle+|3\rangle+|4\rangle+|5\rangle+|6\rangle+|7\rangle\right),\end{aligned}\tag{11.24}$$

and thus a 3-qubit register can represent the eight decimal numbers 0–7 simultaneously. For any decimal number a given by $a = a_0 2^0 + a_1 2^1 + \cdots + a_{N-1} 2^{N-1}$, where the coefficients a_j, $j = 0, 1, \ldots N$, are restricted to 0 or 1, we can write

$$|a\rangle = |a_{N-1}\rangle \otimes |a_{N-2}\rangle \otimes \cdots \otimes |a_1\rangle \otimes |a_0\rangle \equiv |a_{N-1}a_{N-2}\ldots a_0\rangle\,. \tag{11.25}$$

The most general N-qubit quantum register state is

$$|\psi_N\rangle = \sum_{a=0}^{2^N-1} c_a\,|a\rangle. \tag{11.26}$$

With an N-qubit register, all the decimal numbers from 0 to $2^N - 1$, a total of 2^N numbers, can be represented simultaneously in a state of the form of Eq. (11.26).

To do any kind of processing, we need logic gates just as in a classical computer. The important difference between the quantum logic gates and classical ones is that the former are always reversible, in that the outputs uniquely depend on the inputs, which does not hold for classical gates. It holds for the quantum gates because they are implemented by unitary transformations. We begin with single-qubit gates and then proceed to 2-qubit gates.

The first gate we consider is the Hadamard gate. This is a gate that creates balanced superpositions of the basis states $|0\rangle$ and$|1\rangle$ and is represented by the letter H. With $x \in \{0, 1\}$, the Hadamard gate, described by the unitary operator $\hat{U}_{\rm H}$, effects the transformation

$$\hat{U}_{\rm H}\,|x\rangle = \frac{1}{\sqrt{2}}[(-1)^x\,|x\rangle + |1-x\rangle]. \tag{11.27}$$

Thus $\hat{U}_{\rm H}|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $\hat{U}_{\rm H}|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Note that the left-hand side of Eq. (11.24) may be obtained from Hadamard transformations on three bits each in a state $|0\rangle$. If the qubits are realized by the ground and excited states of a two-level atom, then the Hadamard transformation is just a $\pi/2$-pulse. The circuit symbol for the Hadamard gate is given in Fig. 11.11(a).

A second single-qubit gate, the phase gate represented by $\hat{U}_{\rm PG}(\phi)$, effects the transformation

$$\hat{U}_{\rm PG}\,(\phi)\,|x\rangle = e^{i\,x\,\phi}\,|x\rangle \tag{11.28}$$

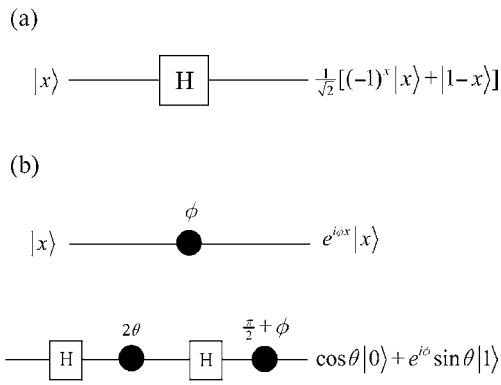


Fig. 11.11. (a) Circuit symbol of a Hadamard gate. (b) Circuit symbol of the 1-qubit phase gate.

Fig. 11.12. Circuit for the generation of the state of Eq. (11.29) using phase and Hadamard gates.

and thus $\hat{U}_{\rm PG}(\phi)|0\rangle = |0\rangle$ and $\hat{U}_{\rm PG}(\phi)|1\rangle = e^{i\,\phi}|1\rangle$. The circuit symbol for this gate is given in Fig.11.11(b). It turns out that a combination of Hadamard and phase gates can produce various superposition states of single qubit. The sequence

$$\hat{U}_{\rm PG}\left(\phi+\frac{\pi}{2}\right)\hat{U}_{\rm H}\hat{U}_{\rm PG}\,(2\theta)\,\hat{U}_{\rm H}\,|0\rangle = \cos\theta\,|0\rangle + e^{i\,\phi}\sin\theta\,|1\rangle\,, \tag{11.29}$$

as represented in Fig. 11.12, yields the most general single-qubit pure state.

So far, we have not encountered registers in entangled states. One possible entangled 2-qubit register state is $(|00\rangle + |11\rangle)/\sqrt{2}$. How can these, as a matter of principle, be made? We have already encountered ways of generating entangled states in the earlier parts of this book through some very specific interactions. At the moment we are only interested in formulating a rather general procedure, without regard to any specific realization. Thus we introduce a 2-qubit gate called the controlled not (C-NOT) denoted by the unitary operator $\hat{U}_{\text{C-NOT}}$. If we let the first qubit be the control qubit and the second be the target, then we can represent the action of this gate symbolically by

$$\hat{U}_{\text{C-NOT}}\,|x\rangle\,|y\rangle = |x\rangle\,|\mathrm{mod}_2\,(x+y)\rangle\,, \tag{11.30}$$

where $x, y \in \{0, 1\}$. The gate is represented diagrammatically by the circuit in Fig. 11.13(a). If the control is in the state $|0\rangle$ and the target also in state $|0\rangle$, the target is unchanged. But if the control is in $|1\rangle$ then the target "flips" from $|0\rangle$ to $|1\rangle$. But suppose the control qubit is prepared, by the Hadamard transformation, in the superposition state $(|0\rangle + |1\rangle)/\sqrt{2}$. Then we shall have, for the target initially in $|0\rangle$ and $|1\rangle$ respectively,

$$\begin{aligned}
\hat{U}_{\text{C-NOT}}\hat{U}_{\text{H1}}\,|0\rangle\,|0\rangle &= \hat{U}_{\text{C-NOT}}\frac{1}{\sqrt{2}}\,(|0\rangle+|1\rangle)\,|0\rangle = \frac{1}{\sqrt{2}}\,(|00\rangle+|11\rangle)\,,\\
\hat{U}_{\text{C-NOT}}\hat{U}_{\text{H1}}\,|0\rangle\,|1\rangle &= \hat{U}_{\text{C-NOT}}\frac{1}{\sqrt{2}}\,(|0\rangle+|1\rangle)\,|1\rangle = \frac{1}{\sqrt{2}}\,(|01\rangle+|10\rangle)\,,
\end{aligned} \tag{11.31}$$

where the H1 means that the Hadamard gate acts only on the first qubit. In both cases we end up with entangled states. The entire process, for the target initially in the state $|0\rangle$, is represented by the diagram in Fig. 11.13(b).

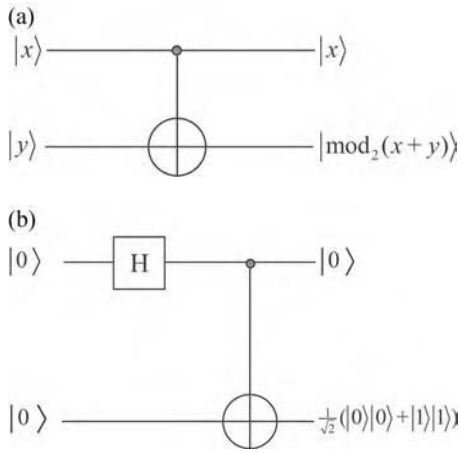


Fig. 11.13. (a) Circuit symbol for a C-NOT. (b) Circuit for creating a 2-qubit entangled state using a Hadamard gate and a C-NOT.

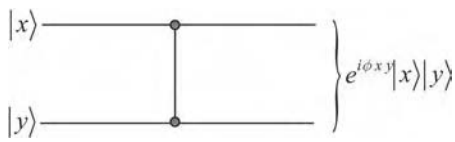


Fig. 11.14. Circuit symbol for a controlled phase gate.

Another 2-qubit gate is the controlled phase gate $\hat{U}_{\mathrm{CPG}}(\phi)$ acting according to

$$\hat{U}_{\mathrm{CPG}}(\phi)\,|x\rangle\,|y\rangle = e^{i\,\phi\,x\,y}\,|x\rangle\,|y\rangle\,, \tag{11.32}$$

represented graphically in Fig. 11.14. The phase factor $e^{i\,\phi}$ appears only when $x = 1 = y$.

Out of these gates, essentially any quantum computational algorithm can be implemented with the proper concatenation. Networks can be devised to bring about certain kinds of output. One, very simple, circuit has already been discussed and others will be addressed in the problems.

A function in a computer is a mapping from one set of numbers to another:

$$f : \{0, 1, \ldots, 2^m - 1\} \to \{0, 1, \ldots, 2^n - 1\}\,, \tag{11.33}$$

where m and n are integers. A classical computer evaluates the function at each input $0, 1, \ldots, 2^m - 1$, to obtain the set of numbers $f(0), f(1), \ldots, f(2^m - 1)$. But the computation is generally irreversible and thus it is not possible to associate a given output with a particular input. Quantum mechanically, it is not possible to evolve an input register $|x\rangle$ into an output register $|f(x)\rangle$ by a unitary transformation. However, by using two registers, one to contain the output and another to retain the input, one can implement the quantum-mechanical evaluation of a function as a unitary transformation according to

$$\hat{U}_f\,|x\rangle\,|0\rangle = |x\rangle\,|f\,(x)\rangle\,, \tag{11.34}$$

where different input and output states will be orthogonal: $\langle x|x'\rangle = \delta_{x\,x'}$, $\langle f|f'\rangle = \delta_{f\,f'}$.

As an example of a quantum computation, simple enough to describe here, yet able to illustrate the potential power of a quantum computer taking full advantage of quantum entanglement, we now describe the quantum algorithm known as Deutsch's problem [15], the first quantum algorithm proposed, and one of the simplest. The problem proposed by Deutsch is this. Consider a function $f: \{0, 1\} \to \{0, 1\}$. There are only four *possible* outcomes and these are $f(0) = 0$, $f(0) = 1, f(1) = 0$, and $f(1) = 1$. Given an unknown function f, determine, *with only one operation*, whether or not the function is constant, meaning that $f(0) = f(1)$, or that it is varying, $f(0) \neq f(1)$. The quantum computer will be required to implement the transformation

$$|x\rangle\,|y\rangle \longrightarrow |x\rangle\,|\mathrm{mod}_2\,(y + f\,(x))\rangle\,, \tag{11.35}$$

where $|x\rangle$ is the input qubit and $|y\rangle$ represents the qubit of the quantum computer hardware. The idea is to prepare these qubits, using Hadamard gates, in the product state

$$\begin{aligned} |\Psi_{\mathrm{in}}\rangle &= |x\rangle\,|y\rangle = \frac{1}{\sqrt{2}}\,(|0\rangle + |1\rangle)\,\frac{1}{\sqrt{2}}\,(|0\rangle - |1\rangle) \\ &= \frac{1}{2}\,(|0\rangle\,|0\rangle - |0\rangle\,|1\rangle + |1\rangle\,|0\rangle - |1\rangle\,|1\rangle)\,. \end{aligned} \tag{11.36}$$

The minus signs in this last equation are important. If we now allow this state to undergo the transformation of Eq. (11.35) we obtain the output state

$$\begin{aligned} |\Psi_{\mathrm{out}}\rangle &= \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|\bar{f}(0)\rangle + |1\rangle|f(1)\rangle - |1\rangle|\bar{f}(1)\rangle), \\ &= \frac{1}{2}[|0\rangle(|f(0)\rangle - |\bar{f}(0)\rangle) + |1\rangle(|f(1)\rangle - |\bar{f}(1)\rangle)], \end{aligned} \tag{11.37}$$

where the bar means to invert: $\bar{0} = 1$, $\bar{1} = 0$. Notice that Eq. (11.37) is the result of simultaneous, parallel, processing on both states of the first qubit in the superposition $|x\rangle \sim |0\rangle + |1\rangle$. This is a display for the potential power of a quantum computer: the possibility of massive parallel processing. If f is a constant function, then the output is

$$|\Psi_{\mathrm{out}}\rangle_{f\,\mathrm{const}} = \frac{1}{2}[(|0\rangle + |1\rangle)(|f(0)\rangle - |\bar{f}(0)\rangle)], \tag{11.38}$$

whereas if f is not constant

$$|\Psi_{\mathrm{out}}\rangle_{f\,\mathrm{not\ const}} = \frac{1}{2}[(|0\rangle - |1\rangle)(|f(0)\rangle - |\bar{f}(0)\rangle)]. \tag{11.39}$$

Note that in the two cases the first qubit states are orthogonal. If we apply a Hadamard transformation to this first qubit we have

$$\begin{aligned} \hat{U}_{\mathrm{H1}}|\Psi_{\mathrm{out}}\rangle_{f\mathrm{const}} &= |0\rangle(|f(0)\rangle - |\bar{f}(0)\rangle), \\ \hat{U}_{\mathrm{H1}}|\Psi_{\mathrm{out}}\rangle_{f\mathrm{not\ const}} &= |1\rangle(|f(0)\rangle - |\bar{f}(0)\rangle). \end{aligned} \tag{11.40}$$

Thus, a single measurement on the first qubit determines whether or not f is constant.

The Deutsch algorithm is, obviously, very simple. A slight generalization of this algorithm by Deutsch and Jozsa [23] has been implemented in an ion-trap quantum computer containing two trapped ions.

A much more involved algorithm is that of Shor's for factoring large numbers into primes [17]. We shall not review that algorithm here but refer to the reader to specialized reviews of quantum computing. We point out again though that Shor's algorithm has been implemented, on a nuclear magnetic resonance (NMR) quantum computer [18], in order to find the prime factors of $15 = 3 \times 5$, a demonstration of the algorithm and obviously not a demonstration of the power of a quantum computer. Unfortunately, the NMR approach appears not to be scalable. Another important quantum algorithm is Grover's search algorithm [24].

11.11 An optical realization of some quantum gates

There have been many proposals for realizing the quantum gates in quantum optical systems such as polarized photons, cavity QED, and in systems of trapped ions. Here we shall discuss all-optical realizations of some gates based on beam splitters, Kerr interactions, and optical phase shifters. Two photonic modes are required and these form the basis of a "dual-rail" rail realization of a quantum computer discussed by Chuang and Yamamoto [25]. Only single photon states are involved and, to make clear the mapping of the dual photon states onto qubit bases $|0\rangle$ and $|1\rangle$, we start with the realization of the Hadamard gate in terms of a 50:50 beam splitter, as pictured in Fig. 11.15(a). The input mode operators we take as $\hat{a}$ and $\hat{b}$ and the beam splitter is chosen to be of the type whose output mode operators $\hat{a}'$ and $\hat{b}'$ are related to the input ones according to the transformation

$$\hat{a}' = \frac{1}{\sqrt{2}}(\hat{a} + \hat{b}), \ \hat{b}' = \frac{1}{\sqrt{2}}(\hat{a} - \hat{b}) \tag{11.41}$$

or, when inverted,

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{a}' + \hat{b}'), \hat{b} = \frac{1}{\sqrt{2}}(\hat{a}' - \hat{b}'). \tag{11.42}$$

We assign the computational state $|0\rangle$ to the input product state $|0\rangle_a|1\rangle_b$, that is, we take $|0\rangle \equiv |0\rangle_a|1\rangle_b$, and similarly we take $|1\rangle = |1\rangle_a|0\rangle_b$. From the rules developed in Chapter 6, and letting $\hat{U}_{\rm BS}$ represent the beam-splitter transformation operator, it follows that

$$\begin{aligned}
\hat{U}_{\rm BS}|0\rangle &= \hat{U}_{\rm BS}|0\rangle_a|1\rangle_b \\
&= \frac{1}{\sqrt{2}}(|0\rangle_a|1\rangle_b + |1\rangle_a|0\rangle_b) = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\
\hat{U}_{\rm BS}|1\rangle &= \hat{U}_{\rm BS}|1\rangle_a|0\rangle_b \\
&= \frac{1}{\sqrt{2}}(|0\rangle_a|1\rangle_b - |1\rangle_a|0\rangle_b) = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),
\end{aligned} \tag{11.43}$$

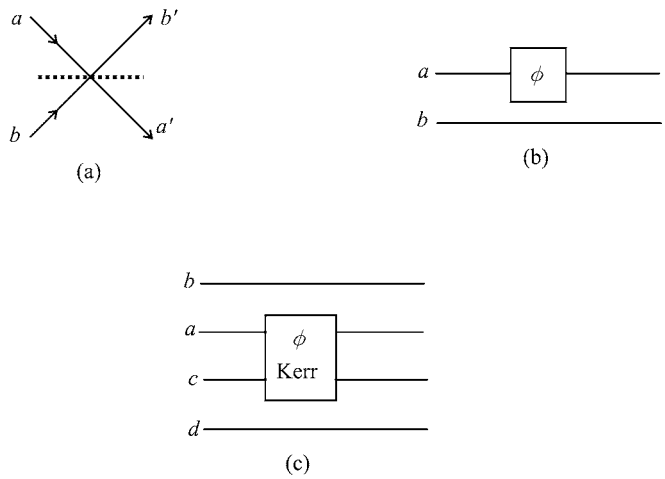


Fig. 11.15. (a) Designation of the input and output modes of the a 50:50 beam splitter in a dual-rail optical implementaion of the Hadamard transformation. (b) Optical dual-rail implementation of a phase gate. (c) Optical implementation of a dual-rail 2-qubit controlled phase gate.

where we have dropped the primes for the states after the beam splitter. Clearly, this device provides a realization of the Hadamard gate. We wish to make it clear that, despite the appearance of two photonic modes containing either the vacuum or a single photon, those are not the basis of the qubits. It is the product of the states of the two modes that map onto the computational basis. We are still dealing with a 1-qubit gate.

The single-qubit phase gate can be realized by placing a phase shifter in the a-beam, the operator for this device being $\hat{U}_{\rm PG} = \exp(i\,\phi\,\hat{a}^\dagger\hat{a})$. It is easy to see that

$$\begin{aligned}\hat{U}_{\rm PG}\,|0\rangle &= \exp(i\,\phi\,\hat{a}^\dagger\hat{a})\,|0\rangle_a\,|1\rangle_b = |0\rangle\,,\\ \hat{U}_{\rm PG}\,|1\rangle &= \exp(i\,\phi\,\hat{a}^\dagger\hat{a})\,|1\rangle_a\,|0\rangle_b = e^{i\,\phi}\,|1\rangle\,.\end{aligned} \tag{11.44}$$

The device is represented graphically in Fig.11.15(b).

For 2-qubit gates we must have two sets of two photonic modes, say (a, b) and (c, d). Following the convention we having been using, we have the identifications

$$\begin{aligned}|0\rangle\,|0\rangle &= |0\rangle_a\,|1\rangle_b\,|0\rangle_c\,|1\rangle_d\,,\\ |0\rangle\,|1\rangle &= |0\rangle_a\,|1\rangle_b\,|1\rangle_c\,|0\rangle_d\,,\\ |1\rangle\,|0\rangle &= |1\rangle_a\,|0\rangle_b\,|0\rangle_c\,|1\rangle_d\,,\\ |1\rangle\,|1\rangle &= |1\rangle_a\,|0\rangle_b\,|1\rangle_c\,|0\rangle_d\,.\end{aligned} \tag{11.45}$$

The controlled phase gate (CPG) can, in principle, be realized by the cross-Kerr interaction between the modes a and c as indicated in Fig. 11.15(c). The interaction is represented by the operator $\hat{U}_{\rm CPG} = \exp(i\,\phi\,\hat{a}^\dagger\hat{a}\,\hat{c}^\dagger\hat{c})$. It is easy to verify that, for 2-qubit states, only the state $|1\rangle|1\rangle$ acquires the phase shift $e^{i\,\phi}$. However, there is a problem in generating arbitrary, especially large, phase shifts in this manner. Large phase shifts ϕ will require high $\chi^{(3)}$ nonlinear susceptibilities which are generally not available. On the other hand, as we have previously

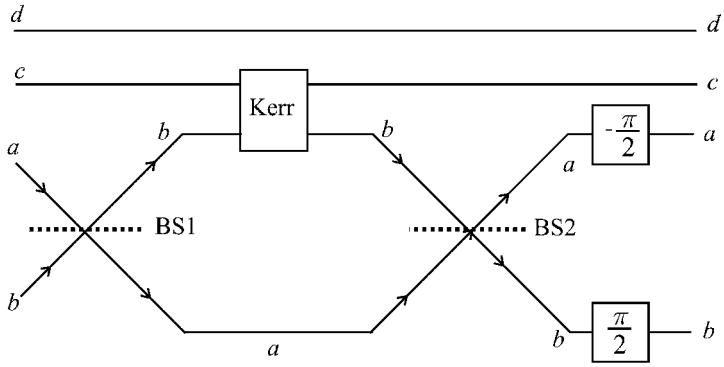


Fig. 11.16. An optical realization of a C-NOT using a Fredkin gate. The state modes (*a, b*) and (*c, d*) represent the two qubits. The phase shifters in the outputs of the Mach–Zehnder interferometer are to adjust the phase factors of the outputs.

mentioned, techniques to generate giant Kerr susceptibilities with electromagnetically induced transparency are actively being pursued.

Lastly, we discuss a possible optical realization of the C-NOT gate. This can be done, at least in principle, with a Mach–Zehnder interferometer coupled to an external mode, which we take to be the c-mode as pictured in Fig. 11.16, via a cross-Kerr interaction. The Kerr "evolution" is described by the unitary operator

$$\hat{U}_{\text{Kerr}}(\eta) = \exp(i\,\eta\,\hat{b}^\dagger\hat{b}\,\hat{d}^\dagger\hat{d}), \tag{11.46}$$

where the parameter η is proportional to the nonlinear susceptibility $\chi^{(3)}$. The first beam splitter will be taken as a type described by the transformatiom

$$\hat{U}_{\text{BS1}} = \exp\Big[i\frac{\pi}{4}(\hat{a}^\dagger\hat{b} + \hat{a}\,\hat{b}^\dagger)\Big], \tag{11.47}$$

and the second beam splitter we take to be conjugate to the first such that $\hat{U}_{\text{BS2}} = \hat{U}^\dagger_{\text{BS1}}$. Thus the operator describing the evolution just after the second beam splitter must be given by

$$\hat{U}_{\text{F}}(\eta) = \hat{U}^\dagger_{\text{BS1}}\hat{U}_{\text{Kerr}}(\eta)\,\hat{U}_{\text{BS1}}, \tag{11.48}$$

where the label F stands for Fredkin. The gadget represented by this operator is the quantum optical form of what is known as the Fredkin [26] gate. A quantum optical realization of the Fredkin gate was first discussed by Milburn [27]. After a bit of algebra, which we leave to the reader, we arrive at

$$\hat{U}_{\text{F}}(\eta) = \exp\Big[i\frac{\eta}{2}\hat{c}^\dagger\hat{c}(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b})\Big]\exp\Big[\frac{\eta}{2}\hat{c}^\dagger\hat{c}(\hat{a}^\dagger\hat{b} - \hat{b}^\dagger\hat{a})\Big]. \tag{11.49}$$

The first term will just yield a phase factor depending on whether or not mode c contains a photon. The second term though is a conditional transformation. If the c-mode is in the vacuum, no transformation will occur. But if there is a single photon in the c-mode, a transformation must occur. For our purposes we must take $\eta = \pi$. Using the identifications in Eq. (11.45), and taking the qubits of the

c–d modes to be the control and those of the a–b as the target, reversed in order compared with the above discussion, we can easily show that

$$\begin{aligned}
\hat{U}_{\mathrm{F}}(\eta)\,|0\rangle\,|0\rangle &= |0\rangle\,|0\rangle\,,\\
\hat{U}_{\mathrm{F}}(\eta)\,|1\rangle\,|0\rangle &= |1\rangle\,|0\rangle\,,\\
\hat{U}_{\mathrm{F}}(\eta)\,|0\rangle\,|1\rangle &= i\,|1\rangle\,|1\rangle\,,\\
\hat{U}_{\mathrm{F}}(\eta)\,|1\rangle\,|1\rangle &= -i\,|0\rangle\,|1\rangle\,,
\end{aligned} \tag{11.50}$$

where we have used the transformation

$$\exp\Big[\frac{\pi}{2}(\hat{a}^\dagger\hat{b}-\hat{b}^\dagger\hat{a})\Big]\begin{cases}|0\rangle_a\,|1\rangle_b = i\,|1\rangle_a\,|0\rangle_b\,,\\ |1\rangle_a\,|0\rangle_b = -i\,|0\rangle_a\,|1\rangle_b\,,\end{cases} \tag{11.51}$$

as the reader can verify. To remove the factors i and $-i$ we can, if necessary, insert appropriate phase shifters into the output beams. Thus the unitary operator describing a quantum optical C-NOT is

$$\hat{U}_{\text{C-NOT}} = \exp\Big(-i\frac{\pi}{2}\hat{a}^\dagger\hat{a}\Big)\exp\Big(i\frac{\pi}{2}\hat{b}^\dagger\hat{b}\Big)\,\hat{U}_{\mathrm{F}}(\pi)\,. \tag{11.52}$$

Unfortunately, as we have already remarked, Kerr media with large nonlinearities would be required to realize these gates. Recently, a C-NOT gate has been realized experimentally with polarized photons [28]. Knill *et al*. [29] have proposed a different approach using linear optics but where state reduction is an essential ingredient during the computation, not just at the outputs.

There are many other possible interactions that can be used to make a quantum C-NOT gate, including cavity QED [30], laser-cooled trapped ions [31], atoms in optical lattices interacting via collisions [32], and so on. All of them have the feature that the interaction of one qubit with another is conditional on the state of the partner qubit. For example, if an atom in an optical lattice trap is brought close to a second atom in an adjacent well of the lattice, then the energy levels of the valence electrons in each of the atoms are shifted owing to the fluctuations of the dipole moment of the other atom. The shifts are greater for higher levels. Thus an electron in an excited state picks up phase more slowly than one in the ground state. This conditional phase shift (i.e. the amount of phase that depends on the state of the control electron) forms the basis of this kind of conditional state change performed by a quantum C-NOT gate.

The atoms must be controlled very carefully, though. It turns out that the fidelity of gate operation has to be about one part in a thousand or better in order to make a quantum computer realistic. This is but one of the hurdles to harnessing the undoubted quantum potential.

11.12 Decoherence and quantum error correction

Quantum coherence is fragile. For the reasons we have discussed in Chapter 8 and elsewhere, it is difficult to maintain superposition states of many particles

in which each particle may be physically separated from all the others. Entanglement is delicate. The reason for this is that all systems, quantum or classical, are not isolated. They interact with everything around them – local fluctuating electromagnetic fields, the presence of impurity ions, coupling of unobserved degrees of freedom to the system containing the qubit, etc. These fluctuations destroy quantum interference. A simple, but useful, analogy is the interference of optical waves in Young's double-slit experiment. In that apparatus, waves from two spatially separated portions of a beam are brought together. If the two parts of the beam have the same phase, then the fringe pattern remains stable. But if the phase of one part of the beam is drifting with respect to the other, then the fringe pattern will be washed out. And the more slits you have in the screen, the lower the visibility for the same amount of phase randomization per pair of slits.

It might appear that the situation is hopeless: how could one hope to control the phases and amplitudes of superpositions involving many qubits? But, amazingly, quantum mechanics provides a way to solve this problem, through even higher levels of entanglement. In classical information processing, inevitable environmental noise is dealt with by error correction. In its simplest form, this involves repeating the message transmission or calculation until a majority result is obtained. But there are more efficient ways, for example through the use of a parity check on a block of bits. A similar notion of encoding logical "1"s and "0"s into blocks of qubits and performing a parity check to reveal errors can be applied to a quantum register [33]. But there is one difficulty: if you measure the register, you would destroy the superposition state encoded in it. So how do we determine what might be wrong with the register qubits without looking at them? Simple: entangle them with an ancillary register, and measure the ancilla! Because the two registers are correlated, the results of the measurement of the ancilla will tell you how to fix the errors in the processing register, without destroying any coherent superpositions in the processing register itself.

Another way to retain register coherence is to know a little about the sort of noise that is acting on it [34]. If the noise has some very slow components (or those with very long wavelengths), then it is sometimes possible to find certain combinations of qubit states for which the noise on one qubit exactly cancels the noise on another. These qubit states live in a "decoherence free subspace" or DFS [35]. If you can use only computational states that lie in this DFS, then your computer will be immune to environmental perturbations.

It is the possibility of combining these tools to combat noise that leads researchers to believe that a quantum computer can be built even though decoherence lurks around every corner.

Problems

1. Given the output state in Eq. (11.2) of the Mach–Zehnder interferometer of Fig. 11.2 with twin single photons as inputs, verify Eq. (11.4).

2. Suppose that the inputs to the Mach–Zehnder interferometer are the twin two-photon states $|2\rangle_a|2\rangle_b$. Obtain the uncertainty of the phase measurement in this case.
3. Suppose the first beam splitter of the Mach–Zehnder interferometer is replaced by a "magical" device that can somehow produce the maximally entangled state of Eq. (11.6). Show that for this state, the uncertainty in the measurement of phase is exactly $\Delta\theta = 1/N$ for all N. Caution: you must incorporate the effect of the phase shifter in the interferometer and you must take into account the second beam splitter.
4. Verify the statement made in Section 11.3 to the effect that teleporting a state with the shared statistical mixture
$$\hat{\rho}_{\mathrm{AB}} = \frac{1}{2}\left(|0\rangle_{\mathrm{A}}|0\rangle_{\mathrm{B}}\,{}_{\mathrm{A}}\langle 0|\,{}_{\mathrm{B}}\langle 0| + |1\rangle_{\mathrm{A}}\,|1\rangle_{\mathrm{B}}\,{}_{\mathrm{A}}\langle 1|\,{}_{\mathrm{B}}\langle 1|\right)$$
is impossible.
5. Suppose that Alice, Bob, and Claire each share one of the "particles" A, B, or C prepared in the three-particle entangled state
$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{\mathrm{A}}\,|0\rangle_{\mathrm{B}}\,|0\rangle_{\mathrm{C}} - |1\rangle_{\mathrm{A}}\,|1\rangle_{\mathrm{B}}\,|1\rangle_{\mathrm{C}}\right)$$
as indicated by the labels. Suppose Alice has the unknown state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$. Investigate the prospects for teleporting the unknown state to both Bob and Claire.
6. Show that the unitary operator for the Hadamard gate can be represented as
$$\hat{U}_{\mathrm{H}} = \frac{1}{\sqrt{2}}\left[|0\rangle\,\langle 0| + |1\rangle\,\langle 0| + |0\rangle\,\langle 1| - |1\rangle\,\langle 1|\right].$$
7. The 1-qubit operator $\hat{X} = |0\rangle\langle 1| + |1\rangle\langle 0|$ represents a logical NOT gate (also known as an inverter) because $\hat{X}|0\rangle = |1\rangle$and $\hat{X}|1\rangle = |0\rangle$. Show that the controlled-not (C-NOT) gate is represented by the 2-qubit operator
$$\hat{U}_{\text{C-NOT}} = |0\rangle\langle 0| \otimes \hat{I} + |1\rangle\langle 1| \otimes \hat{X},$$
where $\hat{I} = |0\rangle\langle 0| + |1\rangle\langle 1|$ is the identity operator and where the first qubit is the control.
8. The 3-qubit gate known as the Toffoli gate is given by
$$\hat{U}_{\mathrm{TG}}\,|x_1\rangle\,|x_2\rangle\,|y\rangle = |x_1\rangle\,|x_2\rangle\,|\mathrm{mod}_2\,(x_1x_2 + y)\rangle\,.$$
Show that this gate is a controlled-controlled-not gate. Draw a circuit diagram to represent the gate. Find the representation of the operator $\hat{U}_{\mathrm{TG}}$ in terms of the projection and inversion operator of the three qubits.
9. Design an optical realization of the Toffoli gate described in the previous problem.
10. Let $\hat{U}_{\mathrm{Cl}}$ represent a 2-qubit operator that "clones" any quantum state according to $\hat{U}_{\mathrm{Cl}}|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$. Show that such an operator does not, in fact, exist.

References

[1] G. E. Moore, *Electronics*, **38** (1965), April 19 issue. This paper can be found at www.intel.com/research/silicon/mooreslaw/htm

[2] Reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press, 1987), p. 139.
[3] J. J. Bollinger, W. M. Itano, D. J. Wineland and D. J. Heinzen, *Phys. Rev. A*, **54** (1996), R4649; C. C. Gerry, *Phys. Rev. A*, **61** (2000), 043811.
[4] W. Heitler, *The Quantum Theory of Radiation*, 3rd edition (London: Oxford University Press, 1954).
[5] P. Kok, H. Lee and J. P. Dowling, *Phys. Rev*. *A*, **65** (2002), 052104.
[6] C. C. Gerry and R. A. Campos, *Phys. Rev. A*, **64** (2001), 063814; C. C. Gerry, A. Benmoussa and R. A. Campos, *Phys. Rev. A*, **66** (2002), 013804.
[7] See, for example, K. Banaszek and I. A. Walmsley, *Opt. Lett.* **28** (2003), 52; M. J. Fitch, B. C. Jacobs, T. B. Pittman and J. D. Franson, *Phys. Rev. A*, **68** (2003), 043814.
[8] M. J. Holland and K. Burnett, *Phys. Rev. Lett.*, **71** (1993), 1355.
[9] R. A. Campos, C. C. Gerry and A. Benmoussa, *Phys. Rev. A*, **68** (2003), 023810.
[10] C. H. Bennett, G. Brassrad, C. Crepeau, R. Jozsa, A. Peres and W. K. Wooters, *Phys. Rev. Lett.*, **70** (1993), 1895.
[11] D. Boschi, S. Branca, F. De Martini and L. Hardy, *Phys. Rev. Lett.*, **80** (1998), 1121.
[12] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger, *Nature*, **390** (1997), 575.
[13] A. Furuwasa, J. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble and E. S. Polzik, *Science*, **282** (1998), 706.
[14] R. Rivest, A. Shamir and L. Adleman, "On Digital Signatures and Public-Key Cryptosystems", *MIT Laboratory for Computer Science Technical Report*, MIT/LCS/TR-212 (January 1979).
[15] D. Deutsch, *Proc. R. Soc. Lond.* A **400** (1985), 97.
[16] D. Deutsch and R. Jozsa, *Proc. R. Soc. Lond.* A **439** (1992), 553.
[17] P. W. Shor, in *Proceedings of the 35th Symposium on Foundations of Computer Science, Los Alamitos*, edited by S. Goldwasser (1994) (IEEE Computer Society Press), p. 124.
[18] L. M. K Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood and I. L. Chuang, *Nature*, **414** (2001), 883.
[19] C. H. Bennett and G. Brassard, in *Proc. of IEEE Conference on Computers, Systems and Signal Processing* (New York: IEEE Press, 1984). See also C. H. Bennett. F. Besette, G. Brassard, L. Salvail and J. Smolin, *J. Cryptol*., **5** (1992), 3.
[20] C. H. Bennett, *Phys. Rev. Lett.*, **68** (1992), 3121.
[21] A. K. Ekert, *Phys. Rev. Lett*., **67** (1991), 661.
[22] N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, *Rev. Mod. Phys*., **74** (2002), 145.
[23] S. Guide, M. Riebe, G. P. T. Lancaster, C. Becher, J. Eschner, H. Häffner, F. Schmidt-Kaler, I. L. Chuang and R. Blatt, *Nature*, **421** (2003), 48.
[24] L. K. Grover, *Phys. Rev. Lett*., **79** (1997), 325.
[25] I. L. Chuang and Y. Yamamoto, *Phys. Rev. A*, **52** (1995), 3489.
[26] E. Fredkin and T. Toffoli, *Int. J. Theor. Phys.*, **21** (1982), 219.
[27] G. J. Milburn, *Phys. Rev. A*, **62** (1989), 2124.
[28] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph and D. Branning, *Nature*, **426** (2003), 264.
[29] E. Knill, R. Laflamme and G. Milburn, *Nature*, **409** (2001), 46.

[30] See T. Sleator and H. Weinfurter, *Phys. Rev. Lett.*, **74** (1995), 4087; P. Domokos, J. M. Raimond, M. Brune and S. Haroche, *Phys. Rev. A*, **52** (1995), 3554.
[31] See C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano and D. J. Wineland, *Phys. Rev. Lett.*, **75** (1995), 4714.
[32] See D. R. Meacher, *Contemp. Phys.*, **39** (1998), 329; I. H. Deutsch and G. K. Brennen, *Forts. Phys.*, **48** (2000), 925.
[33] A. Calderbank and P. Shor, *Phys. Rev. A*, **52** (1995), R2493; *Phys. Rev. A*, **54** (1996), 1098; A. M. Steane, *Phys. Rev. Lett*., **77** (1995), 793.
[34] G. M. Palma, K.-A. Suominen and A. K. Ekert, *Proc. R. Soc. Lond*., A **452** (1996), 567.
[35] D. Lidar, I. L. Chuang and B. Whaley, *Phys. Rev. Lett*., **81** (1998), 2594.

Bibliography

A recent book covering many aspects of cryptography is

P. Garrett, *Making and Breaking Codes: An Introduction to Cryptology* (Upper Saddle River: Prentice Hall, 2001).

A fun book to read on the history of cryptography is

S. Singh, *The Code Book: The Science of Secrecy from Ancient Eygpt to Quantum Cryptography* (New York: Anchor Books, 1999).

Useful books that have been published on all aspects of quantum information are

H.-K. Lo, S. Popescu and T. Spliller (editors), *Introduction to Quantum Computation and Quantum Information* (Singapore: World Scientific, 1998).
D. Bouwmeester, A. K. Ekert and A. Zeilinger (editors), *The Physics of Quantum Information* (Berlin: Springer, 2000).
M. Nielsen and I. L. Chuang, *Quantum Information and Quantum Computation* (Cambridge: Cambridge University Press, 2001).
S. L. Braunstein and H.-K. Lo, *Scalable Quantum Computers* (Berlin: Wiley-VCH, 2001).

Below we list articles that we have found to be accessible and with many references to the original literature.

A. Barenco, "Quantum physics and computers", *Contemp. Phys*., **37** (1996), 375.
D. P. DiVincenzo, "Quantum Computation", *Science*, **270** (1995), 255.
R. J. Hughes, D. M. Alde, P. Dyer, G. G. Luther, G. L. Morgan and M. Schauer, "Quantum cryptography", *Contemp. Phys*., **36** (1995), 149.
S. J. D. Phoenix and P. D. Townsend, "Quantum cryptography: how to beat the code breakers using quantum mechanics", *Contemp. Phys*., **36** (1995), 165.
M. B. Plenio and V. Vedral, "Teleportation, entanglement and thermodynamics in the quantum world", *Contemp. Phys*., **39** (1998), 431.

Appendix A

The density operator, entangled states, the Schmidt decomposition, and the von Neumann entropy

A.1 The density operator

Quantum-mechanical state vectors $|\psi\rangle$ convey the maximal amount of information about a system allowed by the laws of quantum mechanics. Typically, the information consists of quantum numbers associated with a set of commuting observables. Futhermore, if $|\psi_1\rangle$ and $|\psi_2\rangle$ are two possible quantum states then so is their coherent superposition

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle \tag{A1}$$

if the coefficients c_1 and c_2 are known. If the states $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal ($\langle\psi_2 \mid \psi_1\rangle = 0$) then we must have $|c_1|^2 + |c_2|^2 = 1$. But there are frequently, in fact, more often than not, situations where the state vector is not precisely known. There are, for example, cases where the system of interest is interacting with some other system, possibly a very large system, e.g. a reservoir, with which it becomes entangled. It may be possible to write state vectors for the multicomponent system but not for the subsystem of interest. For example, for a system of two spin-1/2 particles with, the eigenstates of, say, the z-component of the spin denoted $|\uparrow\rangle$ for spin up and $|\downarrow\rangle$ for spin down, a possible state vector of the combined system is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2\right], \tag{A2}$$

the so-called singlet state (total angular momentum zero), also known as one of the "Bell" states. Equation (A2) is an example of an entangled state. An entangled state cannot be factored, in any basis, into a product of states of the two subsystems, i.e.

$$|\psi\rangle \neq |\text{spin } 1\rangle\, |\text{spin } 2\rangle \quad \text{(for an entangled two-spin state)}\,. \tag{A3}$$

Entanglement is, apart from the superposition principle itself, an essential mystery of quantum mechanics, as was pointed out in 1935 by Schrödinger himself. Note though that entanglement follows from the superposition principle and is not something imposed on the theory. So Feynman's dictum that the superposition principle contains "the only mystery" is still correct.

Quantum states described by state vectors are said to be *pure* states. States that cannot be described by state vectors are said to be in *mixed* states. Mixed states are described by the density operator

$$\hat{\rho} = \sum_i |\psi_i\rangle\, p_i\, \langle\psi_i| = \sum_i p_i\, |\psi_i\rangle\, \langle\psi_i|, \tag{A4}$$

where the sum is over an ensemble (in the sense of statistical mechanics) where p_i is the probability of the system being in the ith state of the ensemble $|\psi_i\rangle$, where $\langle\psi_i \mid \psi_i\rangle = 1$. The probabilities satisfy the obvious relations

$$0 \le p_i \le 1, \quad \sum_i p_i = 1, \quad \sum_i p_i^2 \le 1. \tag{A5}$$

For the special case where all the p_i vanish except, say, the jth one, $p_i = \delta_{ij}$, we obtain

$$\hat{\rho} = |\psi_j\rangle\langle\psi_j|, \tag{A6}$$

the density operator for the pure state $|\psi_j\rangle$. Note that the density operator for this case is just the projection operator onto the state $|\psi_j\rangle$, and for the more general case of Eq. (A4), the density operator is a sum of the projection operators over the ensemble, weighted with the probabilities of each member of the ensemble.

We now introduce a complete, orthonormal, basis $\{|\varphi_n\rangle\}$ $(\sum_n |\varphi_n\rangle\langle\varphi_n| = \hat{I})$, eigenstates of some observable. Then for the ith member of the ensemble we may write

$$|\psi_i\rangle = \sum_n |\varphi_n\rangle\, \langle\varphi_n \mid \psi_i\rangle = \sum_n c_n^{(i)}\, |\varphi_n\rangle, \tag{A7}$$

where $c_n^{(i)} = \langle\varphi_n \mid \psi_i\rangle$. The matrix element of $\hat{\rho}$ between the n and n' eigenstates is

$$\langle\varphi_n\, |\hat{\rho}|\, \varphi_{n'}\rangle = \sum_i \langle\varphi_n|\psi_i\rangle\, p_i\, \langle\psi_i|\varphi_{n'}\rangle = \sum_i p_i c_n^{(i)} c_{n'}^{(i)*}. \tag{A8}$$

The quantities $\langle\varphi_n|\hat{\rho}|\varphi_{n'}\rangle$ form the elements of the density matrix. Taking the trace of this matrix we have

$$\begin{aligned} \mathrm{Tr}\hat{\rho} &= \sum_n \langle\varphi_n\, |\hat{\rho}|\, \varphi_n\rangle = \sum_i \sum_n \langle\varphi_n|\psi_i\rangle\, p_i\, \langle\psi_i|\varphi_n\rangle \\ &= \sum_i \sum_n p_i\, \langle\psi_i|\varphi_n\rangle\, \langle\varphi_n|\psi_i\rangle = \sum_i p_i = 1. \end{aligned} \tag{A9}$$

Since $\hat{\rho}$ is Hermitian (as is evident from its construction in Eq. (A4)), the diagonal elements $\langle\varphi_n|\hat{\rho}|\varphi_n\rangle$ must be real, and it follows from Eq. (A9) that

$$0 \le \langle\varphi_n|\, \hat{\rho}\, |\varphi_n\rangle \le 1. \tag{A10}$$

Now let us consider the square of the density operator: $\hat{\rho}^2 = \hat{\rho} \bullet \hat{\rho}$. For a pure state where $\hat{\rho} = |\psi\rangle\langle\psi|$ it follows that

$$\hat{\rho}^2 = |\psi\rangle\, \langle\psi \mid \psi\rangle\, \langle\psi| = |\psi\rangle\, \langle\psi| = \hat{\rho} \tag{A11}$$

and thus

$$\mathrm{Tr}\hat{\rho}^2 = \mathrm{Tr}\hat{\rho} = 1. \tag{A12}$$

For a statistical mixture

$$\hat{\rho}^2 = \sum_i \sum_j p_i p_j |\psi_i\rangle\langle\psi_i \mid \psi_j\rangle\langle\psi_j|. \tag{A13}$$

Taking the trace we have

$$\begin{aligned}
\mathrm{Tr}\hat{\rho}^2 &= \sum_n \langle\varphi_n| \hat{\rho}^2 |\varphi_n\rangle \\
&= \sum_n \sum_i \sum_j p_i p_j \langle\varphi_n \mid \psi_i\rangle\langle\psi_i \mid \psi_j\rangle\langle\psi_j \mid \varphi_n\rangle \\
&= \sum_i \sum_j p_i p_j |\langle\psi_i \mid \psi_j\rangle|^2 \\
&\leq \left[\sum_i p_i\right]^2 = 1.
\end{aligned} \tag{A14}$$

The equality holds only if $|\langle\psi_i \mid \psi_j\rangle|^2 = 1$ for *every* pair of states $|\psi_i\rangle$ and $|\psi_j\rangle$. This is possible only if all the $|\psi_i\rangle$ are collinear in Hilbert space, i.e. equivalent up to an overall phase factor. Thus we have the following criteria for pure and mixed states:

$$\begin{aligned}
&\mathrm{Tr}\hat{\rho}^2 = 1, \text{for a pure state,} \\
&\mathrm{Tr}\hat{\rho}^2 < 1, \text{ for a mixed state.}
\end{aligned} \tag{A15}$$

Perhaps a simple example is warranted at this point. Consider a superposition of, say, the vacuum and one-photon number states

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle), \tag{A16}$$

where ϕ is just some phase. The density operator associated with this state is given by

$$\hat{\rho}_\psi = |\psi\rangle\langle\psi| = \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1| + e^{i\phi}|1\rangle\langle 0| + e^{-i\phi}|0\rangle\langle 1|]. \tag{A17}$$

On the other hand, the density operator for an equally populated mixture of vacuum and one-photon states is

$$\hat{\rho}_\mathrm{M} = \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1|]. \tag{A18}$$

The two density operators differ by the presence of the "off-diagonal", or "coherence", terms in the former, such terms being absent in the case of the mixture. The absence of the coherence terms is, of course, what makes the distinction between a state exhibiting full quantum-mechanical behavior and one that does not. It is easy to check that $\mathrm{Tr}\hat{\rho}_\mathrm{M}^2 = 1/2$.

For one of the states of the ensemble $|\psi_i\rangle$, by itself pure, the expectation value of some operator $\hat{O}$ is given by

$$\langle\hat{O}\rangle_i = \langle\psi_i|\hat{O}|\psi_i\rangle. \tag{A19}$$

For the statistical mixture, the ensemble average is given by

$$\langle \hat{O} \rangle = \sum_i p_i \, \langle \psi_i | \, \hat{O} \, | \psi_i \rangle, \tag{A20}$$

which is just the average of the quantum-mechanical expectation values weighted with the probabilities p_i. Formally we may write

$$\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\rho} \, \hat{O}) \tag{A21}$$

since

$$\begin{aligned} \mathrm{Tr}(\hat{\rho} \, \hat{O}) &= \sum_n \langle \varphi_n | \, \hat{\rho} \, \hat{O} \, | \varphi_n \rangle \\ &= \sum_n \sum_i p_i \, \langle \varphi_n \mid \psi_i \rangle \, \langle \psi_i | \, \hat{O} \, | \varphi_n \rangle \\ &= \sum_i \sum_n p_i \, \langle \psi_i | \, \hat{O} \, | \varphi_n \rangle \, \langle \varphi_n \mid \psi_i \rangle \\ &= \sum_i p_i \, \langle \psi_i | \, \hat{O} \, | \psi_i \rangle \, . \end{aligned} \tag{A22}$$

A.2 Two-state system and the Bloch sphere

For a two-state system, be it a spin-$1/2$ particle, a two-level atom, or the polarizations of a single photon, there always exists a description in terms of the Pauli operators $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3$ satisfying the commutation relations

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i \varepsilon_{ijk} \hat{\sigma}_k. \tag{A23}$$

In a basis where $\hat{\sigma}_3$ and $\hat{\sigma}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \hat{\sigma}_3^2$ are diagonal, these operators can be written in matrix form as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A24}$$

Any Hermitian 2×2 matrix can be expressed in terms of the Pauli matrices and the 2×2 identity matrix $\hat{I}_2$, and this includes, of course, the density operator. That is, we can write

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + s_3 & s_1 + i \, s_2 \\ s_1 - i \, s_2 & 1 - s_3 \end{pmatrix} = \frac{1}{2} (\hat{I}_2 + \mathbf{s} \cdot \sigma), \tag{A25}$$

where the vector $\mathbf{s} = \{s_1, s_2, s_3\}$ is known as the Bloch vector. For a pure state $\hat{\rho} = |\Psi\rangle\langle\Psi|$, the Bloch vector has unit length, $\sum_i |s_i|^2 = 1$, and points in some direction specified by the spherical coordinate angles θ and ϕ in a three-dimensional Euclidean space. The associated quantum state can be represented in terms of these angles as

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right) e^{-i\phi/2} \, |\uparrow\rangle + \sin\left(\frac{\theta}{2}\right) e^{i\phi/2} \, |\downarrow\rangle \, . \tag{A26}$$

In general, and including the case of mixed states where $|\mathbf{s}| \leq 1$, the density

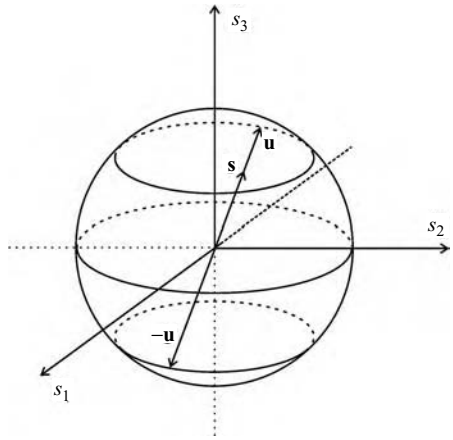


Fig. A.1. Representation of the density matrix of a two-state system in terms of the Bloch sphere and Bloch vector. The three components of the Bloch vector $\mathbf{s} = \{s_1, s_2, s_3\}$ specify the density operator according to the parameterization of Eq. (A25). The two eigenvalues are $(1 \pm |\mathbf{s}|)/2$ and the two eigenvectors are specified by $\mathbf{u}$ and $-\mathbf{u}$.

operator of the form of Eq. (A25) has two eigenvalues

$$g_1 = \frac{1}{2}\left[1 + \sqrt{s_1^2 + s_2^2 + s_3^2}\right] = \frac{1}{2}\left[1 + |\mathbf{s}|\right],$$
$$g_2 = \frac{1}{2}\left[1 - \sqrt{s_1^2 + s_2^2 + s_3^2}\right] = \frac{1}{2}\left[1 - |\mathbf{s}|\right], \tag{A27}$$

and its eigenvectors are determined by the two vectors $\mathbf{u}$ and $-\mathbf{u}$ shown in the Bloch sphere in Fig. A.1. For pure states, where $|\mathbf{s}| = 1$, $\mathbf{u}$ coincides with $\mathbf{s}$ and its tip lies on the surface of the Bloch sphere. For a mixed state where $|\mathbf{s}| < 1$, vector $\mathbf{u}$ points in the same direction as $\mathbf{s}$ but unlike $\mathbf{s}$ maintains unit length so that its tip always lies on the surface of the Bloch sphere. Equation (A26) can be used to express $\mathbf{u}$ and $-\mathbf{u}$ in terms of the vectors in state space.

A.3 Entangled states

Let us now consider a two-particle (or two-mode) system (known also as a bipartite system) and, for simplicity, let us assume that each particle can be in either of two one-particle states $|\psi_1\rangle$ or $|\psi_2\rangle$. Using the notation

$$\left|\psi_1^{(1)}\right\rangle, \text{ particle 1 in state 1,}$$
$$\left|\psi_2^{(1)}\right\rangle, \text{ particle 1 in state 2,}$$
$$\left|\psi_1^{(2)}\right\rangle, \text{ particle 2 in state 1,}$$
$$\left|\psi_2^{(2)}\right\rangle, \text{ particle 2 in state 2,}$$

we consider a pure two-particle superposition state (in general an entangled state)

$$|\Psi\rangle = C_1\left|\psi_1^{(1)}\right\rangle \otimes \left|\psi_2^{(2)}\right\rangle + C_2\left|\psi_2^{(1)}\right\rangle \otimes \left|\psi_1^{(2)}\right\rangle \tag{A28}$$

an example of which is given in Eq. (A2). (We have inserted the direct product symbol here for emphasis, although we generally assume it to be understood

and clear from context throughout this book.) Clearly, this can be extended for multiparticle (multipartite) systems. For such multipartite systems, we can define reduced density operators for each of the subsystems by tracing the density operator over the states of all the other systems. In the present case, with the density operator of the two-particle system given by $\hat{\rho} = |\Psi\rangle\langle\Psi|$, the reduced density operator for particle 1 is

$$\begin{aligned}\hat{\rho}^{(1)} &= \mathrm{Tr}_2\hat{\rho} = \left\langle\psi_1^{(2)}\middle|\hat{\rho}\middle|\psi_1^{(2)}\right\rangle + \left\langle\psi_2^{(2)}\middle|\hat{\rho}\middle|\psi_2^{(2)}\right\rangle \\ &= |C_1|^2\left|\psi_1^{(1)}\right\rangle\left\langle\psi_1^{(1)}\right| + |C_2|^2\left|\psi_2^{(1)}\right\rangle\left\langle\psi_2^{(1)}\right|. \end{aligned} \tag{A29}$$

This has the form for a mixed state for particle 1 as long as $C_i \neq 0,\ i = 1, 2$. Similarly, for particle 2,

$$\hat{\rho}^{(2)} = \mathrm{Tr}_1\hat{\rho} = |C_1|^2\left|\psi_1^{(2)}\right\rangle\left\langle\psi_1^{(2)}\right| + |C_2|^2\left|\psi_2^{(2)}\right\rangle\left\langle\psi_2^{(2)}\right|. \tag{A30}$$

Evidently, when one of the particles is considered without regard to the other, it is generally in a mixed state. Thus one may characterize the degree of entanglement according to the degree of purity of either of the subsystems. If $\mathrm{Tr}[\hat{\rho}^{(2)}]^2 = 1$, the state $|\Psi\rangle$ is not an entangled state; but if $\mathrm{Tr}[\hat{\rho}^{(2)}]^2 < 1$ we may conclude that $|\Psi\rangle$ describes an entanglement between subsystems 1 and 2.

A.4 Schmidt decomposition

There is another convenient way to approach the problem of characterizing entanglement, at least for cases where there are only two subsystems. We refer to the von Neumann entropy, which we introduce in the next section. But first, as a preliminary, we introduce the Schmidt decomposition [1, 2]. To keep the discussion general we do not restrict the dimensions of the Hilbert spaces of the subsystems.

Suppose we let $\{|a_i\rangle, i = 1, 2, 3, \ldots\}$ and $\{|b_j\rangle, j = 1, 2, 3, \ldots\}$ form orthonormal bases for subsystems $\mathcal{U}$ and $\mathcal{V}$ respectively. If the Hilbert spaces of these systems are denoted $\mathcal{H}_\mathcal{U}$ and $\mathcal{H}_\mathcal{V}$ respectively, then any state $|\Psi\rangle \in \mathcal{H} = \mathcal{H}_\mathcal{U} \otimes \mathcal{H}_\mathcal{V}$ can be written as

$$|\Psi\rangle = \sum_{i,j=1} c_{ij}|a_i\rangle \otimes |b_j\rangle. \tag{A31}$$

In these bases, the density operator for the composite system is

$$\begin{aligned}\hat{\rho} = |\Psi\rangle\langle\Psi| &= \sum_{i,j,k,l} c_{ij}c_{kl}^*\,|a_i\rangle\langle a_i| \otimes \left|b_j\right\rangle\langle b_l| \\ &= \sum_{i,j,k,l} \rho_{ij,kl}\,|a_i\rangle\langle a_i| \otimes \left|b_j\right\rangle\langle b_l|\,, \end{aligned} \tag{A32}$$

where the $\rho_{ij,kl} = c_{ij}c_{kl}^*$. The density operators of each of the subsystems are

given by the partial traces

$$\begin{aligned}\hat{\rho}_u &= \mathrm{Tr}_v\hat{\rho} = \sum_j \langle b_j|\hat{\rho}|b_j\rangle = \sum_{i,j,k} \rho_{i\,j,k\,j}|a_i\rangle\langle a_k|,\\ \hat{\rho}_v &= \mathrm{Tr}_u\hat{\rho} = \sum_i \langle a_i|\hat{\rho}|a_i\rangle = \sum_{i,j,k} \rho_{i\,j,i\,l}|b_j\rangle\langle b_l|.\end{aligned} \tag{A33}$$

But it is possible, for any bipartite system in a *pure* state, to write that state in the form

$$|\Psi\rangle = \sum_i g_i\,|u_i\rangle \otimes |v_i\rangle \tag{A34}$$

where the $\{|u_i\rangle\}$ and $\{|v_i\rangle\}$are othonormal bases in $\mathcal{H}_\mathcal{U}$and $\mathcal{H}_\mathcal{V}$ respectively and where $\sum_i |g_i|^2 = 1$. This is the Schmidt decomposition. The density operator for the composite system is then

$$\hat{\rho} = \sum_{i,k} g_i g_k^*\,|u_i\rangle\,\langle u_k| \otimes |v_i\rangle\,\langle v_k|. \tag{A35}$$

The proof of this is as follows. Let us choose $\{|u_i\rangle\}$ such that $\hat{\rho}_u$ is diagonal in this basis and the dimension of $\mathcal{H}_\mathcal{U}$ is not greater than the dimension of $\mathcal{H}_\mathcal{V}$. Let $\{|v_i'\rangle\}$ be an arbitrary basis in $\mathcal{H}_\mathcal{V}$. Then we can write

$$\hat{\rho} = |\Psi\rangle\,\langle\Psi| = \sum_{i,j,k,l} c_{i\,j}c_{k\,l}^*\,|u_i\rangle\,\langle u_k| \otimes \left|v_k'\right\rangle\!\left\langle v_l'\right|. \tag{A36}$$

However, $\hat{\rho}_u = \sum_{i,k,j} c_{i\,j}c_{k\,j}^*|u_i\rangle\langle u_k|$ must be diagonal, therefore $\sum_j c_{i\,j}c_{k\,j}^* = |g_i|^2\delta_{i\,k}$. This condition allows us to switch to another orthogonal basis in $\mathcal{H}_\mathcal{U}$. For each $|g_i| \neq 0$ and for each $|u_i\rangle$ we can define the relative state

$$|v_i\rangle = \sum_j \left(\frac{c_{ij}}{g_i}\right)|v_j'\rangle \tag{A37}$$

where $i = 1, 2, \ldots N$, and N is not greater than the dimension of $\mathcal{H}_\mathcal{U}$. The density operator of the composite system may now be written as

$$\begin{aligned}\hat{\rho} &= \sum_{i,j,k,l} g_i g_k^*\,|u_i\rangle\,\langle u_k| \otimes \left(\frac{c_{i\,j}}{g_i}\right)|v_j'\rangle\langle v_l'|\left(\frac{c_{k\,l}^*}{g_k^*}\right)\\ &= \sum_{i,k} g_i g_k^*\,|u_i\rangle\,\langle u_k| \otimes |v_i\rangle\,\langle v_k|\,,\end{aligned} \tag{A38}$$

where only N basis vectors $\{|v_i\rangle\}$ from $\mathcal{H}_\mathcal{V}$ are sufficient to express the state vector or the density operator of the composite system. This completes the proof.

Before proceeding, some remarks are in order. (i) The summation of the single index in the Schmidt decomposition goes to the smaller of the dimensions of the two Hilbert spaces $\mathcal{H}_\mathcal{U}$ and $\mathcal{H}_\mathcal{V}$. (ii) The decomposition is not unique and cannot, in general, be extended to multipartite systems with greater than two subsystems. (iii) In the bases chosen for the Schmidt decomposition, both $\hat{\rho}_u$ and $\hat{\rho}_v$ are diagonal and have the same positive spectrum:

$$\begin{aligned}\hat{\rho}_u &= \sum_i |g_i|^2\,|u_i\rangle\,\langle u_i|,\\ \hat{\rho}_v &= \sum_i |g_i|^2\,|v_i\rangle\,\langle v_i|.\end{aligned} \tag{A39}$$

It is this feature that facilitates the use of the von Neumann entropy, introduced in Section A.5, to characterize entanglement in a bipartite system. Again, the summation over the single index in the Schmidt decomposition goes to the smaller of the dimensionalities of the two Hilbert spaces $\mathcal{H}_{\mathcal{U}}$ and $\mathcal{H}_{\mathcal{V}}$. This means that if a two-state system is entangled with a system of N dimensions, where $N > 2$, the Schmidt decomposition has only two terms. An example of this for a two-level atom interacting with a single-mode field is discussed at the end of Chapter 4.

Here we give a simple example of the Schmidt decomposition for a bipartite system. We consider a spin singlet (SS) state of the form

$$|\mathrm{SS}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle_1|\downarrow_z\rangle_2 - |\downarrow_z\rangle_1|\uparrow_z\rangle_2\right), \tag{A40}$$

where $|\uparrow_z\rangle_1$ is the state for particle 1 with spin up along the z-axis, etc. Clearly this state already possesses a Schmidt decomposition. But because the spin singlet state is rotationally invariant, it is particularly easy to find other Schmidt decompositions. For example, we can write

$$|\mathrm{SS}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle_1|\downarrow_x\rangle_2 - |\downarrow_x\rangle_1|\uparrow_x\rangle_2\right), \tag{A41}$$

or more generally

$$|\mathrm{SS}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{\mathbf{n}}\rangle_1|\downarrow_{\mathbf{n}}\rangle_2 - |\downarrow_{\mathbf{n}}\rangle_1|\uparrow_{\mathbf{n}}\rangle_2\right), \tag{A42}$$

where $\mathbf{n}$ is a unit vector for a quantization axis in some arbitrary direction.

A.5 von Neumann entropy

The concept of entropy is familiar from thermodynamics and is commonly understood as a measure of disorder: the greater the disorder within a system the greater the entropy. From the point of view of statistical mechanics and information theory, entropy can be thought of as a measure of missing information, the information that would be gained if a complete measurement of a system could be performed [3].

The von Neumann definition of entropy closely parallels the statistical mechanical one. For a density operator $\hat{\rho}$, the von Neumann entropy is defined as [4]

$$S(\hat{\rho}) = -\mathrm{Tr}\left[\hat{\rho}\ln\hat{\rho}\right]. \tag{A43}$$

For a pure state $S(\hat{\rho}_{\mathrm{pure}}) = 0$ and thus repeated measurements of the state in question yield no new information. But for a mixed state we shall have $S(\hat{\rho}_{\mathrm{mixed}}) > 0$. In general, the entropy is not easy to calculate. However, in a basis for which the density operator is diagonal, such as the Schmidt basis, the entropy may be evaluated from the diagonal elements according to

$$S(\hat{\rho}) = -\sum_k \rho_{kk}\ln\rho_{kk}. \tag{A44}$$

Because all of the numbers ρ_{kk} must be real and $0 \leq \rho_{kk} \leq 1$, it follows that $S(\hat{\rho})$ must be positive semidefinite.

As an example, consider the bipartite state

$$|\psi\rangle = \frac{1}{\sqrt{1+|\xi|^2}} \left(|0\rangle_1 |0\rangle_2 + \xi |1\rangle_1 |1\rangle_2\right), \tag{A45}$$

which clearly has the form of a Schmidt decomposition. The density operators for each of the subsystem are

$$\begin{aligned} \hat{\rho}_1 &= \frac{1}{(1+|\xi|^2)}[|0\rangle_{11}\langle 0| + |\xi|^2 |1\rangle_{11}\langle 1|], \\ \hat{\rho}_2 &= \frac{1}{(1+|\xi|^2)}[|0\rangle_{22}\langle 0| + |\xi|^2 |1\rangle_{22}\langle 1|]. \end{aligned} \tag{A46}$$

Then it follows that

$$\begin{aligned} S(\hat{\rho}_1) = -\Bigg\{ &\frac{1}{(1+|\xi|^2)} \ln\left[\frac{1}{(1+|\xi|^2)}\right] \\ &+ \frac{|\xi|^2}{(1+|\xi|^2)} \ln\left[\frac{|\xi|^2}{(1+|\xi|^2)}\right] \Bigg\} = S(\hat{\rho}_2). \end{aligned} \tag{A47}$$

For the case where $\xi = 0$ we have the nonentangled product state $|0\rangle_1 |0\rangle_2$ and $S(\hat{\rho}_1) = S(\hat{\rho}_2) = 0$, as expected. But for $|\xi| = 1$ we have $S(\hat{\rho}_1) = S(\hat{\rho}_2) = \ln 2$, which represents maximal entanglement for this state indicating maximal *quantum* correlations. Information about these correlations is destroyed upon tracing over one of the subsystems.

Finally, in the case of a pure state of a bipartite system where the subsystems are of different dimensions, as is easy to check, the entropies of the of the subsystems are equal. This is a consequence of the Schmidt basis enforcing the same dimensionality on each of the subsystems, as mentioned above. An example of such is provided, in the case of the Jaynes–Cummings model, at the end of Chapter 4.

A.6 Dynamics of the density operator

Finally, in the absence of dissipative interactions and in the absence of an explicitly time dependent interaction, the density operator evolves unitarily according to

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{H}], \tag{A48}$$

easily proved by using the fact that each of the states $|\psi_i\rangle$ of the ensemble satisfies the Schrödinger equation

$$i\hbar \frac{d|\psi_i\rangle}{dt} = \hat{H}|\psi_i\rangle. \tag{A49}$$

Alternatively, we may write

$$\hat{\rho}(t) = \hat{U}(t,0)\,\hat{\rho}(0)\,\hat{U}^{\dagger}(t,0) \tag{A50}$$

where the unitary evolution operator $\hat{U}(t, 0)$ satisfies the equation

$$i\hbar \frac{d\hat{U}}{dt} = \hat{H}\hat{U}. \tag{A51}$$

Equation (A48) is known as the von Neumann equation, the quantum mechanical analog of the Liouville equation associated with the evolution phase-space probability distributions in statistical mechanics. The equation is sometimes written as

$$i\hbar \frac{d\hat{\rho}(t)}{dt} = [\hat{H}, \hat{\rho}(t)] \equiv \mathcal{L}\hat{\rho}(t), \tag{A52}$$

where $\hat{\mathcal{L}}$ is known as the Liouvillian superoperator.

References

[1] E. Schmidt, *Math. Annalen*, **63** (1907), 433. For an early application in quantum mechanics see the chapter by H. Everett, III, in *The Many Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham (Princeton: Princeton University Press, 1973), p. 3.

[2] A. Ekert and P. L. Knight, *Am. J. Phys*., **63** (1995), 415.

[3] See, for example, the discussion in J. Machta, *Am. J. Phys.*, **67** (1999), 1074.

[4] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press, 1955).

Bibliography

Two useful review articles on the density operator that we recommend are:

U. Fano, "Description of states in quantum mechanics by density matrix and operator techniques", *Rev. Mod. Phys.*, **29** (1957), 74.

D. Ter Haar, "Theory and application of the density matrix", *Rep. Prog. Phys.*, **24** (1961), 304.

See also:

M. Weissbluth, *Atoms and Molecules* (New York: Academic Press, 1978), chapter 13.

A review on entanglement and its characterization is:

D. Bruß, "Characterizing entanglement", *J. Math. Phys.*, **43** (2003), 4237.

A review of entropy in physics and information theory can be found in:

A. Wehrl, "General properties of entropy", *Rev. Mod. Phys.*, **50** (1978), 221.

Appendix B
Quantum measurement theory in a (very small) nutshell

The following is in no way meant to be an extensive review of the subject of quantum measurement theory. For more details, the reader should consult the bibliography. Here we present only those aspects of the theory necessary to understand the results of state reductive measurements, particularly for situations involving entangled states.

Suppose the operator $\hat{Q}$ is a Hermitian operator representing some observable Q. Further, we let the states $\{|q_n\rangle; n \text{ integer}\}$ be the eigenstates of $\hat{Q}$ with the eigenvalues being $q_n : \hat{Q}|q_n\rangle = q_n|q_n\rangle$. The eigenstates are complete and resolve unity according to

$$\sum_n |q_n\rangle\langle q_n| = \hat{I}. \tag{B1}$$

Each of the terms $|q_n\rangle\langle q_n|$ forms a projection operator onto the state with eigenvalue q_n, $\hat{P}(q_n) := |q_n\rangle\langle q_n|$. A state $|\psi\rangle$ (assumed normalized: $\langle\psi \mid \psi\rangle = 1$) can be expanded in terms of the eigenstates of $\hat{Q}$ according to

$$|\psi\rangle = \hat{I}\,|\psi\rangle = \sum_n |q_n\rangle\,\langle q_n \mid \psi\rangle = \sum_n c_n\,|q_n\rangle \tag{B2}$$

where the coefficient $c_n = \langle q_n \mid \psi\rangle$ is, in general, a complex number, and is known as a probability amplitude. If $|\psi\rangle$ is the state vector just before a measurement of the observable Q, the probability of obtaining the outcome q_n, $P(q_n)$, is

$$P(q_n) = \langle\psi|\hat{P}(q_n)|\psi\rangle = |c_n|^2. \tag{B3}$$

According to the orthodox (or Copenhagen) interpretation of quantum mechanics, the system described by the state vector $|\psi\rangle$ is *not* in one of the eigenstates of $\hat{Q}$ prior to the measurement, but rather, the state vector *collapses*, or *reduces*, to one of the eigenstates upon measurement, this processes sometimes represented by the symbols

$$|\psi\rangle \xrightarrow[\text{of } \hat{Q}]{\text{measurement}} |q_n\rangle. \tag{B4}$$

Prior to the measurement, the value of the observable associated with the operator $\hat{Q}$ in the system described by the state vector $|\psi\rangle$ is objectively indefinite. It is not merely a matter of only knowing what the value of the observable is

with a probability given by Eq. (B3), as would be the case if we were talking about statistical mechanics, but rather that one cannot assign definite values to observables for systems in superposition states without leading to conflict with experimental observations. The state reduction processes throws the system into a state with a definite value of the observable and subsequent measurements of that observable will return the same state. The dynamics of state reduction is not described by the Schrödinger equation. The interaction of the detector with the particle whose quantum state is being measured may destroy that particle. For example, when a photo-detector "clicks" it does so because the photon itself has been destroyed (absorbed).

It is sometimes possible to arrange experiments to perform selective, or filtered, measurements, sometimes called von Neumann projections. Such measurements are invoked in the text in Chapter 9 where polarization filters are employed and in Chapter 10 where field ionization is used in the context of cavity quantum electrodynamics. A projective measurement on one of the eigenstates $|q_n\rangle$ requires a filtration of all the other eigenstates of the operator. Mathematically we may represent this projection as $\hat{P}(q_n)|\psi\rangle$, which in normalized form reads

$$|\psi\rangle \xrightarrow[\text{measurement}]{\text{projective}} |q_n\rangle = \frac{\hat{P}(q_n)\,|\psi\rangle}{\langle\psi|\,\hat{P}(q_n)\,|\psi\rangle^{1/2}}. \tag{B5}$$

The filtering need not project onto an eigenstate of $\hat{Q}$. Suppose we could somehow filter the system so that the superposition state $|\psi_s\rangle = \frac{1}{\sqrt{2}}(|q_1\rangle + |q_2\rangle)$ is measured. The projection operator associated with this state is $\hat{P}_s = |\psi_s\rangle\langle\psi_s|$ and we have the projection

$$|\psi\rangle \xrightarrow[\text{measurement onto } |\psi_s\rangle]{\text{projective}} e^{i\varphi}\,|\psi_s\rangle = \frac{\hat{P}_s\,|\psi\rangle}{\langle\psi|\,\hat{P}_s\,|\psi\rangle^{1/2}}, \tag{B6}$$

where $\hat{P}_s|\psi\rangle = |\psi_s\rangle\langle\psi_s \mid \psi\rangle = |\psi_s\rangle(|c_1 + c_2|e^{i\varphi})/\sqrt{2}$, and $\langle\psi|\hat{P}_s|\psi\rangle = |c_1 + c_2|^2/2$. The phase φ is an irrelevant overall phase factor.

As a specific example, consider a single photon in the polarization state

$$|\theta\rangle = |H\rangle\cos\theta + |V\rangle\sin\theta \tag{B7}$$

where we have used the convention given in Chapter 9. If a Polaroid filter is placed in the beam and oriented along the horizontal (vertical) direction, the state vector of Eq. (B7) reduces to $|H\rangle(|V\rangle)$. On then other hand, if we try to project onto the state

$$|\theta^{\perp}\rangle = -|H\rangle\sin\theta + |V\rangle\cos\theta, \tag{B8}$$

where $\theta^{\perp} = \theta + \pi/2$, by orienting the filter along that direction, we find that we get no photons at all passing through: $\langle\theta^{\perp} \mid \theta\rangle = 0$. But if we orient the filter along some direction ϑ to the horizontal axis, the photon will be projected onto the polarization state

$$|\vartheta\rangle = |H\rangle\cos\vartheta + |V\rangle\sin\vartheta, \tag{B9}$$

as the reader can easily check.

Let us now turn to the case of two-mode states. If we have two modes labeled 1 and 2, we may write a general two-mode state as

$$|\psi\rangle = \sum_{n,m} c_{n\,m}\,|q_n\rangle_1\,|s_m\rangle_2\,, \tag{B10}$$

where the states $|q_n\rangle_1$ are eigenstates of $\hat{Q}_1$ defined in the Hilbert space of mode 1 and the states $|s_m\rangle$ are eigenstates of some operator $\hat{S}_2$ defined in the Hilbert space of mode 2. Certain choices of the coefficients $c_{n\,m}$ will render the state an entangled state. In either case, an unfiltered measurement of the observables $\hat{Q}_1$ and $\hat{S}_2$ reduces the state vector to

$$|\psi\rangle \xrightarrow[\text{of}\hat{Q}_1\ \text{and}\hat{S}_2]{\text{measurement}} |q_n\rangle_1\,|s_m\rangle_2\,, \tag{B11}$$

whereas, say, a projective measurement onto $|q_n\rangle_1$ reduces the state vector of Eq. (B10) to

$$\begin{aligned} |\psi\rangle \xrightarrow[\text{onto}\ |q_n\rangle_1]{\text{projection}} |\psi, q_n\rangle &= \frac{\hat{P}_1(q_n)|\psi\rangle}{\langle\psi|\hat{P}_1(q_n)|\psi\rangle^{1/2}} \\ &= |q_n\rangle_1 \frac{\sum\limits_m c_{n\,m}|s_m\rangle}{\sum\limits_m |c_{n\,m}|^2}. \end{aligned} \tag{B12}$$

The projection holds whether or not we have an entangled state, as do the more general projections of the type discussed above. Notice that if the original state *is* entangled, the projective measurement creates a factorized state. As an example, let us consider a state of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle_1\,|V\rangle_2 - |V\rangle_1\,|H\rangle_2\right), \tag{B13}$$

one of the Bell states. Suppose that in mode 1 a Polaroid filter is placed at an angle of $\pi/4$ to the horizontal. This causes a projection onto the state

$$|\pi/4\rangle_1 = \frac{1}{\sqrt{2}}(|H\rangle_1 + |V\rangle_1) \tag{B14}$$

in that mode and reduces the original state vector to

$$|\psi\rangle \xrightarrow[\text{onto}|\pi/4\rangle_1]{\text{projection}} \frac{\hat{P}_1(\frac{\pi}{4})|\psi\rangle}{\langle\psi|\hat{P}_1(\frac{\pi}{4})|\psi\rangle^{1/2}} = |\pi/4\rangle_1 \frac{1}{\sqrt{2}}(|V\rangle_2 - |H\rangle_2). \tag{B15}$$

Note that a projective measurement of one part of an entangled system projects the other part into a particular state in a manner that is entirely predictable. The correlations exhibited by these projections are of a highly nonclassical nature and are ultimately responsible for the violations of Bell's inequalities as described in Chapter 9.

Bibliography

Excellent modern discussions on the issues raised in this appendix can be found in:

C. J. Isham, *Lectures on Quantum Theory* (London: Imperial College Press, 1995).

One should also see:

K. Gottfried, *Quantum Mechanics Volume I: Fundamentals* (Reading: Addison-Wesley, 1989), chapter IV.

Of course, one should consult:

J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press, 1955), especially chapter VI.

Appendix C
Derivation of the effective Hamiltonian for dispersive (far off-resonant) interactions

Here we derive the effective Hamiltonian for interactions where there is a large detuning between the relevant frequencies involved. In order to keep the discussion rather general, we start with the full Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \hat{H}_\mathrm{I} \tag{C1}$$

where $\hat{H}_0$ is the interaction-free Hamiltonian and H_I is the interaction Hamiltonian. We take the latter to be of the general form

$$\hat{H}_\mathrm{I} = \hbar g (\hat{A} + \hat{A}^\dagger), \tag{C2}$$

where $\hat{A}$ represents the product of operators describing the interaction, and g is the coupling constant. The operator product $\hat{A}$ is assumed to have no explicit time dependence. In the case of atom–field interactions we have $\hat{A} = \hat{a}\,\hat{\sigma}_+$ and $\hat{H}_0 = \hbar\omega\hat{a}^\dagger\hat{a} + \hbar\omega_0\hat{\sigma}_3/2$. This specific case was discussed by Schneider *et al.* [1]. The time-dependent Schrödinger equation in the Schrödinger picture (SP) is

$$i\,\hbar\frac{d}{dt}\,|\psi_\mathrm{SP}(t)\rangle = \left(\hat{H}_0 + \hat{H}_\mathrm{I}\right)|\psi_\mathrm{SP}(t)\rangle \tag{C3}$$

where $|\psi_\mathrm{SP}(t)\rangle$ is the state vector in that picture.

We now transform to the interaction picture (IP) using the interaction-free part of the Hamiltonian and the operator $\hat{U}_0 = \exp(-i\,\hat{H}_0\,t/\hbar)$ [2]. The state vector in the IP is given by

$$|\psi_\mathrm{IP}(t)\rangle = \hat{U}_0^{-1}\,|\psi_\mathrm{SP}(t)\rangle \tag{C4}$$

and the Schrödinger equation becomes

$$i\,\hbar\frac{d}{dt}\,|\psi_\mathrm{IP}(t)\rangle = \hat{H}_\mathrm{IP}(t)\,|\psi_\mathrm{IP}(t)\rangle\,, \tag{C5}$$

where

$$\begin{aligned}\hat{H}_\mathrm{IP}(t) &= \hat{U}_0^{-1}\hat{H}\hat{U}_0 - i\,\hbar\hat{U}_0^{-1}\frac{d\hat{U}_0}{dt}\\ &= \hbar g(\hat{A}e^{i\Delta t} + \hat{A}^\dagger e^{-i\Delta t}),\end{aligned} \tag{C6}$$

and where Δ is the detuning whose form depends on the form of $\hat{A}$. For the atom–field interaction we have $\Delta = \omega_0 - \omega$. The detuning will be assumed large in a sense described below.

The solution to Eq. (C5) can be written formally as

$$|\psi_{\rm IP}(t)\rangle = \hat{\mathcal{T}}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t')\right)\right]|\psi_{\rm IP}(0)\rangle\,, \tag{C7}$$

where $|\psi_{\rm IP}(0)\rangle = |\psi_{\rm SP}(0)\rangle$ and $\hat{\mathcal{T}}$ is the time-ordering operator. We make the perturbative expansion

$$\begin{aligned}
&\hat{\mathcal{T}}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t')\right)\right] \\
&= \hat{\mathcal{T}}\left[\hat{1} - \frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t') - \frac{1}{2\hbar^2}\int_0^t dt'\int_0^t dt''\,\hat{H}_{\rm IP}(t')\hat{H}_{\rm IP}(t'') + \cdots\right] \\
&= \hat{1} - \frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t') - \frac{1}{2\hbar^2}\hat{\mathcal{T}}\left[\int_0^t dt'\int_0^t dt''\,\hat{H}_{\rm IP}(t')\hat{H}_{\rm IP}(t'')\right] + \cdots
\end{aligned} \tag{C8}$$

The last term, when time ordered, reads

$$\hat{\mathcal{T}}\left[\int_0^t dt'\int_0^t dt''\,\hat{H}_{\rm IP}(t')\hat{H}_{\rm IP}(t'')\right] = 2\int_0^t dt'\,\hat{H}_{\rm IP}(t')\int_0^{t'} dt''\,\hat{H}_{\rm IP}(t''). \tag{C9}$$

The second term in Eq. (C8) yields

$$\begin{aligned}
\int_0^t dt'\,\hat{H}_{\rm IP}(t') &= \hbar g\left[\hat{A}\frac{e^{i\Delta t'}}{i\Delta}\bigg|_0^t - \hat{A}^\dagger\frac{e^{-i\Delta t'}}{i\Delta}\bigg|_0^t\right] \\
&= \frac{\hbar g}{i\Delta}[\hat{A}(e^{i\Delta t}-1) - \hat{A}^\dagger(e^{-i\Delta t}-1)].
\end{aligned} \tag{C10}$$

The second-order term now becomes

$$\begin{aligned}
\int_0^t dt'\,\hat{H}_{\rm IP}(t')\int_0^{t'} dt''\,\hat{H}_{\rm IP}(t'') = \frac{\hbar^2 g^2}{i\Delta}\int_0^t dt'\Big[&\hat{A}^2 e^{2i\Delta t'} - \hat{A}^2 e^{i\Delta t'} \\
&- \hat{A}^{\dagger 2} e^{-2i\Delta t'} + \hat{A}^{\dagger 2} e^{-i\Delta t'} + \hat{A}^\dagger\hat{A}(1-e^{-i\Delta t}) - \hat{A}\hat{A}^\dagger(1-e^{i\Delta t})].
\end{aligned} \tag{C11}$$

Clearly, the integration will give rise to terms that go as g^2/Δ^2 and these will be small for large detuning. Those terms we drop. This is essentially the rotating-wave approximation as those terms will be rapidly oscillating and will average to zero. What remains gives us

$$\int_0^t dt'\,\hat{H}_{\rm IP}(t')\int_0^{t'} dt''\,\hat{H}_{\rm IP}(t'') \approx \frac{i\hbar^2 g^2 t}{\Delta}[\hat{A},\,\hat{A}^\dagger]. \tag{C12}$$

Thus to second order we have

$$\hat{\mathcal{T}}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t')\right)\right] \approx \hat{1} - \frac{g}{\Delta}[\hat{A}(e^{i\Delta t}-1) - \hat{A}^\dagger(e^{-i\Delta t}-1)] - \frac{i\,g^2 t}{\Delta}[\hat{A},\hat{A}^\dagger]. \tag{C13}$$

If the mean "excitation" $\langle\hat{A}\rangle \approx \langle\hat{A}^\dagger\hat{A}\rangle^{1/2}$ is not too large and if

$$\left|\frac{g}{\Delta}\langle\hat{A}^\dagger\hat{A}\rangle^{1/2}\right| \ll 1, \tag{C14}$$

assumed valid because of the large detuning, then the second term of Eq. (C13) can be dropped and thus we have

$$\hat{\mathcal{T}}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\,\hat{H}_{\rm IP}(t')\right)\right] \approx \hat{1} - \frac{i}{\hbar}t\,\hat{H}_{\rm eff} \tag{C15}$$

where

$$\hat{H}_{\rm eff} = \frac{\hbar g^2}{\Delta}[\hat{A},\hat{A}^\dagger]. \tag{C16}$$

For the Jaynes–Cummings interaction we have $\hat{A} = \hat{a}\,\hat{\sigma}_+$ so that

$$[\hat{A},\hat{A}^\dagger] = \hat{\sigma}_+\hat{\sigma}_- + \hat{a}^\dagger\hat{a}\,\hat{\sigma}_3. \tag{C17}$$

Thus the effective Hamiltonian in this case is

$$\hat{H}_{\rm eff} = \hbar\chi(\hat{\sigma}_+\hat{\sigma}_- + \hat{a}^\dagger\hat{a}\,\hat{\sigma}_3), \tag{C18}$$

where $\chi = g^2/\Delta$, in agreement with the expression given in Eq. (4.184). The term $\hbar\chi\hat{\sigma}_+\hat{\sigma}_-$, present even in the absence of photons, is a kind of cavity induced atomic Kerr effect giving rise to an energy shift on the bare excited atomic state $|e\rangle$.

Finally, as another example, suppose we consider the case of the degenerate parametric down-converter with a quantized pump field. The full Hamiltonian for this process is given in Eq. (7.84). In our notation here we have $\hat{H}_0 = \hbar\omega\,\hat{a}^\dagger\hat{a} + \hbar\omega_p\hat{b}^\dagger\hat{b}$, where ω_p is the pump frequency, and $\hat{A} = i\,\hat{a}^{\dagger 2}\hat{b}$ with $g = \chi^{(2)}$. The effective interaction for large detuning is then

$$\hat{H}_{\rm eff} = -\frac{\hbar\left[\chi^{(2)}\right]^2}{\Delta}(4\hat{a}^\dagger\hat{a}\,\hat{b}^\dagger\hat{b} + 2\hat{b}^\dagger\hat{b} - \hat{a}^{\dagger 2}\hat{a}^2), \tag{C19}$$

where $\Delta = 2\omega - \omega_p$. The dispersive form of the interaction is a Kerr-like term $(\hat{a}^{\dagger 2}\hat{a}^2)$, a cross-Kerr-like term $(4\hat{a}^\dagger\hat{a}\,\hat{b}^\dagger\hat{b})$ and a "frequency-pulling" term $(2\hat{b}^\dagger\hat{b})$. Thus in the case of large detuning between pump and signal fields, the parametric down-converter mimics a Kerr medium. Klimov *et al.* [3] have discussed such interactions though they have used Lie algebraic methods to arrive at the effective Hamiltonians.

References

[1] S. Schneider, A. M. Herkommer, U. Leonhardt and W. Schleich, *J. Mod. Opt*., **44** (1997), 2333.

[2] See, for example, J. J. Sakurai, *Advanced Quantum Mechanics* (Reading: Addison-Wesley, 1967); see chapter 4.

[3] See A. B. Klimov, L. L. Sánchez-Soto and J. Delgado, *Opt. Commun*., **191** (2001), 419, and references therein.

Appendix D
Nonlinear optics and spontaneous parametric down-conversion

Down-conversion results from a nonlinear interaction of pump radiation with media where the induced polarization is so strongly effected by the radiation that it deforms beyond the linear response that generates the usual dispersion and absorption. For our purposes, we can expand the nonlinear polarization in a power series in the applied radiation field. Crystals commonly used in nonlinear optics are highly anisotropic and their response is described in tensorial form according to

$$\hat{P}_i = \chi^{(1)}_{i,j}\hat{E}_j + \chi^{(2)}_{i,j,k}\hat{E}_j\hat{E}_k + \chi^{(3)}_{i,j,k,l}\hat{E}_j\hat{E}_k\hat{E}_l + \cdots \tag{D1}$$

where $\chi^{(m)}$ is the mth-order electric susceptibility tensor [1] and where repeated indices imply a sum. The energy density is then $\varepsilon_0 E_i P_i$ and thus the *second-order* contribution to the Hamiltonian, the interaction Hamiltonian, is

$$\hat{H}^{(2)} = \varepsilon_0 \int_V d^3\mathbf{r}\,\chi^{(2)}_{i,j,k}\hat{E}_i\hat{E}_j\hat{E}_k, \tag{D2}$$

where the integral is over the interaction volume. We now represent the components of the fields as Fourier integrals of the form

$$\hat{E}(\mathbf{r},t) = \int d^3\mathbf{k}\Big[\hat{E}^{(-)}(\mathbf{k})\,e^{-i[\omega(\mathbf{k})\,t-\mathbf{k}\cdot\mathbf{r}]} + \hat{E}^{(+)}(\mathbf{k})\,e^{i[\omega(\mathbf{k})\,t-\mathbf{k}\cdot\mathbf{r}]}\Big], \tag{D3}$$

where

$$\hat{E}^{(-)}(\mathbf{k}) = i\sqrt{\frac{2\pi\,\hbar\omega(\mathbf{k})}{V}}\;\hat{a}^{\dagger}(\mathbf{k}), \text{ and } \hat{E}^{(-)}(\mathbf{k}) = i\sqrt{\frac{2\pi\,\hbar\omega(\mathbf{k})}{V}}\;\hat{a}(\mathbf{k}). \tag{D4}$$

The operators $\hat{a}(\mathbf{k})$ and $\hat{a}^{\dagger}(\mathbf{k})$ are the annihilation and creation operators respectively of photons with momentum $\hbar\mathbf{k}$.* If we substitute field expressions of the above form into Eq. (D2) and retain only the terms important for the case when the signal and idler modes are initially in vacuum states, we obtain the interaction Hamiltonian

$$\hat{H}_{\mathrm{I}}(t) = \varepsilon_0 \int_V d^3\mathbf{r} \int d^3\mathbf{k}_s d^3\mathbf{k}_i\,\chi^{(2)}_{lmn} \times \hat{E}^{(+)}_{\mathrm{pl}} e^{i\left[\omega_p(\mathbf{k}_p)\,t-\mathbf{k}_p\cdot\mathbf{r}\right]}\hat{E}^{(-)}_{\mathrm{sm}} e^{-i[\omega_s(\mathbf{k}_s)\,t-\mathbf{k}_s\cdot\mathbf{r}]}\hat{E}^{(-)}_{\mathrm{in}} e^{-i[\omega_i(\mathbf{k}_i)\,t-\mathbf{k}_i\cdot\mathbf{r}]} + H.c. \tag{D5}$$

* The commutation relations for these operators where the wave vector is continuous take the form $[\hat{a}(\mathbf{k}), \hat{a}^{\dagger}(\mathbf{k}')] = \delta^{(3)}(\mathbf{k}-\mathbf{k}')$.

The conversion rates for the process depend on the second or electric susceptibility $\chi^{(2)}$ but typically have efficiencies in the range 10^{-7} to 10^{-11}, extremely low rates. For this reason, in order to obtain significant output in the signal and idler beams it is necessary to pump the medium with a very strong coherent field which we can model as a classical field obtained from a laser as long as we are interested in interactions over a short enough time such that depletion of pump photons can be ignored – the parametric approximation. The pump laser is usually in the ultraviolet while the photons arising from the down-conversion are usually in the visible spectral range.

From time dependent perturbation theory, assuming the initial states of the signal and idler modes are in vacuum states, which we denote for the moment as $|\Psi_0\rangle$, we obtain, to first order, $|\Psi\rangle \approx |\Psi_0\rangle + |\Psi_1\rangle$ where [2]

$$\begin{aligned} |\Psi_1\rangle &= -\frac{i}{\hbar}\int d\,t\hat{H}(t)|\Psi_0\rangle \\ &= \mathcal{N}\int d^3\mathbf{k}_s d^3\mathbf{k}_i \delta(\omega_p - \omega_s(\mathbf{k}_s) - \omega_i(\mathbf{k}_i)) \\ &\quad \times \delta^{(3)}(\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i)\hat{a}_s^\dagger(\mathbf{k}_s)\hat{a}_i^\dagger(\mathbf{k}_i)|\Psi_0\rangle, \end{aligned} \tag{D6}$$

where $\mathcal{N}$ is a normalization factor into which all constants have been absorbed. One sees that the delta functions contain the phase matching conditions

$$\begin{aligned} \omega_p &= \omega_s + \omega_i, \\ \mathbf{k}_p &= \mathbf{k}_s + \mathbf{k}_i. \end{aligned} \tag{D7}$$

In the case of type I phase matching, we end up with the state given by Eq. (9.9), which we arrived at by assuming specific momenta which can be post-selected by the placement of a screen with properly located holes over the output of the down-converter.

References

[1] See R. W. Boyd, *Nonlinear Optics*, 2nd edition (New York: Academic Press, 2003).

[2] See Y. Shih, *Rep. Prog. Phys.*, **66** (2003), 1009.

Index