Lecture Notes in Physics

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Tokyo

Physics and Astronomy ONLINE LIBRARY
http://www.springer.de/phys/

Editorial Policy

The series *Lecture Notes in Physics* (LNP), founded in 1969, reports new developments in physics research and teaching -- quickly, informally but with a high quality. Manuscripts to be considered for publication are topical volumes consisting of a limited number of contributions, carefully edited and closely related to each other. Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to familiarize themselves with the topic concerned. For this reason, traditional proceedings cannot be considered for this series though volumes to appear in this series are often based on material presented at conferences, workshops and schools (in exceptional cases the original papers and/or those not included in the printed book may be added on an accompanying CD ROM, together with the abstracts of posters and other material suitable for publication, e.g. large tables, colour pictures, program codes, etc.).

Acceptance

A project can only be accepted tentatively for publication, by both the editorial board and the publisher, following thorough examination of the material submitted. The book proposal sent to the publisher should consist at least of a preliminary table of contents outlining the structure of the book together with abstracts of all contributions to be included.
Final acceptance is issued by the series editor in charge, in consultation with the publisher, only after receiving the complete manuscript. Final acceptance, possibly requiring minor corrections, usually follows the tentative acceptance unless the final manuscript differs significantly from expectations (project outline). In particular, the series editors are entitled to reject individual contributions if they do not meet the high quality standards of this series. The final manuscript must be camera-ready, and should include both an informative introduction and a sufficiently detailed subject index.

Contractual Aspects

Publication in LNP is free of charge. There is no formal contract, no royalties are paid, and no bulk orders are required, although special discounts are offered in this case. The volume editors receive jointly 30 free copies for their personal use and are entitled, as are the contributing authors, to purchase Springer books at a reduced rate. The publisher secures the copyright for each volume. As a rule, no reprints of individual contributions can be supplied.

Manuscript Submission

The manuscript in its final and approved version must be submitted in camera-ready form. The corresponding electronic source files are also required for the production process, in particular the online version. Technical assistance in compiling the final manuscript can be provided by the publisher's production editor(s), especially with regard to the publisher's own Latex macro package which has been specially designed for this series.

Online Version/ LNP Homepage

LNP homepage (list of available titles, aims and scope, editorial contacts etc.):
http://www.springer.de/phys/books/lnpp/

LNP online (abstracts, full-texts, subscriptions etc.):
http://link.springer.de/series/lnpp/

A. Buchleitner K. Hornberger (Eds.)

Coherent Evolution in Noisy Environments

Springer

Editors

Andreas Buchleitner
MPI für Physik komplexer Systeme
Nöthnitzer Str. 38
01187 Dresden, Germany

Klaus Hornberger
Universität Wien
Institut für Experimentalphysik
Boltzmanngasse 5
1090 Wien, Austria

Cover Picture: (see contribution by K. Wiesenfeld et al. in this volume)

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

CR Subject Classification (1998): D.4.5, E.3, C.2.0, H.2.0, K.6.5, K.4.4

ISSN 0075-8450
ISBN 3-540-44354-1 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

Printed in Germany

Typesetting: Camera-ready by the authors/editor
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
SPIN: 10894710 54/3141/du - 5 4 3 2 1 0

Preface

'Coherent Evolution in Noisy Environments' was the title of an international school at the Max-Planck-Institute for the Physics of Complex Systems in Dresden, Germany, from 2 April to 30 May 2001. Yet, this is also the general theme which a growing community of physicists is contemplating as it comes to monitor, guide, or even control the time evolution of isolated quantum systems. The latter can never be *perfectly* isolated (be it for the purpose of observation) from their environment – which is noisy (since never under complete control) – and will therefore exhibit traces of this, possibly residual, coupling on sufficiently long time scales. However, it is precisely the *coherent* nature of its time evolution which makes an isolated system *quantum*, and it is the detrimental influence of dissipation and of noise fed into the system from the environment, which induces *decoherence* as time evolves.

Given the last two decades' extraordinary progress in the experimental art of isolating single quantum objects (which Schrödinger could only think about in a, by now, famous thought experiment), the theoretical understanding of (de-)coherence and its implications has re-emerged as an important issue of fundamental relevance. Feynman's remarks on the simulation of complex quantum evolution using quantum systems appears to become a more realistic enterprise; moreover quantum cryptography, communication, and computation are identified as emerging key technologies of our young century. If these fields shall guarantee some share holder value on the long run, our theoretical understanding of the coherent evolution of quantum systems in the presence of noise, and, hence, of decoherence, needs a considerable sharpening.

Many and rather distinct subdisciplines of physics and mathematics have their word to say in this context. Whilst quantum opticians arguably come up with the cleanest experimental conditions – which allow for a highly reductionist approach to the quantum world – the condensed matter and mesoscopics communities have to fight with an abundance of imperfections which invites strong input from statistical physics. With some reason one might say that the former can tell us a little more about coherence (and controlled decoherence), whereas the latter are closer to a general theory of decoherence with less stringent simplifications. Nonetheless, both communities are expected to intensify their communication – given the recent realization of simple models of solid state transport theory in quantum optical experiments. Finally, the novel point of view of *quantum information theory* provides a general framework for coherence, decoherence, and

quantum information processing in quantum cryptography, communication, and computation, and receives input from mathematical physics as well as from pure mathematics.

It was the purpose of the school at the MPI-PKS in Dresden to make these different communities listen (and speak) to each other, to convey their different languages, distinct methodologies, and different key challenges to the young and eager in these different fields. If we were able to reach this aim, at least partially, this was the merit of the lecturers of this school. Each of them enthusiastically took on the burden of preparing and delivering between 6 and 10 hours of lectures, sometimes gave additional sessions, and actively participated in the students' seminars and informal discussions, as well as in the other lecturers' courses. Therefore, we should like to thank Hans Briegel, Berge Englert, Gert Ingold, Burkhard Kümmerer, Panagagiotis Lambropoulos, Mark Raizen, Walter Strunz, Steven van Enk, Harald Weinfurter, and Kurt Wiesenfeld, for their crucial contributions to this school.

Towards the end of the event, some of us agreed that the lectures should be conserved, and this idea was strongly encouraged by the two co-organizers of the school, Reinhard Werner and Anton Zeilinger, whom we are very much indebted to for their constructive support in all respects. The result is this present book, which contains a good part of the school lectures, and an additional contribution by Keyl and Werner. It starts out with Ingold's outline of a rather general quantum treatment of dissipation – reflecting the point of view widely spread in the condensed matter and mesoscopics community. Then, Englert and Morigi give a detailed outline of the algebraic treatment of dissipation in the (possibly periodically driven) damped harmonic oscillator, an open quantum system of paradigmatic importance in quantum optics. With the micro-maser as its experimental realization in mind, these lectures constitute – in some respect – the seed for the subsequent chapters by Wiesenfeld et al. and Kümmerer. Both of them deal with stochastic processes, though in rather orthogonal languages, and within rather different contexts. Wiesenfeld et al. discuss the potentially *constructive* role of noise in classical and quantum systems in the presence of some nonlinearity, whilst Kümmerer spells out the mathematical framework of quantum Markov processes. Kümmerer's lecture also provides the general mathematical background of a good part of quantum information theory, and of the corresponding treatment of decoherence as the small system's entanglement with the environment. This latter point of view is elaborated on in Strunz's lecture, which – through the discussion of an experimental realization in the Paris micro-maser setting – is somewhat entangled with the contribution of Englert and Morigi, and rephrases aspects of quantum stochastic calculus already touched upon by Kümmerer. Strunz' treatment of decoherence in phase space is complemented by Aschauer and Briegel, who directly address decoherence in the context of quantum communication, and notably its detrimental influence on quantum entanglement. In particular, they develop efficient strategies to counteract decoherence through the *controlled disentanglement* of the (quantum) carrier from the environment. Finally, again in a more mathematical language, Keyl and Werner

show how quantum data can be protected against decoherence when sent through noisy quantum channels. They also come up with quantitative bounds on the tolerable error rate for such strategies to work.

We believe that this collection of contributions from quite distinct areas nicely illustrates how those areas are slowly getting closer – propelled by some progresses made in physics and mathematics during the last couple of decades – and that we witness how a common language emerges in this exciting area of fundamental research.

Let us finally express our gratitude to all those who made possible the school, and as a direct product thereof this book, through their support and concrete efforts behind the scene: Claudia Poenisch, Christian Caron, Helmut Deggelmann, Torsten Goerke, Heidi Naether, Christa and Klaus Quedenbaum, Andreas Schneider, Hubert Scherrer, Andreas Wagner, Jan-Michael Rost, and the Max-Planck Society.

Dresden and Wien,
August 2002

Andreas Buchleitner
Klaus Hornberger

List of Contributors

Hans Aschauer
Theoretische Physik
Ludwig-Maximilians-Universität
Theresienstr. 37,
D-80333 München, Germany

Andreas Buchleitner
Max-Planck-Institut für Physik
komplexer Systeme
Nöthnitzer Str. 38,
D-01187 Dresden, Germany

Hans J. Briegel
Theoretische Physik
Ludwig-Maximilians-Universität
Theresienstr. 37,
D-80333 München, Germany

Berthold-Georg Englert
Max-Planck-Institut für
Quantenoptik
Hans-Kopfermann-Straße 1,
D-85748 Garching, Germany

Gert-Ludwig Ingold
Institut für Physik
Universität Augsburg,
D-86135 Augsburg, Germany

Michael Keyl
Institut für Mathematische Physik
TU Braunschweig,
Mendelssohnstr. 3,
D-38106 Braunschweig, Germany

Burkhard Kümmerer
Fachbereich Mathematik
Technische Universität Darmstadt
Schloßgartenstraße 7,
D-64289 Darmstadt, Germany

Giovanna Morigi
Max-Planck-Institut für
Quantenoptik
Hans-Kopfermann-Straße 1,
D-85748 Garching, Germany

Walter T. Strunz
Fakultät für Physik
Universität Freiburg
Hermann-Herder-Str. 3,
D-79104 Freiburg, Germany

Thomas Wellens
Max-Planck-Institut für Physik
komplexer Systeme
Nöthnitzer Str. 38,
D-01187 Dresden, Germany

Reinhard F. Werner
Institut für Mathematische Physik
TU Braunschweig,
Mendelssohnstr. 3,
D-38106 Braunschweig, Germany

Kurt Wiesenfeld
Georgia Institute of Technology
Atlanta, GA 30332, USA

Contents

1 Path Integrals and Their Application to Dissipative Quantum Systems

Gert-Ludwig Ingold

Institut für Physik, Universität Augsburg, D-86135 Augsburg, Germany

1.1 Introduction

The coupling of a system to its environment is a recurrent subject in this collection of lecture notes. The consequences of such a coupling are threefold. First of all, energy may irreversibly be transferred from the system to the environment thereby giving rise to the phenomenon of *dissipation*. In addition, the fluctuating force exerted by the environment on the system causes *fluctuations* of the system degree of freedom which manifest itself for example as Brownian motion. While these two effects occur both for classical as well as quantum systems, there exists a third phenomenon which is specific to the quantum world. As a consequence of the entanglement between system and environmental degrees of freedom a coherent superposition of quantum states may be destroyed in a process referred to as *decoherence*. This effect is of major concern if one wants to implement a quantum computer. Therefore, decoherence is discussed in detail in Chap. 5.

Quantum computation, however, is by no means the only topic where the coupling to an environment is relevant. In fact, virtually no real system can be considered as completely isolated from its surroundings. Therefore, the phenomena listed in the previous paragraph play a role in many areas of physics and chemistry and a series of methods has been developed to address this situation. Some approaches like the master equations discussed in Chap. 2 are particularly well suited if the coupling to the environment is weak, a situation desired in quantum computing. On the other hand, in many solid state systems, the environmental coupling can be so strong that weak coupling theories are no longer valid. This is the regime where the path integral approach has proven to be very useful.

It would be beyond the scope of this chapter even to attempt to give a complete overview of the use of path integrals in the description of dissipative quantum systems. In particular for a two-level system coupled to harmonic oscillator degrees of freedom, the so-called spin-boson model, quite a number of approximations have been developed which are useful in their respective parameter regimes. This chapter rather attempts to give an introduction to path integrals for readers unfamiliar with but interested in this method and its application to dissipative quantum systems.

In this spirit, Sect. 1.2 gives an introduction to path integrals. Some aspects discussed in this section are not necessarily closely related to the problem of dissipative systems. They rather serve to illustrate the path integral approach and to convey to the reader the beauty and power of this approach. In Sect. 1.3

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 1–53, 2002.

we elaborate on the general idea of the coupling of a system to an environment. The path integral formalism is employed to eliminate the environmental degrees of freedom and thus to obtain an effective description of the system degree of freedom. The results provide the basis for a discussion of the damped harmonic oscillator in Sect. 1.4. Starting from the partition function we will examine several aspects of this dissipative quantum system.

Readers interested in a more in-depth treatment of the subject of quantum dissipation are referred to existing textbooks. In particular, we recommend the book by U. Weiss [1] which provides an extensive presentation of this topic together with a comprehensive list of references. Chapter 4 of [2] may serve as a more concise introduction complementary to the present chapter. Path integrals are discussed in a whole variety of textbooks with an emphasis either on the physical or the mathematical aspects. We only mention the book by H. Kleinert [3] which gives a detailed discussion of path integrals and their applications in different areas.

1.2 Path Integrals

1.2.1 Introduction

The most often used and taught approach to nonrelativistic quantum mechanics is based on the Schrödinger equation which possesses strong ties with the the Hamiltonian formulation of classical mechanics. The nonvanishing Poisson brackets between position and momentum in classical mechanics lead us to introduce noncommuting operators in quantum mechanics. The Hamilton function turns into the Hamilton operator, the central object in the Schrödinger equation. One of the most important tasks is to find the eigenfunctions of the Hamilton operator and the associated eigenvalues. Decomposition of a state into these eigenfunctions then allows us to determine its time evolution.

As an alternative, there exists a formulation of quantum mechanics based on the Lagrange formalism of classical mechanics with the action as the central concept. This approach, which was developed by Feynman in the 1940's [4,5], avoids the use of operators though this does not necessarily mean that the solution of quantum mechanical problems becomes simpler. Instead of finding eigenfunctions of a Hamiltonian one now has to evaluate a functional integral which directly yields the propagator required to determine the dynamics of a quantum system. Since the relation between Feynman's formulation and classical mechanics is very close, the path integral formalism often has the important advantage of providing a much more intuitive approach as we will try to convey to the reader in the following sections.

1.2.2 Propagator

In quantum mechanics, one often needs to determine the solution $|\psi(t)\rangle$ of the time-dependent Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle \ , \tag{1.1}$$

where H is the Hamiltonian describing the system. Formally, the solution of (1.1) may be written as

$$|\psi(t)\rangle = \mathcal{T} \exp\left(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t' H(t')\right)|\psi(0)\rangle \; . \tag{1.2}$$

Here, the time ordering operator $\mathcal{T}$ is required because the operators corresponding to the Hamiltonian at different times do not commute in general. In the following, we will restrict ourselves to time-independent Hamiltonians where (1.2) simplifies to

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}Ht\right)|\psi(0)\rangle \; . \tag{1.3}$$

As the inspection of (1.2) and (1.3) demonstrates, the solution of the time-dependent Schrödinger equation contains two parts: the initial state $|\psi(0)\rangle$ which serves as an initial condition and the so-called propagator, an operator which contains all information required to determine the time evolution of the system.

Writing (1.3) in position representation one finds

$$\langle x|\psi(t)\rangle = \int \mathrm{d}x' \langle x| \exp\left(-\frac{\mathrm{i}}{\hbar}Ht\right)|x'\rangle\langle x'|\psi(0)\rangle \tag{1.4}$$

or

$$\psi(x,t) = \int \mathrm{d}x' K(x,t,x',0)\psi(x',0) \tag{1.5}$$

with the propagator

$$K(x,t,x',0) = \langle x| \exp\left(-\frac{\mathrm{i}}{\hbar}Ht\right)|x'\rangle \; . \tag{1.6}$$

It is precisely this propagator which is the central object of Feynman's formulation of quantum mechanics. Before discussing the path integral representation of the propagator, it is therefore useful to take a look at some properties of the propagator.

Instead of performing the time evolution of the state $|\psi(0)\rangle$ into $|\psi(t)\rangle$ in one step as was done in equation (1.3), one could envisage to perform this procedure in two steps by first propagating the initial state $|\psi(0)\rangle$ up to an intermediate time t_1 and taking the new state $|\psi(t_1)\rangle$ as initial state for a propagation over the time $t - t_1$. This amounts to replacing (1.3) by

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}H(t-t_1)\right)\exp\left(-\frac{\mathrm{i}}{\hbar}Ht_1\right)|\psi(0)\rangle \tag{1.7}$$

or equivalently

$$\psi(x,t) = \int \mathrm{d}x' \int \mathrm{d}x'' K(x,t,x'',t_1)K(x'',t_1,x',0)\psi(x',0) \; . \tag{1.8}$$

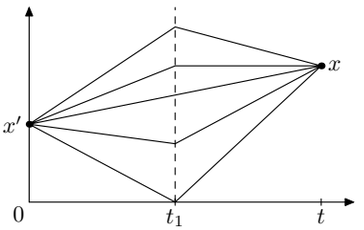


Fig. 1.1. According to the semigroup property (1.9) the propagator $K(x,t,x',0)$ may be decomposed into propagators arriving at some time t_1 at an intermediate point x'' and propagators continuing from there to the final point x

Comparing (1.5) and (1.8), we find the semigroup property of the propagator

$$K(x,t,x',0)=\int \mathrm{d}x'' K(x,t,x'',t_1)K(x'',t_1,x',0)\ . \tag{1.9}$$

This result is visualized in Fig. 1.1 where the propagators between space-time points are depicted by straight lines connecting the corresponding two points. At the intermediate time t_1 one has to integrate over all positions x''. This insight will be of use when we discuss the path integral representation of the propagator later on.

The propagator contains the complete information about the eigenenergies E_n and the corresponding eigenstates $|n\rangle$. Making use of the completeness of the eigenstates, one finds from (1.6)

$$K(x,t,x',0)=\sum_n \exp\left(-\frac{\mathrm{i}}{\hbar}E_n t\right)\psi_n(x)\psi_n(x')^*\ . \tag{1.10}$$

Here, the star denotes complex conjugation. Not only does the propagator contain the eigenenergies and eigenstates, this information may also be extracted from it. To this end, we introduce the retarded Green function

$$G_\mathrm{r}(x,t,x',0)=K(x,t,x',0)\Theta(t) \tag{1.11}$$

where $\Theta(t)$ is the Heaviside function which equals 1 for positive argument t and is zero otherwise. Performing a Fourier transformation, one ends up with the spectral representation

$$\begin{aligned} G_\mathrm{r}(x,x',E) &= -\frac{\mathrm{i}}{\hbar}\int_0^\infty \mathrm{d}t \exp\left(\frac{\mathrm{i}}{\hbar}Et\right) G_\mathrm{r}(t) \\ &= \sum_n \frac{\psi_n(x)\psi_n(x')^*}{E-E_n+\mathrm{i}\varepsilon}\ , \end{aligned} \tag{1.12}$$

where ε is an infinitely small positive quantity. According to (1.12), the poles of the energy-dependent retarded Green function indicate the eigenenergies while the corresponding residua can be factorized into the eigenfunctions at positions x and x'.

1.2.3 Free Particle

An important step towards the path integral formulation of quantum mechanics can be made by considering the propagator of a free particle of mass m. The eigenstates of the corresponding Hamiltonian

$$H = \frac{p^2}{2m} \tag{1.13}$$

are momentum eigenstates

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{\mathrm{i}}{\hbar} px\right) \tag{1.14}$$

with a momentum eigenvalue p out of a continuous spectrum. Inserting these eigenstates into the representation (1.10) of the propagator, one finds by virtue of

$$\int_{-\infty}^{\infty} \mathrm{d}x \exp(-\mathrm{i}ax^2) = \sqrt{\frac{\pi}{\mathrm{i}a}} = \sqrt{\frac{\pi}{a}} \exp\left(-\mathrm{i}\frac{\pi}{4}\right) \tag{1.15}$$

for the propagator of the free particle the result

$$\begin{aligned} K(x_\mathrm{f}, t, x_\mathrm{i}, 0) &= \frac{1}{2\pi\hbar} \int \mathrm{d}p \exp\left(-\frac{\mathrm{i}}{\hbar}\frac{p^2}{2m}t\right) \exp\left(\frac{\mathrm{i}}{\hbar}p(x_\mathrm{f} - x_\mathrm{i})\right) \\ &= \sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}} \exp\left(\frac{\mathrm{i}}{\hbar}\frac{m(x_\mathrm{f} - x_\mathrm{i})^2}{2t}\right) . \end{aligned} \tag{1.16}$$

It was already noted by Dirac [6] that the quantum mechanical propagator and the classical properties of a free particle are closely related. In order to demonstrate this, we evaluate the action of a particle moving from x_i to x_f in time t. From the classical path

$$x_\mathrm{cl}(s) = x_\mathrm{i} + (x_\mathrm{f} - x_\mathrm{i})\frac{s}{t} \tag{1.17}$$

obeying the boundary conditions $x_\mathrm{cl}(0) = x_\mathrm{i}$ and $x_\mathrm{cl}(t) = x_\mathrm{f}$, the corresponding classical action is found as

$$S_\mathrm{cl} = \frac{m}{2}\int_0^t \mathrm{d}s \dot{x}_\mathrm{cl}^2 = \frac{m}{2}\frac{(x_\mathrm{f} - x_\mathrm{i})^2}{t} . \tag{1.18}$$

This result enables us to express the propagator of a free particle entirely in terms of the classical action as

$$K(x_\mathrm{f}, t, x_\mathrm{i}, 0) = \left(-\frac{1}{2\pi\mathrm{i}\hbar}\frac{\partial^2 S_\mathrm{cl}(x_\mathrm{f}, t, x_\mathrm{i}, 0)}{\partial x_\mathrm{f}\partial x_\mathrm{i}}\right)^{1/2} \exp\left(\frac{\mathrm{i}}{\hbar}S_\mathrm{cl}(x_\mathrm{f}, t, x_\mathrm{i}, 0)\right) . \tag{1.19}$$

This result is quite remarkable and one might suspect that it is due to a peculiarity of the free particle. However, since the propagation in a general potential (in the absence of delta function contributions) may be decomposed into a series of short-time propagations of a free particle, the result (1.19) may indeed be employed to construct a representation of the propagator where the classical action appears in the exponent. In the prefactor, the action appears in the form shown in (1.19) only within the semiclassical approximation (cf. Sect. 1.2.8) or for potentials where this approximation turns out to be exact.

1.2.4 Path Integral Representation of Quantum Mechanics

While avoiding to go too deeply into the mathematical details, we nevertheless want to sketch the derivation of the path integral representation of the propagator. The main idea is to decompose the time evolution over a finite time t into N slices of short time intervals $\Delta t = t/N$ where we will eventually take the limit $N \to \infty$. Denoting the operator of the kinetic and potential energy by T and V, respectively, we thus find

$$\exp\left(-\frac{\mathrm{i}}{\hbar}Ht\right) = \left[\exp\left(-\frac{\mathrm{i}}{\hbar}(T+V)\Delta t\right)\right]^N . \tag{1.20}$$

For simplicity, we will assume that the Hamiltonian is time-independent even though the following derivation may be generalized to the time-dependent case. We now would like to decompose the short-time propagator in (1.20) into a part depending on the kinetic energy and another part containing the potential energy. However, since the two operators do not commute, we have to exercise some caution. From an expansion of the Baker-Hausdorff formula one finds

$$\exp\left(-\frac{\mathrm{i}}{\hbar}(T+V)\Delta t\right) \approx \exp\left(-\frac{\mathrm{i}}{\hbar}T\Delta t\right)\exp\left(-\frac{\mathrm{i}}{\hbar}V\Delta t\right) + \frac{1}{\hbar^2}[T,V](\Delta t)^2 \tag{1.21}$$

where terms of order $(\Delta t)^3$ and higher have been neglected. Since we are interested in the limit $\Delta t \to 0$, we may neglect the contribution of the commutator and arrive at the Trotter formula

$$\exp\left(-\frac{\mathrm{i}}{\hbar}(T+V)t\right) = \lim_{N\to\infty}\left[U(\Delta t)\right]^N \tag{1.22}$$

with the short time evolution operator

$$U(\Delta t) = \exp\left(-\frac{\mathrm{i}}{\hbar}T\Delta t\right)\exp\left(-\frac{\mathrm{i}}{\hbar}V\Delta t\right) . \tag{1.23}$$

What we have presented here is, of course, at best a motivation and certainly does not constitute a mathematical proof. We refer readers interested in the details of the proof and the conditions under which the Trotter formula holds to the literature [7].

In position representation one now obtains for the propagator

$$K(x_\mathrm{f},t,x_\mathrm{i},0) = \lim_{N\to\infty}\int_{-\infty}^{\infty}\left(\prod_{j=1}^{N-1}\mathrm{d}x_j\right)\langle x_\mathrm{f}\,|U(\Delta t)|\,x_{N-1}\rangle\ldots \times\langle x_1\,|U(\Delta t)|\,x_\mathrm{i}\rangle . \tag{1.24}$$

Since the potential is diagonal in position representation, one obtains together with the expression (1.16) for the propagator of the free particle for the matrix

element

$$\begin{aligned}\langle x_{j+1}\,|U(\Delta t)|\,x_j\rangle &= \left\langle x_{j+1}\left|\exp\left(-\frac{\mathrm{i}}{\hbar}T\Delta t\right)\right|x_j\right\rangle \exp\left(-\frac{\mathrm{i}}{\hbar}V(x_j)\Delta t\right)\\ &= \sqrt{\frac{m}{2\pi\mathrm{i}\hbar\Delta t}}\exp\left[\frac{\mathrm{i}}{\hbar}\left(\frac{m}{2}\frac{(x_{j+1}-x_j)^2}{\Delta t}-V(x_j)\Delta t\right)\right]\ .\end{aligned} \tag{1.25}$$

We thus arrive at our final version of the propagator

$$\begin{aligned}K(x_\mathrm{f},t,x_\mathrm{i},0) = \lim_{N\to\infty}\sqrt{\frac{m}{2\pi\mathrm{i}\hbar\Delta t}}\int_{-\infty}^{\infty}\left(\prod_{j=1}^{N-1}\mathrm{d}x_j\sqrt{\frac{m}{2\pi\mathrm{i}\hbar\Delta t}}\right)\\ \times\exp\left[\frac{\mathrm{i}}{\hbar}\sum_{j=0}^{N-1}\left(\frac{m}{2}\left(\frac{x_{j+1}-x_j}{\Delta t}\right)^2-V(x_j)\right)\Delta t\right]\end{aligned} \tag{1.26}$$

where x_0 and x_N should be identified with x_i and x_f, respectively. The discretization of the propagator used in this expression is a consequence of the form (1.21) of the Baker-Hausdorff relation. In lowest order in Δt, we could have used a different decomposition which would have led to a different discretization of the propagator. For a discussion of the mathematical subtleties we refer the reader to [8].

Remarking that the exponent in (1.26) contains a discretized version of the action

$$S[x] = \int_0^t \mathrm{d}s\left(\frac{m}{2}\dot{x}^2 - V(x)\right)\ , \tag{1.27}$$

we can write this result in short notation as

$$K(x_\mathrm{f},t,x_\mathrm{i},0) = \int \mathcal{D}x\exp\left(\frac{\mathrm{i}}{\hbar}S[x]\right)\ . \tag{1.28}$$

The action (1.27) is a functional which takes as argument a function $x(s)$ and returns a number, the action $S[x]$. The integral in (1.28) therefore is a functional integral where one has to integrate over all functions satisfying the boundary conditions $x(0) = x_\mathrm{i}$ and $x(t) = x_\mathrm{f}$. Since these functions represent paths, one refers to this kind of functional integrals also as path integral.

The three lines shown in Fig. 1.2 represent the infinity of paths satisfying the boundary conditions. Among them the thicker line indicates a special path corresponding to an extremum of the action. According to the principal of least action such a path is a solution of the classical equation of motion. It should be noted, however, that even though sometimes there exists a unique extremum, in general there may be more than one or even none. A demonstration of this fact will be provided in Sect. 1.2.7 where we will discuss the driven harmonic oscillator.

The other paths depicted in Fig. 1.2 may be interpreted as quantum fluctuations around the classical path. As we will see in Sect. 1.2.8, the amplitude of

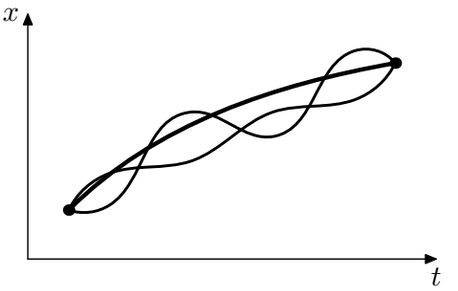


Fig. 1.2. The thick line represents a classical path satisfying the boundary conditions. The thinner lines are no solutions of the classical equation of motion and may be associated with quantum fluctuations

these fluctuations is typically of the order of $\sqrt{\hbar}$. In the classical limit $\hbar \to 0$ therefore only the classical paths survive as one should expect.

Before explicitly evaluating a path integral, we want to discuss two examples which will give us some insight into the difference of the approaches offered by the Schrödinger and Feynman formulation of quantum mechanics.

1.2.5 Particle on a Ring

We confine a particle of mass m to a ring of radius R and denote its angular degree of freedom by ϕ. This system is described by the Hamiltonian

$$H = -\frac{\hbar^2}{2mR^2}\frac{\partial^2}{\partial\phi^2} \; . \tag{1.29}$$

Requiring the wave function to be continuous and differentiable, one finds the stationary states

$$\psi_\ell(\phi) = \frac{1}{\sqrt{2\pi}} \exp\left(\mathrm{i}\ell\phi\right) \tag{1.30}$$

with $\ell = 0, \pm 1, \pm 2, \ldots$ and the eigenenergies

$$E_\ell = \frac{\hbar^2\ell^2}{2mR^2} \; . \tag{1.31}$$

These solutions of the time-independent Schrödinger equation allow us to construct the propagator

$$K(\phi_\mathrm{f}, t, \phi_\mathrm{i}, 0) = \frac{1}{2\pi}\sum_{\ell=-\infty}^{\infty} \exp\left(\mathrm{i}\ell(\phi_\mathrm{f} - \phi_\mathrm{i}) - \mathrm{i}\frac{\hbar\ell^2}{2mR^2}t\right) \; . \tag{1.32}$$

We now want to derive this result within the path integral formalism. To this end we will employ the propagator of the free particle. However, an important difference between a free particle and a particle on a ring deserves our attention.

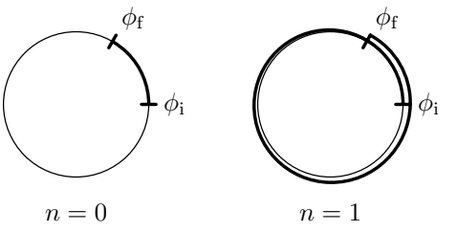


Fig. 1.3. On a ring, the angles ϕ_f and $\phi_\mathrm{f}+2\pi n$ have to be identified. As a consequence, there exist infinitely many classical paths connecting two points on a ring, which may be identified by their winding number n

Due to the ring topology we have to identify all angles $\phi + 2\pi n$, where n is an integer, with the angle ϕ. As a consequence, there exist infinitely many classical paths connecting ϕ_i and ϕ_f. All these paths are topologically different and can be characterized by their winding number n. As an example, Fig. 1.3 shows a path for $n = 0$ and $n = 1$. Due to their different topology, these two paths (and any two paths corresponding to different winding numbers) cannot be continuously transformed into each other. This implies that adding a fluctuation to one of the classical paths will never change its winding number.

Therefore, we have to sum over all winding numbers in order to account for all possible paths. The propagator thus consists of a sum over free propagators corresponding to different winding numbers

$$K(\phi_\mathrm{f},t,\phi_\mathrm{i},0) = \sum_{n=-\infty}^{\infty} R\sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}}\exp\left(\frac{\mathrm{i}}{\hbar}\frac{mR^2}{2}\frac{(\phi_\mathrm{f}-\phi_\mathrm{i}-2\pi n)^2}{t}\right) . \tag{1.33}$$

Here, the factor R accounts for the fact that, in contrast to the free particle, the coordinate is given by an angle instead of a position.

The propagator (1.33) is 2π-periodic in $\phi_\mathrm{f}-\phi_\mathrm{i}$ and can therefore be expressed in terms of a Fourier series

$$K(\phi_\mathrm{f},t,\phi_\mathrm{i},0) = \sum_{\ell=-\infty}^{\infty} c_\ell \exp\left[\mathrm{i}\ell(\phi_\mathrm{f}-\phi_\mathrm{i})\right] . \tag{1.34}$$

The Fourier coefficients are found to read

$$c_\ell = \frac{1}{2\pi}\exp\left(-\mathrm{i}\frac{\hbar\ell^2}{2mR^2}t\right) \tag{1.35}$$

which proves the equivalence of (1.33) with our previous result (1.32). We thus have obtained the propagator of a free particle on a ring both by solving the Schrödinger equation and by employing path integral methods. These two approaches make use of complementary representations. In the first case, this is the angular momentum representation while in the second case, one works in the phase representation and sums over winding numbers.

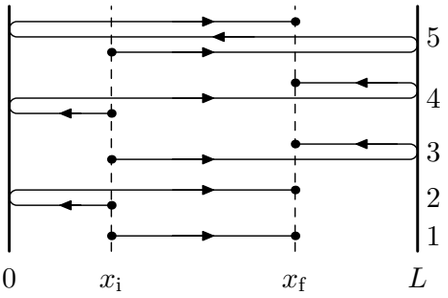


Fig. 1.4. The reflection at the walls of a box leads to an infinite number of possible trajectories connecting two points in the box

1.2.6 Particle in a Box

Another textbook example in standard quantum mechanics is the particle in a box of length L confined by infinitely high walls at $x = 0$ and $x = L$. From the eigenvalues

$$E_j = \frac{\hbar^2\pi^2 j^2}{2mL^2} \tag{1.36}$$

with $j = 1, 2, \ldots$ and the corresponding eigenfunctions

$$\psi_j(x) = \sqrt{\frac{2}{L}} \sin\left(\pi j \frac{x}{L}\right) \tag{1.37}$$

the propagator is immediately obtained as

$$K(x_\mathrm{f}, t, x_\mathrm{i}, 0) = \frac{2}{L}\sum_{j=1}^{\infty} \exp\left(-\mathrm{i}\frac{\hbar\pi^2 j^2}{2mL^2}t\right) \sin\left(\pi j \frac{x_\mathrm{f}}{L}\right) \sin\left(\pi j \frac{x_\mathrm{i}}{L}\right) \ . \tag{1.38}$$

It took some time until this problem was solved within the path integral approach [9,10]. Here, we have to consider all paths connecting the points x_i and x_f within a period of time t. Due to the reflecting walls, there again exist infinitely many classical paths, five of which are depicted in Fig. 1.4. However, in contrast to the case of a particle on a ring, these paths are no longer topologically distinct. As a consequence, we may deform a classical path continuously to obtain one of the other classical paths.

If, for the moment, we disregard the details of the reflections at the wall, the motion of the particle in a box is equivalent to the motion of a free particle. The fact that paths are folded back on themselves can be accounted for by taking into account replicas of the box as shown in Fig. 1.5. Now, the path does not necessarily end at $x_\mathrm{f}^{(0)} = x_\mathrm{f}$ but at one of the mirror images $x_\mathrm{f}^{(n)}$ where n is an arbitrary integer. In order to obtain the propagator, we will have to sum over all replicas. Due to the different geometry we need to distinguish between those paths arising from an even and an odd number of reflections. From Fig. 1.5 one

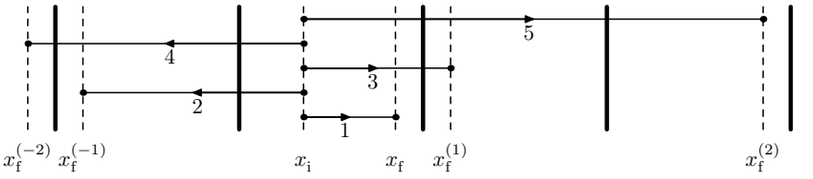


Fig. 1.5. Instead of a particle getting reflected at the walls of the box one may think of a free particle moving from the starting point in the box to the end point in one of the replicas of the box

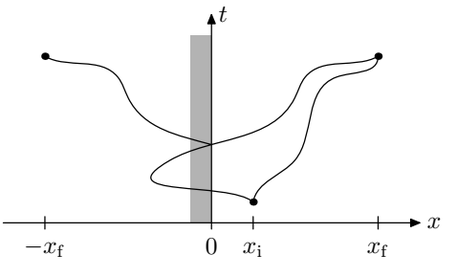


Fig. 1.6. A path crossing the wall is cancelled by a path running to the mirror point of the end point

can see that for an odd number $2n-1$ of reflections, the end point lies at $2nL-x_{\rm f}$ and the contribution to the full propagator therefore is given by

$$K^{(2n-1)}(x_{\rm f},t,x_{\rm i},0)=\sqrt{\frac{m}{2\pi{\rm i}\hbar t}}\exp\left(\frac{\rm i}{\hbar}\frac{m(2nL-x_{\rm f}-x_{\rm i})^2}{2t}\right)\ . \tag{1.39}$$

On the other hand, for an even number $2n$ of reflections, the end point is located at $2nL+x_{\rm f}$ and we find

$$K^{(2n)}(x_{\rm f},t,x_{\rm i},0)=\sqrt{\frac{m}{2\pi{\rm i}\hbar t}}\exp\left(\frac{\rm i}{\hbar}\frac{m(2nL+x_{\rm f}-x_{\rm i})^2}{2t}\right)\ . \tag{1.40}$$

However, it is not obvious that just summing up the propagators (1.39) and (1.40) for all n will do the job.

In order to clarify this point, we start with the somewhat simpler situation of just one wall and take a look at all paths running between $x_{\rm i}$ and $x_{\rm f}$ in time t. As can be seen from the space-time diagram in Fig. 1.6 there are paths which do not cross the wall and which therefore contribute to the path integral. On the other hand, there exist also paths which cross the wall an even number of times. Since these paths spend some time in the forbidden region, they do not contribute to the path integral.

It requires some thinking to ensure that only paths not crossing the wall are taken into account. Our strategy will consist in first writing down a propagator $K_{\rm free}$ which disregards the wall. Then, we have to subtract off the contributions of all the paths which cross the wall. This can be done by constructing a path

with the same action as the original path. To this end we take the original path up to the last crossing with the wall and then continue along the mirror image of the original path. We thus end up at the mirror image $-x_\mathrm{f}$ of the original end point x_f. Note that a path running from x_i to $-x_\mathrm{f}$ necessarily crosses the wall at least once. As a consequence, subtracting the propagator between these two points eliminates all original paths which do not remain in the region $x > 0$. We therefore obtain our desired result, the propagator K_wall in the presence of a wall, by subtracting a propagator going to the reflected end point from the unconstrained propagator to the original end point [9,10,11]

$$K_\mathrm{wall}(x_\mathrm{f},t,x_\mathrm{i},0) = K_\mathrm{free}(x_\mathrm{f},t,x_\mathrm{i},0) - K_\mathrm{free}(-x_\mathrm{f},t,x_\mathrm{i},0)\ . \tag{1.41}$$

This result bears much resemblance with the method of image charges in electrostatics. After giving it some thought, this should not be too surprising since the free Schrödinger equation and the Poisson equation are formally equivalent. According to the method of image charges one may account for a metallic plate (i.e. the wall) by putting a negative charge (i.e. the mirrored end point) complementing the positive charge (i.e. the original end point). For the propagator this results in the difference appearing in (1.41).

Let us now come back to our infinitely deep potential well with two walls. This problem corresponds to the electrostatics of a charge between two parallel metal plates. In this case, the method of image charges leads to an infinite number of charges of alternating signs. The original positive charge gives rise to two negative charges which are each an image corresponding to one of the two metal plates. In addition, however, these images have mirror images corresponding to the other metal plate and this process has to be carried on ad infinitum.

Expressing the propagator of the particle in the box in terms of the free propagator works in exactly the same way. A path intersecting both walls is subtracted twice, i.e. one time too often. Therefore, one contribution has to be restored which is done by adding another end point. Continuing the procedure one ends up with an infinite number of end points, some of which we have shown in Fig. 1.5. As a consequence, we can attribute a sign to each end point in this figure. The general rule which follows from these considerations is that each reflection at a wall leads to factor -1. The propagator therefore can be written as

$$\begin{aligned} K(x_\mathrm{f},t,x_\mathrm{i},0) = \sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}} \sum_{n=-\infty}^{\infty} \Bigg[&\exp\left(\frac{\mathrm{i}}{\hbar}\frac{m(2nL+x_\mathrm{f}-x_\mathrm{i})^2}{2t}\right) \\ &- \exp\left(\frac{\mathrm{i}}{\hbar}\frac{m(2nL-x_\mathrm{f}-x_\mathrm{i})^2}{2t}\right)\Bigg]\ . \end{aligned} \tag{1.42}$$

The symmetries

$$K(x_\mathrm{f}+2L,t,x_\mathrm{i},0) = K(x_\mathrm{f},t,x_\mathrm{i},0) \tag{1.43}$$

$$K(-x_\mathrm{f},t,x_\mathrm{i},0) = -K(x_\mathrm{f},t,x_\mathrm{i},0) \tag{1.44}$$

suggest to expand the propagator into the Fourier series

$$K(x_\mathrm{f},t,x_\mathrm{i},0)=\sum_{j=1}^{\infty}a_j(x_\mathrm{i},t)\sin\left(\pi j\frac{x_\mathrm{f}}{L}\right)\ . \tag{1.45}$$

Its Fourier coefficients are obtained from (1.42) as

$$\begin{aligned}a_j(x_\mathrm{i},t)&=\frac{1}{L}\int_{-L}^{L}\mathrm{d}x_\mathrm{f}\sin\left(\pi j\frac{x_\mathrm{f}}{L}\right)K(x_\mathrm{f},t,x_\mathrm{i},0)\\&=\frac{2}{L}\sin\left(\pi j\frac{x_\mathrm{i}}{L}\right)\exp\left(-\frac{\mathrm{i}}{\hbar}E_jt\right)\end{aligned} \tag{1.46}$$

where the energies E_j are the eigenenergies of the box defined in (1.36). Inserting (1.46) into (1.45) we thus recover our previous result (1.38).

1.2.7 Driven Harmonic Oscillator

Even though the situations dealt with in the previous two sections have been conceptually quite interesting, we could in both cases avoid the explicit calculation of a path integral. In the present section, we will introduce the basic techniques needed to evaluate path integrals. As an example, we will consider the driven harmonic oscillator which is simple enough to allow for an exact solution. In addition, the propagator will be of use in the discussion of damped quantum systems in later sections.

Our starting point is the Lagrangian

$$L=\frac{m}{2}\dot{x}^2-\frac{m}{2}\omega^2x^2+xf(t) \tag{1.47}$$

of a harmonic oscillator with mass m and frequency ω. The force $f(t)$ may be due to an external field, e.g. an electric field coupling via dipole interaction to a charged particle. In the context of dissipative quantum mechanics, the harmonic oscillator could represent a degree of freedom of the environment under the influence of a force exerted by the system.

According to (1.28) we obtain the propagator $K(x_\mathrm{f},t,x_\mathrm{i},0)$ by calculating the action for all possible paths starting at time zero at x_i and ending at time t at x_f. It is convenient to decompose the general path

$$x(s)=x_\mathrm{cl}(s)+\xi(s) \tag{1.48}$$

into the classical path x_cl satisfying the boundary conditions $x_\mathrm{cl}(0)=x_\mathrm{i}$, $x_\mathrm{cl}(t)=x_\mathrm{f}$ and a fluctuating part ξ vanishing at the boundaries, i.e. $\xi(0)=\xi(t)=0$. The classical path has to satisfy the equation of motion

$$m\ddot{x}_\mathrm{cl}+m\omega^2x_\mathrm{cl}=f(s) \tag{1.49}$$

obtained from the Lagrangian (1.47).

For an exactly solvable problem like the driven harmonic oscillator, we could replace $x_{\rm cl}$ by any path satisfying $x(0) = x_{\rm i}$, $x(t) = x_{\rm f}$. We leave it as an exercise to the reader to perform the following calculation with $x_{\rm cl}(s)$ of the driven harmonic oscillator replaced by $x_{\rm i}+(x_{\rm f}-x_{\rm i})s/t$. However, it is important to note that within the semiclassical approximation discussed in Sect. 1.2.8 an expansion around the classical path is essential since this path leads to the dominant contribution to the path integral.

With (1.48) we obtain for the action

$$\begin{aligned} S &= \int_0^t {\rm d}s \left(\frac{m}{2}\dot{x}^2 - \frac{m}{2}\omega^2 x^2 + xf(s)\right) \\ &= \int_0^t {\rm d}s \left(\frac{m}{2}\dot{x}_{\rm cl}^2 - \frac{m}{2}\omega^2 x_{\rm cl}^2 + x_{\rm cl}f(s)\right) + \int_0^t {\rm d}s \left(m\dot{x}_{\rm cl}\dot{\xi} - m\omega^2 x_{\rm cl}\xi + \xi f(s)\right) \\ &\quad + \int_0^t {\rm d}s \left(\frac{m}{2}\dot{\xi}^2 - \frac{m}{2}\omega^2\xi^2\right). \end{aligned} \tag{1.50}$$

For our case of a harmonic potential, the third term is independent of the boundary values $x_{\rm i}$ and $x_{\rm f}$ as well as of the external driving. The second term vanishes as a consequence of the expansion around the classical path. This can be seen by partial integration and by making use of the fact that $x_{\rm cl}$ is a solution of the classical equation of motion:

$$\int_0^t {\rm d}s\big(m\dot{x}_{\rm cl}\dot{\xi} - m\omega^2 x_{\rm cl}\xi + \xi f(s)\big) = -\int_0^t {\rm d}s\big(m\ddot{x}_{\rm cl} + m\omega^2 x_{\rm cl} - f(s)\big)\xi = 0\,. \tag{1.51}$$

We now proceed in two steps by first determining the contribution of the classical path and then addressing the fluctuations. The solution of the classical equation of motion satisfying the boundary conditions reads

$$\begin{aligned} x_{\rm cl}(s) &= x_{\rm f}\frac{\sin(\omega s)}{\sin(\omega t)} + x_{\rm i}\frac{\sin(\omega(t-s))}{\sin(\omega t)} \\ &\quad + \frac{1}{m\omega}\left[\int_0^s {\rm d}u\,\sin(\omega(s-u))f(u) - \frac{\sin(\omega s)}{\sin(\omega t)}\int_0^t {\rm d}u\,\sin(\omega(t-u))f(u)\right]. \end{aligned} \tag{1.52}$$

A peculiarity of the harmonic oscillator in the absence of driving is the appearance of conjugate points at times $T_n = (\pi/\omega)n$ where n is an arbitrary integer. Since the frequency of the oscillations is independent of the amplitude, the position of the oscillator at these times is determined by the initial position: $x(T_{2n+1}) = -x_{\rm i}$ and $x(T_{2n}) = x_{\rm i}$, see Fig. 1.7. This also illustrates the fact mentioned on p. 7, that depending on the boundary conditions there may be more than one or no classical solution.

The task of evaluating the action of the classical path may be simplified by a partial integration

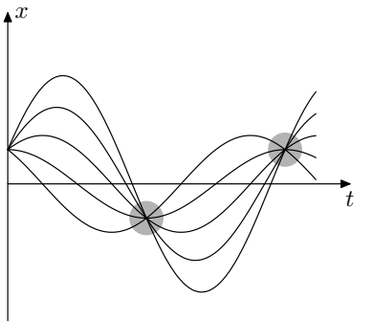


Fig. 1.7. In a harmonic potential all trajectories emerging from the same starting point converge at conjugate points at multiples of half an oscillation period

$$\begin{aligned} S_{\rm cl} &= \int_0^t {\rm d}s \left(\frac{m}{2}\dot{x}_{\rm cl}^2 - \frac{m}{2}\omega^2 x_{\rm cl}^2 + x_{\rm cl} f(s) \right) \\ &= \frac{m}{2} x_{\rm cl}\dot{x}_{\rm cl}\Big|_0^t - \int_0^t {\rm d}s \left(\frac{m}{2} x_{\rm cl}\ddot{x}_{\rm cl} + \frac{m}{2}\omega^2 x_{\rm cl}^2 - x_{\rm cl} f(s) \right) \\ &= \frac{m}{2} \left(x_{\rm f}\dot{x}_{\rm cl}(t) - x_{\rm i}\dot{x}_{\rm cl}(0) \right) + \frac{1}{2}\int_0^t {\rm d}s\, x_{\rm cl}(s) f(s) \end{aligned} \tag{1.53}$$

where we have made use of the classical equation of motion to obtain the third line. From the solution (1.52) of the classical equation of motion we get

$$\dot{x}_{\rm cl}(0) = \omega \frac{x_{\rm f} - x_{\rm i}\cos(\omega t)}{\sin(\omega t)} - \frac{1}{m\sin(\omega t)} \int_0^t {\rm d}s \sin(\omega(t-s)) f(s) \tag{1.54}$$

$$\dot{x}_{\rm cl}(t) = \omega \frac{x_{\rm f}\cos(\omega t) - x_{\rm i}}{\sin(\omega t)} + \frac{1}{m\sin(\omega t)} \int_0^t {\rm d}s \sin(\omega s) f(s)\ . \tag{1.55}$$

Inserting initial and final velocity into (1.53) we find for the classical action

$$\begin{aligned} S_{\rm cl} =& \frac{m\omega}{2\sin(\omega t)} \left[(x_{\rm i}^2 + x_{\rm f}^2)\cos(\omega t) - 2x_{\rm i}x_{\rm f} \right] \\ &+ \frac{x_{\rm f}}{\sin(\omega t)} \int_0^t {\rm d}s \sin(\omega s) f(s) + \frac{x_{\rm i}}{\sin(\omega t)} \int_0^t {\rm d}s \sin(\omega(t-s)) f(s) \\ &- \frac{1}{m\omega\sin(\omega t)} \int_0^t {\rm d}s \int_0^s {\rm d}u \sin(\omega u) \sin(\omega(t-s)) f(s) f(u)\ . \end{aligned} \tag{1.56}$$

As a second step, we have to evaluate the contribution of the fluctuations which is determined by the third term in (1.50). After partial integration this term becomes

$$S^{(2)} = \int_0^t {\rm d}s \left(\frac{m}{2}\dot{\xi}^2 - \frac{m}{2}\omega^2\xi^2 \right) = -\int_0^t {\rm d}s \frac{m}{2}\xi \left(\frac{{\rm d}^2}{{\rm d}s^2} + \omega^2 \right) \xi\ . \tag{1.57}$$

Here, the superscript '(2)' indicates that this term corresponds to the contribution of second order in ξ. In view of the right-hand side it is appropriate to expand the fluctuation

$$\xi(s) = \sum_{n=1}^{\infty} a_n \xi_n(s) \tag{1.58}$$

into eigenfunctions of

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}s^2} + \omega^2\right) \xi_n = \lambda_n \xi_n \tag{1.59}$$

with $\xi_n(0) = \xi_n(t) = 0$. As eigenfunctions of a selfadjoint operator, the ξ_n are complete and may be chosen orthonormal. Solving (1.59) yields the eigenfunctions

$$\xi_n(s) = \sqrt{\frac{2}{t}} \sin\left(\pi n \frac{s}{t}\right) \tag{1.60}$$

and corresponding eigenvalues

$$\lambda_n = -\left(\frac{\pi n}{t}\right)^2 + \omega^2 \,. \tag{1.61}$$

We emphasize that (1.58) is not the usual Fourier series on an interval of length t. Such an expansion could be used in the form

$$\xi(s) = \sqrt{\frac{2}{t}} \sum_{n=1}^{\infty} \left[a_n \left(\cos\left(2\pi n \frac{s}{t}\right) - 1\right) + b_n \sin\left(2\pi n \frac{s}{t}\right)\right] \tag{1.62}$$

which ensures that the fluctuations vanish at the boundaries. We invite the reader to redo the following calculation with the expansion (1.62) replacing (1.58). While at the end the same propagator should be found, it will become clear why the expansion in terms of eigenfunctions satisfying (1.59) is preferable.

The integration over the fluctuations now becomes an integration over the expansion coefficients a_n. Inserting the expansion (1.58) into the action one finds

$$S^{(2)} = -\frac{m}{2} \sum_{n=1}^{\infty} \lambda_n a_n^2 = \frac{m}{2} \sum_{n=1}^{\infty} \left(\left(\frac{\pi n}{t}\right)^2 - \omega^2\right) a_n^2 \,. \tag{1.63}$$

As this result shows, the classical action is only an extremum of the action but not necessarily a minimum although this is the case for short time intervals $t < \pi/\omega$. The existence of conjugate points at times $T_n = n\pi/\omega$ mentioned above manifests itself here as vanishing of the eigenvalue λ_n. Then the action is independent of a_n which implies that for a time interval T_n all paths $x_{\mathrm{cl}} + a_n \xi_n$ with arbitrary coefficient a_n are solutions of the classical equation of motion.

After expansion of the fluctuations in terms of the eigenfunctions (1.60), the propagator takes the form

$$K(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0) \sim \exp\left(\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right) \int \left(\prod_{n=1}^{\infty} \mathrm{d}a_n\right) \exp\left(-\frac{\mathrm{i}}{\hbar} \frac{m}{2} \sum_{n=1}^{\infty} \lambda_n a_n^2\right) \,. \tag{1.64}$$

In principle, we need to know the Jacobi determinant of the transformation from the path integral to the integral over the Fourier coefficients. However, since this Jacobi determinant is independent of the oscillator frequency ω, we may also compare with the free particle. Evaluating the Gaussian fluctuation integrals, we find for the ratio between the prefactors of the propagators K_ω and K_0 of the harmonic oscillator and the free particle, respectively,

$$\frac{K_\omega \exp[-(\mathrm{i}/\hbar)S_{\mathrm{cl},\omega}]}{K_0 \exp[-(\mathrm{i}/\hbar)S_{\mathrm{cl},0}]} = \sqrt{\frac{D_0}{D}} \,. \tag{1.65}$$

Here, we have introduced the fluctuation determinants for the harmonic oscillator

$$D = \det\left(\frac{\mathrm{d}^2}{\mathrm{d}s^2} + \omega^2\right) = \prod_{n=1}^{\infty} \lambda_n \tag{1.66}$$

and the free particle

$$D_0 = \det\left(\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right) = \prod_{n=1}^{\infty} \lambda_n^0 \,. \tag{1.67}$$

The eigenvalues for the free particle

$$\lambda_n^0 = -\left(\frac{\pi n}{t}\right)^2 \tag{1.68}$$

are obtained from the eigenvalues (1.61) of the harmonic oscillator simply by setting the frequency ω equal to zero. With the prefactor of the propagator of the free particle

$$K_0 \exp\left(-\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl},0}\right) = \sqrt{\frac{m}{2\pi \mathrm{i}\hbar t}} \tag{1.69}$$

and (1.65), the propagator of the harmonic oscillator becomes

$$K(x_\mathrm{f}, t, x_\mathrm{i}, 0) = \sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}} \sqrt{\frac{D_0}{D}} \exp\left(\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right) \,. \tag{1.70}$$

For readers unfamiliar with the concept of determinants of differential operators we mention that we may define matrix elements of an operator by projection onto a basis as is familiar from standard quantum mechanics. The operator represented in its eigenbasis yields a diagonal matrix with the eigenvalues on the diagonal. Then, as for finite dimensional matrices, the determinant is the product of these eigenvalues.

Each of the determinants (1.66) and (1.67) by itself diverges. However, we are interested in the ratio between them which is well-defined [12]

$$\frac{D}{D_0} = \prod_{n=1}^{\infty} \left(1 - \left(\frac{\omega t}{\pi n}\right)^2\right) = \frac{\sin(\omega t)}{\omega t} \,. \tag{1.71}$$

Inserting this result into (1.70) leads to the propagator of the driven harmonic oscillator in its final form

$$\begin{aligned} K(x_\mathrm{f},t,x_\mathrm{i},0) &= \sqrt{\frac{m\omega}{2\pi\mathrm{i}\hbar\sin(\omega t)}}\exp\left[\frac{\mathrm{i}}{\hbar}S_\mathrm{cl}\right] \\ &= \sqrt{\frac{m\omega}{2\pi\hbar|\sin(\omega t)|}}\exp\left[\frac{\mathrm{i}}{\hbar}S_\mathrm{cl}-\mathrm{i}\left(\frac{\pi}{4}+n\frac{\pi}{2}\right)\right] \end{aligned} \tag{1.72}$$

with the classical action defined in (1.56). The Morse index n in the phase factor is given by the integer part of $\omega t/\pi$. This phase accounts for the changes in sign of the sine function [13]. Here, one might argue that it is not obvious which sign of the square root one has to take. However, the semigroup property (1.9) allows to construct propagators across conjugate points by joining propagators for shorter time intervals. In this way, the sign may be determined unambiguously [14].

It is interesting to note that the phase factor $\exp(-\mathrm{i}n\pi/2)$ in (1.72) implies that $K(x_\mathrm{f},2\pi/\omega,x_\mathrm{i},0) = -K(x_\mathrm{f},0,x_\mathrm{i},0) = -\delta(x_\mathrm{f}-x_\mathrm{i})$, i.e. the wave function after one period of oscillation differs from the original wave function by a factor -1. The oscillator thus returns to its original state only after two periods very much like a spin-1/2 particle which picks up a sign under rotation by 2π and returns to its original state only after a 4π-rotation. This effect might be observed in the case of the harmonic oscillator by letting interfere the wave functions of two oscillators with different frequency [15].

1.2.8 Semiclassical Approximation

The systems considered so far have been special in the sense that an exact expression for the propagator could be obtained. This is a consequence of the fact that the potential was at most quadratic in the coordinate. Unfortunately, in most cases of interest the potential is more complicated and apart from a few exceptions an exact evaluation of the path integral turns out to be impossible. To cope with such situations, approximation schemes have been devised. In the following, we will restrict ourselves to the most important approximation which is valid whenever the quantum fluctuations are small or, equivalently, when the actions involved are large compared to Planck's constant so that the latter may be considered to be small.

The decomposition of a general path into the classical path and fluctuations around it as employed in (1.48) in the previous section was merely a matter of convenience. For the exactly solvable case of a driven harmonic oscillator it is not really relevant how we express a general path satisfying the boundary conditions. Within the semiclassical approximation, however, it is decisive to expand around the path leading to the dominant contribution, i.e. the classical path. From a more mathematical point of view, we have to evaluate a path integral over $\exp(\mathrm{i}S/\hbar)$ for small $\hbar$. This can be done in a systematic way by the method of stationary phase where the exponent has to be expanded around the extrema of the action S.

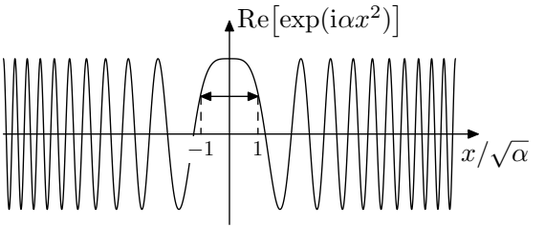


Fig. 1.8. In stationary phase approximation only a small region around the extremum contributes to the integral. For the example shown here, the extremum lies at $x = 0$

At this point it may be useful to give a brief reminder of the method of stationary phase. Suppose we want to evaluate the integral

$$I(\alpha) = \int_{-\infty}^{\infty} \mathrm{d}x g(x) \exp\bigl(\mathrm{i}\alpha f(x)\bigr) \tag{1.73}$$

in the limit of very large α. Inspection of Fig. 1.8, where $f(x) = x^2$, suggests that the dominant contribution to the integral comes from a region, in our example of size $1/\sqrt{\alpha}$, around the extremal (or stationary) point of the function $f(x)$. Outside of this region, the integrand is rapidly oscillating and therefore gives to leading order a negligible contribution. Since for large α, the region determining the integral is very small, we may expand the function $f(x)$ locally around the extremum x_0

$$f(x) \approx f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \dots \tag{1.74}$$

and replace $g(x)$ by $g(x_0)$. Neglecting higher order terms, which is allowed if $f''(x_0)$ is of order one, we are left with the Gaussian integral

$$\begin{aligned} I(\alpha) &\approx g(x_0) \exp\bigl(\mathrm{i}\alpha f(x_0)\bigr) \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(\frac{\mathrm{i}}{2} f''(x_0)(x - x_0)^2\right) \\ &= \sqrt{\frac{2\pi}{|f''(x_0)|}} g(x_0) \exp\Bigl[\mathrm{i}\alpha f(x_0) + \mathrm{i}\frac{\pi}{4} \mathrm{sgn}\bigl(f''(x_0)\bigr)\Bigr] \ , \end{aligned} \tag{1.75}$$

where $\mathrm{sgn}(f''(x_0))$ denotes the sign of $f''(x_0)$. If $f(x)$ possesses more than one extremum, one has to sum over the contributions of all extrema unless one extremum can be shown to be dominant.

We now apply the stationary phase approximation to path integrals where $1/\hbar$ plays the role of the large parameter. Since the action is stationary at classical paths, we are obliged to express the general path as

$$x(s) = x_{\mathrm{cl}}(s) + \xi(s) \ , \tag{1.76}$$

where x_{cl} is the classical path (or one of several possible paths) satisfying the boundary conditions and ξ represents the fluctuations around the classical path.

With this decomposition the action becomes

$$\begin{aligned} S &= \int_0^t \mathrm{d}s \left(\frac{m}{2}\dot{x}^2 - V(x)\right) \\ &= \int_0^t \mathrm{d}s \left(\frac{m}{2}\dot{x}_{\mathrm{cl}}^2 - V(x_{\mathrm{cl}})\right) + \int_0^t \mathrm{d}s \left(m\dot{x}_{\mathrm{cl}}\dot{\xi} - V'(x_{\mathrm{cl}})\xi\right) \\ &\quad + \int_0^t \mathrm{d}s \left(\frac{m}{2}\dot{\xi}^2 - \frac{1}{2}V''(x_{\mathrm{cl}})\xi^2\right) + \ldots \end{aligned} \tag{1.77}$$

It is instructive to compare this result with the action (1.50) for the driven harmonic oscillator. Again, the first term represents the classical action. The second term vanishes as was shown explicitly in (1.51) for the driven oscillator. In the general case, one can convince oneself by partial integration of the kinetic part and comparison with the classical equation of motion that this term vanishes again. This is of course a consequence of the fact that the classical path, around which we expand, corresponds to an extremum of the action. The third term on the right-hand-side of (1.77) is the leading order term in the fluctuations as was the case in (1.50). There is however an important difference since for anharmonic potentials the second derivative of the potential V'' is not constant and therefore the contribution of the fluctuations depends on the classical path. Finally, in general there will be higher order terms in the fluctuations as indicated by the dots in (1.77). The semiclassical approximation consists in neglecting these higher order terms so that after a partial integration, we get for the action

$$S_{\mathrm{sc}} = S_{\mathrm{cl}} - \frac{1}{2}\int_0^t \mathrm{d}s\, \xi\left(m\frac{\mathrm{d}^2}{\mathrm{d}s^2} + V''(x_{\mathrm{cl}})\right)\xi \tag{1.78}$$

where the index 'sc' indicates the semiclassical approximation.

Before deriving the propagator in semiclassical approximation, we have to discuss the regime of validity of this approximation. Since the first term in (1.78) gives only rise to a global phase factor, it is the second term which determines the magnitude of the quantum fluctuations. For this term to contribute, we should have $\xi^2/\hbar \lesssim 1$ so that the magnitude of typical fluctuations is at most of order $\sqrt{\hbar}$. The term of third order in the fluctuations is already smaller than the second order term by a factor $(\sqrt{\hbar})^3/\hbar = \sqrt{\hbar}$. If Planck's constant can be considered to be small, we may indeed neglect the fluctuation contributions of higher than second order except for one exception: It may happen that the second order term does not contribute, as has been the case at the conjugate points for the driven harmonic oscillator in Sect. 1.2.7. Then, the leading nonvanishing contribution becomes dominant. For the following discussion, we will not consider this latter case.

In analogy to Sect. 1.2.7 we obtain for the propagator in semiclassical approximation

$$K(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0) = \sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}}\sqrt{\frac{D_0}{D}}\exp\left(\frac{\mathrm{i}}{\hbar}S_{\mathrm{cl}}\right) \tag{1.79}$$

where

$$D = \det\left(\frac{\mathrm{d}^2}{\mathrm{d}s^2} + V''(x_{\mathrm{cl}})\right) \tag{1.80}$$

and D_0 is the fluctuation determinant (1.67) of the free particle.

Even though it may seem that determining the prefactor is a formidable task since the fluctuation determinant for a given potential has to be evaluated, this task can be greatly simplified. In addition, the following considerations offer the benefit of providing a physical interpretation of the prefactor. In our evaluation of the prefactor we follow Marinov [16]. The main idea is to make use of the semigroup property (1.9) of the propagator

$$\begin{aligned} &C(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0) \exp\left[\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0)\right] \\ &= \int \mathrm{d}x' C(x_{\mathrm{f}}, t, x', t') C(x', t', x_{\mathrm{i}}, 0) \exp\left[\frac{\mathrm{i}}{\hbar}\left[S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x', t') + S_{\mathrm{cl}}(x', t', x_{\mathrm{i}}, 0)\right]\right] \end{aligned} \tag{1.81}$$

where the prefactor C depends on the fluctuation contribution. We now have to evaluate the x'-integral within the semiclassical approximation. According to the stationary phase requirement discussed above, the dominant contribution to the integral comes from $x' = x_0(x_{\mathrm{f}}, x_{\mathrm{i}}, t, t')$ satisfying

$$\left.\frac{\partial S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x', t')}{\partial x'}\right|_{x'=x_0} + \left.\frac{\partial S_{\mathrm{cl}}(x', t', x_{\mathrm{i}}, 0)}{\partial x'}\right|_{x'=x_0} = 0\,. \tag{1.82}$$

According to classical mechanics these derivatives are related to initial and final momentum by [17]

$$\left(\frac{\partial S_{\mathrm{cl}}}{\partial x_{\mathrm{i}}}\right)_{x_{\mathrm{f}}, t_{\mathrm{f}}, t_{\mathrm{i}}} = -p_{\mathrm{i}} \qquad \left(\frac{\partial S_{\mathrm{cl}}}{\partial x_{\mathrm{f}}}\right)_{x_{\mathrm{i}}, t_{\mathrm{f}}, t_{\mathrm{i}}} = p_{\mathrm{f}} \tag{1.83}$$

so that (1.82) can expressed as

$$p(t' - \varepsilon) = p(t' + \varepsilon)\,. \tag{1.84}$$

The point x_0 thus has to be chosen such that the two partial classical paths can be joined with a continuous momentum. Together they therefore yield the complete classical path and in particular

$$S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0) = S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x_0, t') + S_{\mathrm{cl}}(x_0, t', x_{\mathrm{i}}, 0)\,. \tag{1.85}$$

This relation ensures that the phase factors depending on the classical actions on both sides of (1.81) are equal.

After having identified the stationary path, we have to evaluate the integral over x' in (1.81). Within semiclassical approximation this Gaussian integral leads to

$$\begin{aligned} &\frac{C(x_{\mathrm{f}}, t, x_{\mathrm{i}}, 0)}{C(x_{\mathrm{f}}, t, x_0, t') C(x_0, t', x_{\mathrm{i}}, 0)} \\ &\qquad = \left(\frac{1}{2\pi\mathrm{i}\hbar} \frac{\partial^2}{\partial x_0^2}\left[S_{\mathrm{cl}}(x_{\mathrm{f}}, t, x_0, t') + S_{\mathrm{cl}}(x_0, t', x_{\mathrm{i}}, 0)\right]\right)^{-1/2}. \end{aligned} \tag{1.86}$$

In order to make progress, it is useful to take the derivative of (1.85) with respect to x_f and x_i. Keeping in mind that x_0 depends on these two variables one finds

$$\begin{aligned}\frac{\partial^2 S_\mathrm{cl}(x_\mathrm{f},t,x_\mathrm{i},0)}{\partial x_\mathrm{f}\partial x_\mathrm{i}} &= \frac{\partial^2 S_\mathrm{cl}(x_\mathrm{f},t,x_0,t')}{\partial x_\mathrm{f}\partial x_0}\frac{\partial x_0}{\partial x_\mathrm{i}} + \frac{\partial^2 S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)}{\partial x_\mathrm{i}\partial x_0}\frac{\partial x_0}{\partial x_\mathrm{f}} \\ &\quad + \frac{\partial^2}{\partial x_0^2}\bigl[S_\mathrm{cl}(x_\mathrm{f},t,x_0,t') + S_\mathrm{cl}(x_0,t',x_\mathrm{i},t)\bigr]\frac{\partial x_0}{\partial x_\mathrm{i}}\frac{\partial x_0}{\partial x_\mathrm{f}}\,.\end{aligned} \tag{1.87}$$

Similarly, one finds by taking derivatives of the stationary phase condition (1.82)

$$\frac{\partial x_0}{\partial x_\mathrm{f}} = -\frac{\dfrac{\partial^2}{\partial x_\mathrm{f}x_0}S_\mathrm{cl}(x_\mathrm{f},t,x_0,t')}{\dfrac{\partial^2}{\partial x_0^2}\bigl[S_\mathrm{cl}(x_\mathrm{f},t,x_0,t') + S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)\bigr]} \tag{1.88}$$

and

$$\frac{\partial x_0}{\partial x_\mathrm{i}} = -\frac{\dfrac{\partial^2}{\partial x_\mathrm{i}x_0}S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)}{\dfrac{\partial^2}{\partial x_0^2}\bigl[S_\mathrm{cl}(x_\mathrm{f},t,x_0,t') + S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)\bigr]}\,. \tag{1.89}$$

These expressions allow to eliminate the partial derivatives of x_0 with respect to x_i and x_f appearing in (1.87) and one finally obtains

$$\begin{aligned}&\left(\frac{\partial^2}{\partial x_0^2}\bigl[S_\mathrm{cl}(x_\mathrm{f},t,x_0,t') + S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)\bigr]\right)^{-1} \\ &\qquad = -\frac{\dfrac{\partial^2}{\partial x_\mathrm{i}\partial x_\mathrm{f}}S_\mathrm{cl}(x_\mathrm{f},t,x_\mathrm{i},0)}{\dfrac{\partial^2 S_\mathrm{cl}(x_\mathrm{f},t,x_0,t')}{\partial x_\mathrm{f}\partial x_0}\dfrac{\partial^2 S_\mathrm{cl}(x_0,t',x_\mathrm{i},0)}{\partial x_\mathrm{i}\partial x_0}}\,.\end{aligned} \tag{1.90}$$

Inserting this result into (1.86), the prefactor can be identified as the so-called van Vleck–Pauli–Morette determinant [18,19,20]

$$C(x_\mathrm{f},t,x_\mathrm{i},0) = \left[\frac{1}{2\pi\mathrm{i}\hbar}\left(-\frac{\partial^2 S_\mathrm{cl}(x_\mathrm{f},t,x_\mathrm{i},0)}{\partial x_\mathrm{f}\partial x_\mathrm{i}}\right)\right]^{1/2} \tag{1.91}$$

so that the propagator in semiclassical approximation finally reads

$$\begin{aligned}&K(x_\mathrm{f},t,x_\mathrm{i},0) \\ &\quad = \left(\frac{1}{2\pi\hbar}\left|-\frac{\partial^2 S_\mathrm{cl}(x_\mathrm{f},t,x_\mathrm{i},0)}{\partial x_\mathrm{f}\partial x_\mathrm{i}}\right|\right)^{1/2}\exp\left[\frac{\mathrm{i}}{\hbar}S_\mathrm{cl}(x_\mathrm{f},t,x_\mathrm{i},0) - \mathrm{i}\left(\frac{\pi}{4}+n\frac{\pi}{2}\right)\right]\end{aligned} \tag{1.92}$$

where the Morse index n denotes the number of sign changes of $\partial^2 S_\mathrm{cl}/\partial x_\mathrm{f}\partial x_\mathrm{i}$ [13]. We had encountered such a phase factor before in the propagator (1.72) of the harmonic oscillator.

As we have already mentioned above, derivatives of the action with respect to position are related to momenta. This allows to give a physical interpretation of the prefactor of the propagator as the change of the end point of the path as a function of the initial momentum

$$\left(-\frac{\partial^2 S_{\rm cl}}{\partial x_{\rm i}\partial x_{\rm f}}\right)^{-1} = \frac{\partial x_{\rm f}}{\partial p_{\rm i}} \; . \tag{1.93}$$

A zero of this expression, or equivalently a divergence of the prefactor of the propagator, indicates a conjugate point where the end point does not depend on the initial momentum.

To close this section, we compare the semiclassical result (1.92) with exact results for the free particle and the harmonic oscillator. In our discussion of the free particle in Sect. 1.2.3 we already mentioned that the propagator can be expressed entirely in terms of classical quantities. Indeed, the expression (1.19) for the propagator of the free particle agrees with (1.92).

For the harmonic oscillator, we know from Sect. 1.2.7 that the prefactor does not depend on a possibly present external force. We may therefore consider the action (1.56) in the absence of driving $f(s) = 0$ which then reads

$$S_{\rm cl} = \frac{m\omega}{2\sin(\omega t)}\left[\left(x_{\rm i}^2 + x_{\rm f}^2\right)\cos(\omega t) - 2x_{\rm i}x_{\rm f}\right] \; . \tag{1.94}$$

Taking the derivative with respect to $x_{\rm i}$ and $x_{\rm f}$ one finds for the prefactor

$$C(x_{\rm f}, t, x_{\rm i}, 0) = \left(\frac{m\omega}{2\pi{\rm i}\hbar\sin(\omega t)}\right)^{1/2} \tag{1.95}$$

which is identical with the prefactor in our previous result (1.72). As expected, for potentials at most quadratic in the coordinate the semiclassical propagator agrees with the exact expression.

1.2.9 Imaginary Time Path Integral

In the discussion of dissipative systems we will be dealing with a system coupled to a large number of environmental degrees of freedom. In most cases, the environment will act like a large heat bath characterized by a temperature T. The state of the environment will therefore be given by an equilibrium density matrix. Occasionally, we may also be interested in the equilibrium density matrix of the system itself. Such a state may be reached after equilibration due to weak coupling with a heat bath.

In order to describe such thermal equilibrium states and the dynamics of the system on a unique footing, it is desirable to express equilibrium density matrices in terms of path integrals. This is indeed possible as one recognizes by writing the equilibrium density operator in position representation

$$\rho_\beta(x, x') = \frac{1}{\mathcal{Z}}\langle x|\exp(-\beta H)|x'\rangle \tag{1.96}$$

with the partition function

$$\mathcal{Z} = \int \mathrm{d}x \langle x| \exp(-\beta H)|x\rangle \ . \tag{1.97}$$

Comparing with the propagator in position representation

$$K(x,t,x',0) = \langle x| \exp\left(-\frac{\mathrm{i}}{\hbar}Ht\right)|x'\rangle \tag{1.98}$$

one concludes that apart from the partition function the equilibrium density matrix is equivalent to a propagator in imaginary time $t = -\mathrm{i}\hbar\beta$.

After the substitution $\sigma = \mathrm{i}s$ the action in imaginary time $-\mathrm{i}\hbar\beta$ reads

$$\int_0^{-\mathrm{i}\hbar\beta} \mathrm{d}s \left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2 - V(x)\right] = \mathrm{i}\int_0^{\hbar\beta} \mathrm{d}\sigma \left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\sigma}\right)^2 + V(x)\right] \ . \tag{1.99}$$

Here and in the following, we use greek letters to indicate imaginary times. Motivated by the right-hand side of (1.99) we define the so-called Euclidean action

$$S^{\mathrm{E}}[x] = \int_0^{\hbar\beta} \mathrm{d}\sigma \left[\frac{m}{2}\dot{x}^2 + V(x)\right] \ . \tag{1.100}$$

Even though one might fear a lack of intuition for motion in imaginary time, this results shows that it can simply be thought of as motion in the inverted potential in real time. With the Euclidean action (1.100) we now obtain as an important result the path integral expression for the (unnormalized) equilibrium density matrix

$$\langle x| \exp(-\beta H)|x'\rangle = \int_{\bar{x}(0)=x'}^{\bar{x}(\hbar\beta)=x} \mathcal{D}\bar{x} \exp\left(-\frac{1}{\hbar}S^{\mathrm{E}}[\bar{x}]\right) \ . \tag{1.101}$$

This kind of functional integral was discussed as early as 1923 by Wiener [21] in the context of classical Brownian motion.

As an example we consider the (undriven) harmonic oscillator. There is actually no need to evaluate a path integral since we know already from Sect. 1.2.7 the propagator

$$\begin{aligned} &K(x_{\mathrm{f}},t,x_{\mathrm{i}},0) \\ &\quad = \sqrt{\frac{m\omega}{2\pi\mathrm{i}\hbar\sin(\omega t)}} \exp\left[-\mathrm{i}\frac{m\omega}{2\hbar}\frac{(x_{\mathrm{i}}^2+x_{\mathrm{f}}^2)\cos(\omega t) - 2x_{\mathrm{i}}x_{\mathrm{f}}}{\sin(\omega t)}\right] \ . \end{aligned} \tag{1.102}$$

Transforming the propagator into imaginary time $t \to -\mathrm{i}\hbar\beta$ and renaming x_{i} and x_{f} into x' and x, respectively, one obtains the equilibrium density matrix

$$\begin{aligned} &\rho_\beta(x,x') \\ &\quad = \frac{1}{\mathcal{Z}}\sqrt{\frac{m\omega}{2\pi\hbar\sinh(\hbar\beta\omega)}} \exp\left[-\frac{m\omega}{2\hbar}\frac{(x^2+x'2)\cosh(\hbar\beta\omega) - 2xx'}{\sinh(\hbar\beta\omega)}\right] \ . \end{aligned} \tag{1.103}$$

The partition function is obtained by performing the trace as

$$\mathcal{Z} = \int \mathrm{d}x \langle x| \exp(-\beta H)|x\rangle = \frac{1}{2\sinh(\hbar\beta\omega/2)} \tag{1.104}$$

which agrees with the expression

$$\mathcal{Z} = \sum_{n=0}^{\infty} \exp\left[-\beta\hbar\omega\left(n+\frac{1}{2}\right)\right] \tag{1.105}$$

based on the energy levels of the harmonic oscillator.

Since the partition function often serves as a starting point for the calculation of thermodynamic properties, it is instructive to take a closer at how this quantity may be obtained within the path integral formalism. A possible approach is the one we just have sketched. By means of an imaginary time path integral one first calculates $\langle x| \exp(-\beta H)|x\rangle$ which is proportional to the probability to find the system at position x. Subsequent integration over coordinate space then yields the partition function.

However, the partition function may also be determined in one step. To this end, we expand around the periodic trajectory with extremal Euclidean action which in our case is given by $x(\sigma) = 0$. Any deviation will increase both the kinetic and potential energy and thus increase the Euclidean action. All other trajectories contributing to the partition function are generated by a Fourier series on the imaginary time interval from 0 to $\hbar\beta$

$$x(\sigma) = \frac{1}{\sqrt{\hbar\beta}}\left[a_0 + \sqrt{2}\sum_{n=1}^{\infty}\left(a_n\cos(\nu_n\sigma) + b_n\sin(\nu_n\sigma)\right)\right] \tag{1.106}$$

where we have introduced the so-called Matsubara frequencies

$$\nu_n = \frac{2\pi}{\hbar\beta}n\,. \tag{1.107}$$

This ansatz should be compared with (1.62) for the fluctuations where a_0 was fixed because the fluctuations had to vanish at the boundaries. For the partition function this requirement is dropped since we have to integrate over all periodic trajectories. Furthermore, we note that indeed with the ansatz (1.106) only the periodic trajectories contribute. All other paths cost an infinite amount of action due to the jump at the boundary as we will see shortly.

Inserting the Fourier expansion (1.106) into the Euclidean action of the harmonic oscillator

$$S^{\mathrm{E}} = \int_0^{\hbar\beta} \mathrm{d}\sigma \frac{m}{2}\left(\dot{x}^2 + \omega^2 x^2\right) \tag{1.108}$$

we find

$$S^{\mathrm{E}} = \frac{m}{2}\left[\omega^2 a_0^2 + \sum_{n=1}^{\infty}(\nu_n^2+\omega^2)(a_n^2+b_n^2)\right]\,. \tag{1.109}$$

As in Sect. 1.2.7 we do not want to go into the mathematical details of integration measures and Jacobi determinants. Unfortunately, the free particle cannot serve as a reference here because its partition function does not exist. We therefore content ourselves with remarking that because of

$$\frac{1}{\omega}\prod_{n=1}^{\infty}\frac{1}{\nu_n^2+\omega^2}=\frac{\hbar\beta}{\sum_{n=1}^{\infty}\nu_n^2}\frac{1}{2\sinh(\hbar\beta\omega/2)} \tag{1.110}$$

the result of the Gaussian integral over the Fourier coefficients yields the partition function up to a frequency independent factor. This enables us to determine the partition function in more complicated cases by proceeding as above and using the partition function of the harmonic oscillator as a reference.

Returning to the density matrix of the harmonic oscillator we finally obtain by inserting the partition function (1.104) into the expression (1.103) for the density matrix

$$\begin{aligned}\rho_\beta(x,x')=&\sqrt{\frac{m\omega}{\pi\hbar}\frac{\cosh(\hbar\beta\omega)-1}{\sinh(\hbar\beta\omega)}}\\ &\times\exp\left[-\frac{m\omega}{2\hbar}\frac{(x^2+x'^2)\cosh(\hbar\beta\omega)-2xx'}{\sinh(\hbar\beta\omega)}\right].\end{aligned} \tag{1.111}$$

Without path integrals, this result would require the evaluation of sums over Hermite polynomials.

The expression for the density matrix (1.111) can be verified in the limits of high and zero temperature. In the classical limit of very high temperatures, the probability distribution in real space is given by

$$P(x)=\rho_\beta(x,x)=\sqrt{\frac{\beta m\omega^2}{2\pi}}\exp\left(-\beta\frac{m\omega^2}{2}x^2\right)\sim\exp[-\beta V(x)]\,. \tag{1.112}$$

We thus have obtained the Boltzmann distribution which depends only on the potential energy. The fact that the kinetic energy does not play a role can easily be understood in terms of the path integral formalism. Excursions in a very short time $\hbar\beta$ cost too much action and are therefore strongly suppressed.

In the opposite limit of zero temperature the density matrix factorizes into a product of ground state wave functions of the harmonic oscillator

$$\lim_{\beta\to\infty}\rho_\beta(x,x')=\left[\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}x^2\right)\right]\left[\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}x'^2\right)\right] \tag{1.113}$$

as should be expected.

1.3 Dissipative Systems

1.3.1 Introduction

In classical mechanics dissipation can often be adequately described by including a velocity dependent damping term into the equation of motion. Such a phenomenological approach is no longer possible in quantum mechanics where the Hamilton formalism implies energy conservation for time-independent Hamiltonians. Then, a better understanding of the situation is necessary in order to arrive at an appropriate physical model.

A damped pendulum may help us to understand the mechanism of dissipation. The degree of freedom of interest, the elongation of the pendulum, undergoes a damped motion because it interacts with other degrees of freedom, the molecules in the air surrounding the pendulum's mass. We may consider the pendulum and the air molecules as one large system which, if assumed to be isolated from further degrees of freedom, obeys energy conservation. The energy of the pendulum alone, however, will in general not be conserved. This single degree of freedom is therefore subject to dissipation arising from the coupling to other degrees of freedom.

This insight will allow us in the following section to introduce a model for a system coupled to an environment and to demonstrate explicitly its dissipative nature. In particular, we will introduce the quantities needed for a description which focuses on the system degree of freedom. We are then in a position to return to the path integral formalism and to demonstrate how it may be employed to study dissipative systems. Starting from the model of system and environment, the latter will be eliminated to obtain a reduced description for the system alone. This leaves us with an effective action which forms the basis of the path integral description of dissipation.

1.3.2 Environment as a Collection of Harmonic Oscillators

A suitable model for dissipative quantum systems should both incorporate the idea of a coupling between system and environment and be amenable to an analytic treatment of the environmental coupling. These requirements are met by a model which nowadays is often referred to as Caldeira–Leggett model [22,23] even though it has been discussed in the literature under various names before for harmonic systems [24,25,26,27] and anharmonic systems [28]. The Hamiltonian

$$H = H_{\mathrm{S}} + H_{\mathrm{B}} + H_{\mathrm{SB}} \tag{1.114}$$

consists of three contributions. The Hamiltonian of the system degree of freedom

$$H_{\mathrm{S}} = \frac{p^2}{2m} + V(q) \tag{1.115}$$

models a particle of mass m moving in a potential V. Here, we denote the coordinate by q to facilitate the distinction from the environmental coordinates x_n which we will introduce in a moment. Of course, the system degree of freedom does not have to be associated with a real particle but may be quite abstract. In

fact, a substantial part of the calculations to be discussed in the following does not depend on the detailed form of the system Hamiltonian.

The Hamiltonian of the environmental degrees of freedom

$$H_{\mathrm{B}} = \sum_{n=1}^{N} \left(\frac{p_n^2}{2m_n} + \frac{m_n}{2} \omega_n^2 x_n^2 \right) \tag{1.116}$$

describes a collection of harmonic oscillators. While the properties of the environment may in some cases be chosen on the basis of a microscopic model, this does not have to be the case. Often, a phenomenological approach is sufficient as we will see below. As an example we mention an Ohmic resistor which as a linear electric element should be well described by a Hamiltonian of the form (1.116). On the other hand, the underlying mechanism leading to dissipation, e.g. in a resistor, may be much more complicated than that implied by the model of a collection of harmonic oscillators.

The coupling defined by the Hamiltonian

$$H_{\mathrm{SB}} = -q \sum_{n=1}^{N} c_n x_n + q^2 \sum_{n=1}^{N} \frac{c_n^2}{2m_n \omega_n^2} \tag{1.117}$$

is bilinear in the position operators of system and environment. There are cases where the bilinear coupling is realistic, e.g. for an environment consisting of a linear electric circuit like the resistor just mentioned or for a dipolar coupling to electromagnetic field modes encountered in quantum optics. Within a more general scope, this Hamiltonian may be viewed as linearization of a nonlinear coupling in the limit of weak coupling to the environmental degrees of freedom. As was first pointed out by Caldeira and Leggett, an infinite number of degrees of freedom still allows for strong damping even if each environmental oscillator couples only weakly to the system [22,23].

An environment consisting of harmonic oscillators as in (1.116) might be criticized. If the potential $V(q)$ is harmonic, one may pass to normal coordinates and thus demonstrate that after some time a revival of the initial state will occur. For sufficiently many environmental oscillators, however, this so-called Poincaré recurrence time tends to infinity [29]. Therefore, even with a linear environment irreversibility becomes possible at least for all practical purposes.

The reader may have noticed that in the coupling Hamiltonian (1.117) a term is present which only contains an operator acting in the system Hilbert space but depends on the coupling constants c_n. The physical reason for the inclusion of this term lies in a potential renormalization introduced by the first term in (1.117). This becomes clear if we consider the minimum of the Hamiltonian with respect to the system and environment coordinates. From the requirement

$$\frac{\partial H}{\partial x_n} = m_n \omega_n^2 x_n - c_n q \stackrel{!}{=} 0 \tag{1.118}$$

we obtain

$$x_n = \frac{c_n}{m_n \omega_n^2} q \; . \tag{1.119}$$

Using this result to determine the minimum of the Hamiltonian with respect to the system coordinate we find

$$\frac{\partial H}{\partial q} = \frac{\partial V}{\partial q} - \sum_{n=1}^{N} c_n x_n + q \sum_{n=1}^{N} \frac{c_n^2}{m_n \omega_n^2} = \frac{\partial V}{\partial q} \; . \tag{1.120}$$

The second term in (1.117) thus ensures that this minimum is determined by the bare potential $V(q)$.

After having specified the model, we now want to derive an effective description of the system alone. It was first shown by Magalinskiĭ [24] that the elimination of the environmental degrees of freedom leads indeed to a damped equation of motion for the system coordinate. We perform the elimination within the Heisenberg picture where the evolution of an operator A is determined by

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}[H, A] \; . \tag{1.121}$$

From the Hamiltonian (1.114) we obtain the equations of motion for the environmental degrees of freedom

$$\begin{aligned} \dot{p}_n &= -m_n \omega_n^2 x_n + c_n q \\ \dot{x}_n &= \frac{p_n}{m_n} \end{aligned} \tag{1.122}$$

and the system degree of freedom

$$\begin{aligned} \dot{p} &= -\frac{\partial V}{\partial q} + \sum_{n=1}^{N} c_n x_n - q \sum_{n=1}^{N} \frac{c_n^2}{m_n \omega_n^2} \\ \dot{x} &= \frac{p}{m} \; . \end{aligned} \tag{1.123}$$

The trick for solving the environmental equations of motion (1.122) consists in treating the system coordinate $q(t)$ as if it were a given function of time. The inhomogeneous differential equation then has the solution

$$x_n(t) = x_n(0)\cos(\omega_n t) + \frac{p_n(0)}{m_n \omega_n}\sin(\omega_n t) + \frac{c_n}{m_n \omega_n}\int_0^t \mathrm{d}s \sin\big(\omega_n(t-s)\big) q(s) \; . \tag{1.124}$$

Inserting this result into (1.123) one finds an effective equation of motion for the system coordinate

$$\begin{aligned} m\ddot{q} - \int_0^t \mathrm{d}s \sum_{n=1}^{N} \frac{c_n^2}{m_n \omega_n} \sin\big(\omega_n(t-s)\big) q(s) + \frac{\partial V}{\partial q} + q \sum_{n=1}^{N} \frac{c_n^2}{m_n \omega_n^2} \\ = \sum_{n=1}^{N} c_n \left[x_n(0)\cos(\omega_n t) + \frac{p_n(0)}{m_n \omega_n}\sin(\omega_n t) \right] . \end{aligned} \tag{1.125}$$

By a partial integration of the second term on the left-hand side this equation of motion can be cast into its final form

$$m\ddot{q} + m\int_0^t \mathrm{d}s\gamma(t-s)\dot{q}(s) + \frac{\partial V}{\partial q} = \xi(t) \tag{1.126}$$

with the damping kernel

$$\gamma(t) = \frac{1}{m}\sum_{n=1}^{N}\frac{c_n^2}{m_n\omega_n^2}\cos(\omega_n t) \tag{1.127}$$

and the operator-valued fluctuating force

$$\xi(t) = \sum_{n=1}^{N} c_n\left[\left(x_n(0) - \frac{c_n}{m_n\omega_n^2}q(0)\right)\cos(\omega_n t) + \frac{p_n(0)}{m_n\omega_n}\sin(\omega_n t)\right] . \tag{1.128}$$

The fluctuating force vanishes if averaged over a thermal density matrix of the environment including the coupling to the system

$$\langle\xi(t)\rangle_{\mathrm{B+SB}} = \frac{\mathrm{Tr}_{\mathrm{B}}\left[\xi(t)\exp\bigl(-\beta(H_{\mathrm{B}}+H_{\mathrm{SB}})\bigr)\right]}{\mathrm{Tr}_{\mathrm{B}}\left[\exp\bigl(-\beta(H_{\mathrm{B}}+H_{\mathrm{SB}})\bigr)\right]} = 0 . \tag{1.129}$$

For weak coupling, one may want to split off the transient term $m\gamma(t)q(0)$ which is of second order in the coupling and write the fluctuating force as [30]

$$\xi(t) = \zeta(t) - m\gamma(t)q(0) . \tag{1.130}$$

The so defined force $\zeta(t)$ vanishes if averaged over the environment alone

$$\langle\zeta(t)\rangle_{\mathrm{B}} = \frac{\mathrm{Tr}_{\mathrm{B}}\left[\zeta(t)\exp(-\beta H_{\mathrm{B}})\right]}{\mathrm{Tr}_{\mathrm{B}}\left[\exp(-\beta H_{\mathrm{B}})\right]} = 0 . \tag{1.131}$$

An important quantity to characterize the fluctuating force is the correlation function which again can be evaluated for ξ with respect to $H_{\mathrm{B}} + H_{\mathrm{SB}}$ or equivalently for ζ with respect to H_{B} alone. With (1.128) and (1.130) we get the correlation function

$$\langle\zeta(t)\zeta(0)\rangle_{\mathrm{B}} = \sum_{n,l} c_n c_l\left\langle\left(x_n(0)\cos(\omega_n t) + \frac{p_n(0)}{m_n\omega_i}\sin(\omega_n t)\right)x_l(0)\right\rangle_{\mathrm{B}} . \tag{1.132}$$

In thermal equilibrium the second moments are given by

$$\langle x_n(0)x_l(0)\rangle_{\mathrm{B}} = \delta_{nl}\frac{\hbar}{2m_n\omega_n}\coth\left(\frac{\hbar\beta\omega_n}{2}\right) \tag{1.133}$$

$$\langle p_n(0)x_l(0)\rangle_{\mathrm{B}} = -\frac{\mathrm{i}\hbar}{2}\delta_{nl} , \tag{1.134}$$

so that the noise correlation function finally becomes

$$\langle\zeta(t)\zeta(0)\rangle_{\mathrm{B}} = \sum_{n=1}^{N} \frac{\hbar c_n^2}{2m_n\omega_n}\left[\coth\left(\frac{\hbar\beta\omega_n}{2}\right)\cos(\omega_n t) - \mathrm{i}\sin(\omega_n t)\right] . \tag{1.135}$$

The imaginary part appearing here is a consequence of the fact that the operators $\zeta(t)$ and $\zeta(0)$ in general do not commute. The correlation function (1.135) appears as an integral kernel both in master equations as well as in the effective action derived below (cf. (1.168) and (1.169)).

It is remarkable that within a reduced description for the system alone all quantities characterizing the environment may be expressed in terms of the spectral density of bath oscillators

$$J(\omega) = \pi \sum_{n=1}^{N} \frac{c_n^2}{2m_n\omega_n}\delta(\omega - \omega_n) . \tag{1.136}$$

As an example, the damping kernel may be expressed in terms of this spectral density as

$$\gamma(t) = \frac{1}{m}\sum_{n=1}^{N} \frac{c_n^2}{m_n\omega_n^2}\cos(\omega_n t) = \frac{2}{m}\int_0^\infty \frac{\mathrm{d}\omega}{\pi}\frac{J(\omega)}{\omega}\cos(\omega t) . \tag{1.137}$$

For practical calculations, it is therefore unnecessary to specify all parameters m_n, ω_n and c_n appearing in (1.116) and (1.117). It rather suffices to define the spectral density $J(\omega)$.

The most frequently used spectral density

$$J(\omega) = m\gamma\omega \tag{1.138}$$

is associated with the so-called Ohmic damping. This term is sometimes employed to indicate a proportionality to frequency merely at low frequencies instead of over the whole frequency range. In fact, in any realistic situation the spectral density will not increase like in (1.138) for arbitrarily high frequencies. It is justified to use the term "Ohmic damping" even if (1.138) holds only below a certain frequency provided this frequency is much higher than the typical frequencies appearing in the system dynamics.

From (1.137) one finds the damping kernel for Ohmic damping

$$\gamma(t) = 2\gamma\delta(t) , \tag{1.139}$$

which renders (1.126) memory-free. We thus recover the velocity proportional damping term familiar from classical damped systems. It should be noted that the factor of two in (1.139) disappears upon integration in (1.126) since (1.137) implies that the delta function is symmetric around zero.

At this point, we want to briefly elucidate the origin of the term "Ohmic damping". Let us consider the electric circuit shown in Fig. 1.9 consisting of a

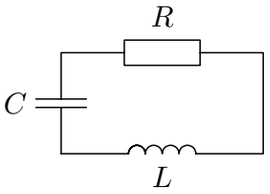


Fig. 1.9. LC oscillator with Ohmic damping due to a resistor R

resistance R, a capacitance C and an inductance L. Summing up the voltages around the loop, one obtains as equation of motion for the charge Q on the capacitor

$$L\ddot{Q} + R\dot{Q} + \frac{Q}{C} = 0 \,, \tag{1.140}$$

which shows that an Ohmic resistor leads indeed to memoryless damping. These considerations demonstrate that even without knowledge of the microscopic origin of dissipation in a resistor, we may employ the Ohmic spectral density (1.138) to account for its dissipative nature.

The spectral density (1.138) for Ohmic damping unfortunately diverges at high frequencies which, as already mentioned, cannot be the case in practice. Even in theoretical considerations this feature of strict Ohmic damping may lead to divergencies and a cutoff is needed for regularization. One possibility is the Drude cutoff, where the spectral density

$$J(\omega) = m\gamma\omega \frac{\omega_{\mathrm{D}}^2}{\omega^2 + \omega_{\mathrm{D}}^2} \tag{1.141}$$

above frequencies of the order of ω_{D} is suppressed. The corresponding damping kernel reads

$$\gamma(t) = \gamma\omega_{\mathrm{D}} \exp(-\omega_{\mathrm{D}}|t|) \,. \tag{1.142}$$

This leads to memory effects in (1.126) for short times $t < \omega_{\mathrm{D}}^{-1}$. For the long-time behaviour, however, only the Ohmic low frequency behaviour of the spectral density (1.141) is relevant. If a Drude cutoff is introduced for technical reasons, the cutoff frequency ω_{D} should be much larger than all other frequencies appearing in the problem in order to avoid spurious effects.

The relation (1.136) between the spectral density and the "microscopic" parameters implies that one may set $c_n = m_n\omega_n^2$ without loss of generality since the frequencies ω_n and the oscillator strengths $c_n^2/2m_n\omega_n$ can still be freely chosen. This special choice for the coupling constants has the advantage of a translationally invariant coupling [31]

$$H = H_S + \sum_{n=1}^{N} \left(\frac{p_n^2}{2m_n} + \frac{m_n}{2}\omega_n^2 (x_n - q)^2 \right) \,. \tag{1.143}$$

Furthermore, we now can determine the total mass of environmental oscillators

$$\sum_{n=1}^{N} m_n = \frac{2}{\pi}\int_0^\infty \mathrm{d}\omega \frac{J(\omega)}{\omega^3} \; . \tag{1.144}$$

If the spectral density of bath oscillators at small frequencies takes the form $J(\omega) \sim \omega^\alpha$, the total mass of bath oscillators is infinite for $\alpha \leq 2$. In particular, this includes the case of Ohmic damping where a free damped particle executes a diffusive motion. In contrast, for $\alpha > 2$, the total mass is finite. In this case, the particle will behave for long times like it were free albeit possessing a renormalized mass due to the environmental coupling [32]. We emphasize that the divergence of the total mass for $\alpha \leq 2$ is due to an infrared divergence and therefore independent of a high-frequency cutoff.

It is also useful to express the potential renormalization introduced in (1.117) in terms of the spectral density of bath oscillators. From (1.136) it is straightforward to obtain

$$q^2 \sum_{n=1}^{N} \frac{c_n^2}{2m_n\omega_n^2} = \frac{q^2}{\pi}\int_0^\infty \mathrm{d}\omega \frac{J(\omega)}{\omega} \; . \tag{1.145}$$

This term is infinite for strictly Ohmic damping but becomes finite when a high-frequency cutoff is introduced.

Finally, one finds for the noise correlation function (1.132)

$$\begin{aligned} K(t) &= \langle \zeta(t)\zeta(0)\rangle_{\mathrm{B}} \\ &= \hbar \int_0^\infty \frac{\mathrm{d}\omega}{\pi} J(\omega) \left[\coth\left(\frac{\hbar\beta\omega}{2}\right)\cos(\omega t) - \mathrm{i}\sin(\omega t)\right] \; . \end{aligned} \tag{1.146}$$

In the classical limit, $\hbar \to 0$, this correlation function reduces to the real-valued expression

$$K(t) = mk_{\mathrm{B}}T\gamma(t) \; , \tag{1.147}$$

where we have made use of (1.137). For Ohmic damping this implies delta correlated, i.e. white, noise.

In the quantum case, the noise correlation function is complex and can be decomposed into its real and imaginary part

$$K(t) = K'(t) + \mathrm{i}K''(t) \; . \tag{1.148}$$

Employing once more (1.137), one immediately finds that the imaginary part is related to the time derivative of the damping kernel by

$$K''(t) = \frac{m\hbar}{2}\frac{\mathrm{d}\gamma}{\mathrm{d}t} \; . \tag{1.149}$$

For Ohmic damping, the real part reads

$$K'(t) = -\frac{\pi m\gamma}{(\hbar\beta)^2}\frac{1}{\sinh^2\left(\dfrac{\pi t}{\hbar\beta}\right)} \tag{1.150}$$

which implies that at zero temperature the noise is correlated even for long times. The noise correlation then only decays algebraically like $1/t^2$ much in contrast to the classical result (1.147).

1.3.3 Effective Action

In the previous section we had eliminated the environmental degrees of freedom to obtain the effective equation of motion (1.126) for the system degree of freedom alone. This section will be devoted to a discussion of the corresponding procedure within the path integral formalism.

We start to illustrate the basic idea by considering the time evolution of the full density matrix of system and environment

$$W(q_\mathrm{f}, x_{n\mathrm{f}}, q'_\mathrm{f}, x'_{n\mathrm{f}}, t) = \int \mathrm{d}q_\mathrm{i}\mathrm{d}q'_\mathrm{i}\mathrm{d}x_{n\mathrm{i}}\mathrm{d}x'_{n\mathrm{i}} K(q_\mathrm{f}, x_{n\mathrm{f}}, t, q_\mathrm{i}, x_{n\mathrm{i}}, 0) \times W(q_\mathrm{i}, x_{n\mathrm{i}}, q'_\mathrm{i}, x'_{n\mathrm{i}}, 0) K^*(q'_\mathrm{f}, x'_{n\mathrm{f}}, t, q'_\mathrm{i}, x'_{n\mathrm{i}}, 0) \tag{1.151}$$

which is induced by the two propagators K. Here, the coordinates q and x_n refer again to the system and bath degrees of freedom, respectively. The environment is assumed to be in thermal equilibrium described by the density matrix W_β^B while the system may be in a nonequilibrium state ρ. If we neglect initial correlations between system and environment, i.e. if we switch on the coupling after preparation of the initial state, the initial density matrix may be written in factorized form

$$W(q_\mathrm{i}, x_{n\mathrm{i}}, q'_\mathrm{i}, x'_{n\mathrm{i}}, 0) = \rho(q_\mathrm{i}, q'_\mathrm{i}) W_\beta^\mathrm{B}(x_{n\mathrm{i}}, x'_{n\mathrm{i}}) \; . \tag{1.152}$$

Since we are only interested in the dynamics of the system degree of freedom, we trace out the environment. Then the time evolution may be expressed as

$$\rho(q_\mathrm{f}, q'_\mathrm{f}, t) = \int \mathrm{d}q_\mathrm{i}\mathrm{d}q'_\mathrm{i} J(q_\mathrm{f}, q'_\mathrm{f}, t, q_\mathrm{i}, q'_\mathrm{i}, 0) \rho(q_\mathrm{i}, q'_\mathrm{i}) \tag{1.153}$$

with the propagating function

$$J(q_\mathrm{f}, q'_\mathrm{f}, t, q_\mathrm{i}, q'_\mathrm{i}, 0) = \int \mathrm{d}x_{n\mathrm{f}}\mathrm{d}x_{n\mathrm{i}}\mathrm{d}x'_{n\mathrm{i}} K(q_\mathrm{f}, x_{n\mathrm{f}}, t, q_\mathrm{i}, x_{n\mathrm{i}}, 0) \times W_\beta^\mathrm{B}(x_{n\mathrm{i}}, x'_{n\mathrm{i}}) K^*(q'_\mathrm{f}, x_{n\mathrm{f}}, t, q'_\mathrm{i}, x'_{n\mathrm{i}}, 0) \; . \tag{1.154}$$

Here, the trace has been performed by setting $x_{n\mathrm{f}} = x'_{n\mathrm{f}}$ and integrating over these coordinates. The propagators may be expressed as real time path integrals while the equilibrium density matrix of the bath is given by a path integral in imaginary time. Performing the path integrals and the conventional integrals appearing in (1.154) one finds a functional depending on the system path. The important point is that this functional contains all information about the environment required to determine the system dynamics.

For factorizing initial conditions, the propagating function J has been calculated by Feynman and Vernon [33] on the basis of the Hamiltonian (1.114).

More general initial conditions taking into account correlations between system and environment may be considered as well [34].

Instead of deriving the propagating function we will demonstrate how to trace the environment out of the equilibrium density matrix of system plus environment. While this task is conceptually similar and leads to the same physical insight, it is considerably less tedious.

We start from the imaginary time path integral representation of the full equilibrium density matrix

$$W_\beta(q, x_n, q', x_n') = \frac{1}{\mathcal{Z}_\beta} \int \mathcal{D}\bar{q} \left(\prod_{n=1}^{N} \mathcal{D}\bar{x}_n \right) \exp\left(-\frac{1}{\hbar} S^{\mathrm{E}}[\bar{q}, \bar{x}_n] \right) \tag{1.155}$$

where the paths run from $\bar{q}(0) = q'$ and $\bar{x}_n(0) = x_n'$ to $\bar{q}(\hbar\beta) = q$ and $\bar{x}_n(\hbar\beta) = x_n$. The Euclidean action corresponding to the model Hamiltonian (1.114) reads in imaginary time

$$S^{\mathrm{E}}[\bar{q}, \bar{x}_n] = S_{\mathrm{S}}^{\mathrm{E}}[\bar{q}] + S_{\mathrm{B}}^{\mathrm{E}}[\bar{x}_n] + S_{\mathrm{SB}}^{\mathrm{E}}[\bar{q}, \bar{x}_n] \tag{1.156}$$

with

$$S_{\mathrm{S}}^{\mathrm{E}}[\bar{q}] = \int_0^{\hbar\beta} \mathrm{d}\tau \left(\frac{m}{2} \dot{\bar{q}}^2 + V(\bar{q}) \right) \tag{1.157}$$

$$S_{\mathrm{B}}^{\mathrm{E}}[\bar{x}_n] = \int_0^{\hbar\beta} \mathrm{d}\tau \sum_{n=1}^{N} \frac{m_n}{2} \left(\dot{\bar{x}}_n^2 + \omega_n^2 \bar{x}_n^2 \right) \tag{1.158}$$

$$S_{\mathrm{SB}}^{\mathrm{E}}[\bar{q}, \bar{x}_n] = \int_0^{\hbar\beta} \mathrm{d}\tau \left(-\bar{q} \sum_{n=1}^{N} c_n \bar{x}_n + \bar{q}^2 \sum_{n=1}^{N} \frac{c_n^2}{2 m_n \omega_n^2} \right) . \tag{1.159}$$

The reduced density matrix of the system is obtained by tracing over the environmental degrees of freedom

$$\begin{aligned} \rho_\beta(q, q') &= \mathrm{tr}_{\mathrm{B}} \left(W_\beta(q, x_n, q', x_n') \right) \\ &= \frac{1}{\mathcal{Z}_\beta} \int \mathcal{D}\bar{q} \int \prod_{n=1}^{N} \mathrm{d}x_n \oint \prod_{n=1}^{N} \mathcal{D}\bar{x}_n \exp\left(-\frac{1}{\hbar} S^{\mathrm{E}}[\bar{q}, \bar{x}_n] \right) \end{aligned} \tag{1.160}$$

where the circle on the second functional integral sign indicates that one has to integrate over closed paths $\bar{x}_n(0) = \bar{x}_n(\hbar\beta) = x_n$ when performing the trace. The dependence on the environmental coupling may be made explicit by writing

$$\rho_\beta(q, q') = \frac{1}{\mathcal{Z}} \int \mathcal{D}\bar{q} \exp\left(-\frac{1}{\hbar} S_{\mathrm{S}}^{\mathrm{E}}[\bar{q}] \right) \mathcal{F}[\bar{q}] \tag{1.161}$$

where the influence functional $\mathcal{F}[\bar{q}]$ describes the influence of the environment on the system. Here, the partition function $\mathcal{Z}$ should not be confused with the partition function $\mathcal{Z}_\beta$ of system plus environment. The relation between the two quantities will be discussed shortly.

Since the bath oscillators are not coupled among each other, the influence functional may be decomposed into factors corresponding to the individual bath oscillators

$$\mathcal{F}[\bar{q}] = \prod_{n=1}^{N} \frac{1}{\mathcal{Z}_n} \mathcal{F}_n[\bar{q}] \tag{1.162}$$

where

$$\mathcal{Z}_n = \frac{1}{2\sinh\left(\hbar\beta\omega_n/2\right)} \tag{1.163}$$

is the partition function of a single bath oscillator. The influence functional of a bath oscillator can be expressed as

$$\mathcal{F}_n[\bar{q}] = \int \mathrm{d}x_n \oint \mathcal{D}\bar{x}_n \exp\left(-\frac{1}{\hbar} S_n^{\mathrm{E}}[\bar{q}, \bar{x}_n]\right) \tag{1.164}$$

with the action

$$S_n^{\mathrm{E}}[\bar{q}, \bar{x}_n] = \int_0^{\hbar\beta} \mathrm{d}\tau \frac{m_n}{2} \left[\dot{\bar{x}}_n^2 + \omega_n^2 \left(\bar{x}_n - \frac{c_n}{m_n\omega_n^2}\bar{q}\right)^2\right] . \tag{1.165}$$

The partition function $\mathcal{Z}$ of the damped system is related to the full partition function $\mathcal{Z}_\beta$ by the partition function of the environmental oscillators $\mathcal{Z}_{\mathrm{B}} = \prod_{n=1}^{N} \mathcal{Z}_n$ according to $\mathcal{Z} = \mathcal{Z}_\beta/\mathcal{Z}_{\mathrm{B}}$. In the limit of vanishing coupling, $c_n = 0$, the influence functional becomes $\mathcal{F}[\bar{q}] = 1$ so that (1.161) reduces to the path integral representation of the density matrix of an isolated system as it should.

Apart from the potential renormalization term proportional to $\bar{q}^2$, the action (1.165) describes a driven harmonic oscillator. We may therefore make use of our results from Sect. 1.2.7. After analytic continuation $t \to -\mathrm{i}\hbar\beta$ in (1.56) and setting $x_{\mathrm{i}} = x_{\mathrm{f}} = x_n$ one finds for the classical Euclidean action

$$\begin{aligned} S_n^{\mathrm{E,cl}}[\bar{q}] = {} & m_n\omega_n \frac{\cosh(\hbar\beta\omega_n) - 1}{\sinh(\hbar\beta\omega_n)} x_n^2 \\ & - c_n \int_0^{\hbar\beta} \mathrm{d}\tau \frac{\sinh(\omega_n\tau) + \sinh(\omega_n(\hbar\beta - \tau))}{\sinh(\hbar\beta\omega_n)} x_n \bar{q}(\tau) \\ & - \frac{c_n^2}{m_n\omega_n} \int_0^{\hbar\beta} \mathrm{d}\tau \int_0^{\tau} \mathrm{d}\sigma \frac{\sinh(\omega_n(\hbar\beta - \tau))\sinh(\omega_n\sigma)}{\sinh(\hbar\beta\omega_n)} \bar{q}(\tau)\bar{q}(\sigma) \\ & + \frac{c_n^2}{2m_n\omega_n^2} \int_0^{\hbar\beta} \mathrm{d}\tau \bar{q}^2(\tau) . \end{aligned} \tag{1.166}$$

In view of the required integration over x_n one completes the square

$$\begin{aligned} S_n^{\mathrm{E,cl}}[\bar{q}] = {} & m_n\omega_n \frac{\cosh(\hbar\beta\omega_n) - 1}{\sinh(\hbar\beta\omega_n)} (x_n - x_n^{(0)})^2 - \int_0^{\hbar\beta} \mathrm{d}\tau \int_0^{\tau} \mathrm{d}\sigma K_n(\tau - \sigma)\bar{q}(\tau)\bar{q}(\sigma) \\ & + \frac{c_n^2}{2m_n\omega_n^2} \int_0^{\hbar\beta} \mathrm{d}\tau \bar{q}^2(\tau) \end{aligned} \tag{1.167}$$

where $x_n^{(0)}$ does not need to be specified since it drops out after integration.

The integral kernel appearing in (1.167) follows from (1.166) as

$$K_n(\tau) = \frac{c_n^2}{2m_n\omega_n} \frac{\cosh\left(\omega_n\big(\frac{\hbar\beta}{2} - \tau\big)\right)}{\sinh\left(\frac{\hbar\beta\omega_n}{2}\right)} = K_n(\hbar\beta - \tau) \tag{1.168}$$

and therefore can be identified as the noise correlation function (1.135) in imaginary time

$$K_n(\tau) = \frac{1}{\hbar}\langle\zeta_n(-\mathrm{i}\tau)\zeta_n(0)\rangle_{\mathrm{B}} \; . \tag{1.169}$$

The term in (1.167) containing this kernel is quite unusual for an action. The double integral describes a nonlocal contribution where the system trajectory interacts with itself. This self-interaction is mediated by the environment as can be seen from the factor c_n^2 in (1.168).

The integral kernel $K_n(\tau)$ is only needed in an interval of length $\hbar\beta$. Periodic continuation outside of this interval therefore allows us to expand the kernel into a Fourier series

$$\begin{aligned} K_n(\tau) &= \frac{c_n^2}{\hbar\beta m_n\omega_n} \sum_{l=-\infty}^{\infty} \frac{\omega_n}{\omega_n^2 + \nu_l^2} \exp(\mathrm{i}\nu_l\tau) \\ &= \frac{c_n^2}{\hbar\beta m_n\omega_n^2} \sum_{l=-\infty}^{\infty} \exp(\mathrm{i}\nu_l\tau) - \frac{c_n^2}{\hbar\beta m_n\omega_n^2} \sum_{l=-\infty}^{\infty} \frac{\nu_l^2}{\omega_n^2 + \nu_l^2} \exp(\mathrm{i}\nu_l\tau) \\ &= \frac{c_n^2}{m_n\omega_n^2} \sum_{j=-\infty}^{\infty} \delta(\tau - j\hbar\beta) - k_n(\tau) \; , \end{aligned} \tag{1.170}$$

where the Matsubara frequencies ν_l have been defined in (1.107). In (1.170), we split the kernel into two parts. The first term contains delta functions which lead to a local contribution to the action. Noting that due to the region of integration in (1.167) only half of the delta function contributes, this local term just cancels the potential renormalization (1.145). We are therefore left with the nonlocal kernel

$$k_n(\tau) = \frac{c_n^2}{\hbar\beta m_n\omega_n^2} \sum_{l=-\infty}^{\infty} \frac{\nu_l^2}{\omega_n^2 + \nu_l^2} \exp(\mathrm{i}\nu_l\tau) \; . \tag{1.171}$$

It can be shown that this kernel no longer contains a local contribution by writing

$$\begin{aligned} &\int_0^{\hbar\beta} \mathrm{d}\tau \int_0^{\tau} \mathrm{d}\sigma k_n(\tau - \sigma)\bar{q}(\tau)\bar{q}(\sigma) \\ &\qquad = -\frac{1}{2}\int_0^{\hbar\beta} \mathrm{d}\tau \int_0^{\tau} \mathrm{d}\sigma k_n(\tau - \sigma)\left[\big(\bar{q}(\tau) - \bar{q}(\sigma)\big)^2 - \big(\bar{q}(\tau)^2 + \bar{q}(\sigma)^2\big)\right] \; . \end{aligned} \tag{1.172}$$

The first term is manifestly nonlocal because it contains the difference $\bar{q}(\tau)-\bar{q}(\sigma)$. Exploiting the symmetry of $k_n(\tau)$, the second term can be expressed as

$$\frac{1}{2}\int_0^{\hbar\beta}\mathrm{d}\tau\int_0^{\tau}\mathrm{d}\sigma k_n(\tau-\sigma)\left(\bar{q}(\tau)^2+\bar{q}(\sigma)^2\right)=\int_0^{\hbar\beta}\mathrm{d}\tau\bar{q}(\tau)^2\int_0^{\hbar\beta}\mathrm{d}\sigma k_n(\sigma)\ . \quad (1.173)$$

Therefore, this term potentially could result in a local contribution. However, the time integral over the interval from 0 to $\hbar\beta$ corresponds to the $l=0$ Fourier component which vanishes for $k_n(\tau)$. As a consequence, the kernel $k_n(\tau)$ indeed gives rise to a purely nonlocal contribution to the action.

Now, we can carry out the Gaussian integral over x_n appearing in the influence functional (1.164). With the action (1.167) we find

$$\mathcal{F}_n[\bar{q}]=\mathcal{Z}_n\exp\left(-\frac{1}{2\hbar}\int_0^{\hbar\beta}\mathrm{d}\tau\int_0^{\hbar\beta}\mathrm{d}\sigma k_n(\tau-\sigma)\bar{q}(\tau)\bar{q}(\sigma)\right)\ . \quad (1.174)$$

The partition function $\mathcal{Z}_n$ arises from the fluctuation contribution and may be shown to be given by (1.163) for example by comparison with the uncoupled case $c_n=0$.

With (1.162) we finally obtain the influence functional

$$\mathcal{F}[\bar{q}]=\exp\left(-\frac{1}{2\hbar}\int_0^{\hbar\beta}\mathrm{d}\tau\int_0^{\hbar\beta}\mathrm{d}\sigma k(\tau-\sigma)\bar{q}(\tau)\bar{q}(\sigma)\right) \quad (1.175)$$

with

$$\begin{aligned}k(\tau)=\sum_{n=1}^{N}k_n(\tau)&=\sum_{n=1}^{N}\frac{c_n^2}{\hbar\beta m_n\omega_n^2}\sum_{l=-\infty}^{\infty}\frac{\nu_l^2}{\omega_n^2+\nu_l^2}\exp(\mathrm{i}\nu_l\tau)\\&=\frac{2}{\hbar\beta}\int_0^{\infty}\frac{\mathrm{d}\omega}{\pi}\frac{J(\omega)}{\omega}\sum_{l=-\infty}^{\infty}\frac{\nu_l^2}{\omega^2+\nu_l^2}\exp(\mathrm{i}\nu_l\tau)\end{aligned} \quad (1.176)$$

where we have made use of the spectral density of bath oscillators (1.136) to obtain the last line.

The kernel $k(\tau)$ may be related to the damping kernel $\gamma(t)$ by observing that the Laplace transform of the latter is given by

$$\begin{aligned}\hat{\gamma}(z)=\int_0^{\infty}\mathrm{d}t\exp(-zt)\gamma(t)&=\frac{2}{m}\int_0^{\infty}\frac{\mathrm{d}\omega}{\pi}\frac{J(\omega)}{\omega}\int_0^{\infty}\mathrm{d}t\exp(-zt)\cos(\omega t)\\&=\frac{2}{m}\int_0^{\infty}\frac{\mathrm{d}\omega}{\pi}\frac{J(\omega)}{\omega}\frac{z}{z^2+\omega^2}\ .\end{aligned} \quad (1.177)$$

In the first line we have employed the relation (1.137) between the damping kernel and the spectral density of bath oscillators. In view of (1.176) and (1.177) we can finally express the kernel as

$$k(\tau)=\frac{m}{\hbar\beta}\sum_{l=-\infty}^{\infty}|\nu_l|\hat{\gamma}(|\nu_l|)\exp(\mathrm{i}\nu_l\tau)\ . \quad (1.178)$$

For strictly Ohmic damping, this kernel is highly singular and we therefore introduce a Drude cutoff. The Laplace transform of the corresponding damping kernel is obtained from (1.142) as

$$\hat{\gamma}(z) = \frac{\gamma\omega_{\mathrm{D}}}{\omega_{\mathrm{D}} + z} \tag{1.179}$$

which reduces to $\hat{\gamma}(z) = \gamma$ for strictly Ohmic damping. Keeping the leading terms in the cutoff frequency ω_{D}, the kernel now reads

$$k(\tau) = m\gamma\omega_{\mathrm{D}} \sum_{n=-\infty}^{\infty} \delta(\tau - n\hbar\beta) - \frac{\pi m\gamma}{(\hbar\beta)^2} \frac{1}{\sin^2\left(\dfrac{\pi\tau}{\hbar\beta}\right)} + O(\omega_{\mathrm{D}}^{-1}) \; . \tag{1.180}$$

For low temperatures, this gives rise to a long range interaction between different parts of the system trajectory. This reminds us of the algebraic decay of the real part (1.150) of the noise correlation function $K(t)$ and in fact it follows from our previous discussion that up to the periodic delta function appearing in (1.180) the kernel $k(\tau)$ equals $-K(-\mathrm{i}\tau)$.

Summarizing this calculation we obtain the important result that the influence of the environment on the system may be taken into account by adding a nonlocal contribution to the action. We then obtain the effective action

$$S_{\mathrm{eff}}^{\mathrm{E}}[\bar{q}] = S_{\mathrm{S}}^{\mathrm{E}}[\bar{q}] + \frac{1}{2}\int_0^{\hbar\beta} \mathrm{d}\tau \int_0^{\hbar\beta} \mathrm{d}\sigma \, k(\tau - \sigma)\bar{q}(\tau)\bar{q}(\sigma) \tag{1.181}$$

with $k(\tau)$ given by (1.178). The elimination of the environment within the real time path integral formalism, e.g. along the lines of the calculation by Feynman and Vernon [33] mentioned at the beginning of this section, would have led to an effective action of a structure similar to (1.181). An important difference consists in the fact that the propagation of a density matrix involves two paths instead of one. In addition, the integral kernel then of course appears in its real time version.

1.4 Damped Harmonic Oscillator

1.4.1 Partition Function

In this final section we will apply the results of the previous sections to the damped harmonic oscillator where exact results may be obtained analytically. The Hamiltonian describing system and environment is given by (1.114)–(1.117) with the potential

$$V(q) = \frac{m}{2}\omega_0^2 q^2 \; . \tag{1.182}$$

As we have seen in Sect. 1.3.2, there is no point in dealing with all the microscopic parameters present in the Hamiltonians (1.116) and (1.117). Instead, it

is sufficient to specify the spectral density of bath oscillators (1.136). In the following, we will mostly assume Ohmic damping, i.e. $J(\omega) = m\gamma\omega$, and introduce a high-frequency cutoff of the Drude type (1.141) when necessary.

In the general discussion of dissipative quantum systems we have concentrated on imaginary time calculations and we will therefore try to take this approach as a starting point of the following considerations. Probably the most important quantity which can be obtained in imaginary time is the partition function which in statistical mechanics can be viewed as a generating function for expectation values. Based on our previous discussion of the partition function of the undamped harmonic oscillator in Sect. 1.2.9 and the effective action in imaginary time in Sect. 1.3.3, it is rather straightforward to obtain the partition function of the damped harmonic oscillator.

According to (1.104) and (1.110) we can express the partition function of the undamped harmonic oscillator as

$$\mathcal{Z}_{\mathrm{u}} = \frac{1}{\hbar\beta\omega_0} \prod_{n=1}^{\infty} \frac{\nu_n^2}{\nu_n^2 + \omega_0^2} \,. \tag{1.183}$$

We remind the reader that the denominator of the product stems from the fluctuation determinant associated with the Euclidean action

$$S^{\mathrm{E}}[q] = \int_0^{\hbar\beta} \mathrm{d}\tau \left(\frac{m}{2}\dot{q}^2 + \frac{m}{2}\omega_0^2 q^2 \right) \,. \tag{1.184}$$

As we have seen in Sect. 1.3.3, the coupling to the environment leads to an additional, nonlocal term to the action. Expanding the fluctuations in a Fourier series as we did on p. 25 and making use of the Fourier decomposition (1.178) of the integral kernel $k(\tau)$, we conclude that an additional term $\nu_n\hat{\gamma}(\nu_n)$ appears in the fluctuation determinant. Modifying (1.183) accordingly, we find for the partition function of the damped harmonic oscillator

$$\mathcal{Z} = \frac{1}{\hbar\beta\omega_0} \prod_{n=1}^{\infty} \frac{\nu_n^2}{\nu_n^2 + \nu_n\hat{\gamma}(\nu_n) + \omega_0^2} \,. \tag{1.185}$$

For strictly Ohmic damping, we have $\hat{\gamma}(\nu_n) = \gamma$. Since infinite products over terms of the form $1 + a/n$ for large n do not converge, we are forced to introduce a high-frequency cutoff in order to obtain a finite result. One possibility is the Drude cutoff (1.142) with $\hat{\gamma}$ given by (1.179).

In the following section we will try to extract some interesting information from the partition function and in the process will get an idea of where the difficulties for strictly Ohmic damping arise from.

1.4.2 Ground State Energy and Density of States

A thermodynamic quantity directly related to the partition function is the free energy which can be obtained from the former by means of

$$F = -\frac{1}{\beta} \ln(\mathcal{Z}) \,. \tag{1.186}$$

In the limit of zero temperature, the free energy becomes the ground state energy of the undamped oscillator shifted due to the coupling to the environment. For the free energy, we find with (1.185)

$$F = \frac{1}{\beta}\ln(\hbar\beta\omega_0) + \frac{1}{\beta}\sum_{n=1}^{\infty}\ln\left(1 + \frac{\hat{\gamma}(\nu_n)}{\nu_n} + \frac{\omega_0^2}{\nu_n^2}\right) . \tag{1.187}$$

In the limit $\beta \to \infty$ the spacing between the Matsubara frequencies ν_n goes to zero and the sum turns into the ground state energy of the damped oscillator given by the integral

$$\varepsilon_0 = \frac{\hbar}{2\pi}\int_0^{\infty} \mathrm{d}\nu \ln\left(1 + \frac{\hat{\gamma}(\nu)}{\nu} + \frac{\omega_0^2}{\nu^2}\right) . \tag{1.188}$$

It is particularly interesting to consider the case of weak coupling where connection can be made to results of perturbation theory. This will also help to understand the physical meaning of a ground state energy of a dissipative system derived from a free energy. An expansion of (1.188) including terms of order γ yields

$$\varepsilon_0 = \frac{\hbar}{2\pi}\int_0^{\infty} \mathrm{d}\nu \ln\left(1 + \frac{\omega_0^2}{\nu^2}\right) + \frac{\hbar}{2\pi}\int_0^{\infty} \mathrm{d}\nu \frac{\nu}{\nu^2 + \omega_0^2}\hat{\gamma}(\nu) . \tag{1.189}$$

Evaluation of the first integral yields the expected result $\hbar\omega_0/2$, i.e. the ground state energy of the undamped harmonic oscillator. The second integral represents the shift due to the coupling to the environmental oscillators and can be expressed in terms of the spectral density of bath oscillators $J(\omega)$. Recalling (1.177) one can perform the integral over ν and the ground state energy (1.189) in the presence of damping becomes

$$\varepsilon_0 = \frac{\hbar\omega_0}{2} + \frac{\hbar}{2\pi m}\int_0^{\infty} \mathrm{d}\omega J(\omega)\frac{1}{\omega(\omega_0 + \omega)} . \tag{1.190}$$

In order to facilitate the physical interpretation, we rewrite this result as

$$\varepsilon_0 = \frac{\hbar\omega_0}{2} - \frac{\hbar}{2\pi m\omega_0}\int_0^{\infty} \mathrm{d}\omega J(\omega)\left(\frac{1}{\omega_0 + \omega} - \frac{1}{\omega}\right) . \tag{1.191}$$

The first term of order γ may be interpreted in analogy to the Lamb shift. There, an atomic level is shifted by creation and subsequent annihilation of a virtual photon as a consequence of the coupling to the electromagnetic vacuum. In our case, the atomic level is replaced by the ground state of the harmonic oscillator and the environmental oscillators are completely equivalent to the modes of the electromagnetic field. The pictorial representation of this process is shown in Fig. 1.10. The coupling Hamiltonian (1.117) allows a transition from the ground state into the first excited state $|1\rangle$ by excitation of the j-th environmental oscillator into its first excited state $|1_j\rangle$.

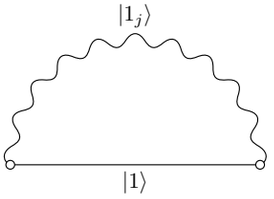


Fig. 1.10. The ground state energy of the harmonic oscillator is shifted by a transition to the first excited state accompanied by a virtual excitation of the j-th environmental mode

The energy shift associated with the diagram depicted in Fig. 1.10 is given by second order perturbation theory as

$$\Delta_0 = \sum_{j=1}^{N} \frac{|\langle 1, 1_j|c_j q x_j|0,0\rangle|^2}{-\hbar\omega_0 - \hbar\omega_j} \tag{1.192}$$

where the denominator is determined by the energy of the intermediate state $|1, 1_j\rangle$. With the matrix element

$$\langle 1, 1_j|qx_j|0,0\rangle = \frac{\hbar}{2(mm_j\omega_0\omega_j)^{1/2}} \tag{1.193}$$

and the relation (1.136) for the spectral density of bath oscillators we get

$$\Delta_0 = -\frac{\hbar}{2\pi m\omega_0}\int_0^\infty \mathrm{d}\omega J(\omega)\frac{1}{\omega_0+\omega} \tag{1.194}$$

which is just the first term of order γ in (1.191). As we know, the bilinear coupling Hamiltonian appearing in (1.192) gives rise to a renormalization of the potential which has been taken care of by the second term in the Hamiltonian (1.117). The result (1.194) contains this potential renormalization since only the bilinear coupling term has been considered. In (1.191), which results from the full Hamiltonian, this effect is subtracted off by the second term under the integral in (1.191) as can be verified by comparison with (1.145).

It is obvious that for strictly Ohmic damping with $J(\omega) = m\gamma\omega$ the correction (1.194) and with it the ground state energy (1.191) will display an ultraviolet divergence which is due to the unphysical behaviour of the spectral density $J(\omega)$ at large frequencies. Assuming a Drude cutoff we find with (1.141) to leading order in the cutoff frequency ω_D the finite result

$$\Delta_0 = -\frac{\hbar\gamma\omega_\mathrm{D}}{4\omega_0} + \frac{\hbar\gamma}{2\pi}\ln\left(\frac{\omega_\mathrm{D}}{\omega_0}\right) + O(\omega_\mathrm{D}^{-1})\,. \tag{1.195}$$

The negative first term corresponds to the potential renormalization which is no longer present in the ground state energy ε_0. The second term, on the other hand, is positive and thus leads to an increase of the ground state energy.

Not only the ground state energy can be derived from the partition function but one may also formally introduce a density of states $\rho(E)$ of the damped system according to [35]

$$\mathcal{Z}(\beta) = \int_0^\infty \mathrm{d}E \rho(E) \exp(-\beta E) \; . \tag{1.196}$$

Inversion of the Laplace transformation allows to determine $\rho(E)$ from the partition function according to

$$\rho(E) = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \mathrm{d}\beta \mathcal{Z}(\beta) \exp(\beta E) \tag{1.197}$$

where the constant c has to be chosen such that the line of integration is to the right of all poles of $\mathcal{Z}(\beta)$.

Once a high-frequency cutoff for the spectral density of bath oscillators is specified, the inverse Laplace transform in (1.197) may be evaluated either numerically or by contour integration. The second approach leads to a series which again has to be evaluated numerically [35]. However, it is not necessary to introduce a cutoff provided we shift the energy by the ground state energy ε_0 which in fact is the only divergent quantity in this problem. Such a shift may be performed by considering $\mathcal{Z} \exp(\beta\varepsilon_0)$ instead of $\mathcal{Z}$ itself. To demonstrate that this procedure renders the cutoff irrelevant, we will restrict ourselves to the limit of weak damping and large cutoff considered before even though a more general treatment is feasible.

In a first step we decompose the infinite product appearing in the partition function (1.185) with Drude cutoff (1.179) into a factor where the limit $\omega_\mathrm{D} \to \infty$ can safely be taken and a factor still containing the cutoff frequency

$$\mathcal{Z} = \frac{1}{\hbar\beta\omega_0} \prod_{n=1}^\infty \frac{\nu_n^2 + \gamma\nu_n}{\nu_n^2 + \gamma\nu_n + \omega_0^2} \prod_{n=1}^\infty \frac{1}{1 + \dfrac{\gamma\omega_\mathrm{D}}{\nu_n(\nu_n + \omega_\mathrm{D})}} \; . \tag{1.198}$$

It is the last product which has to be analyzed with care because it vanishes in the limit $\omega_D \to \infty$. To leading order in γ one finds

$$\begin{aligned} \ln\left(\prod_{n=1}^\infty 1 + \frac{\gamma\omega_\mathrm{D}}{\nu_n(\nu_n + \omega_\mathrm{D})} \right) &= \sum_{n=1}^\infty \frac{\gamma\omega_\mathrm{D}}{\nu_n(\nu_n + \omega_\mathrm{D})} \\ &= \frac{\hbar\beta\gamma}{2\pi} \psi\left(1 + \frac{\hbar\beta\omega_\mathrm{D}}{2\pi} \right) \end{aligned} \tag{1.199}$$

where we have introduced the digamma function [36]

$$\psi(1+z) = -\mathcal{C} + \sum_{n=1}^\infty \frac{z}{n(n+z)} \; . \tag{1.200}$$

Here, $\mathcal{C} = 0.577\ldots$ is the Euler constant. With the leading asymptotic behaviour $\psi(1+z) \sim \ln(z)$ for large arguments z, the partition function for large cutoff

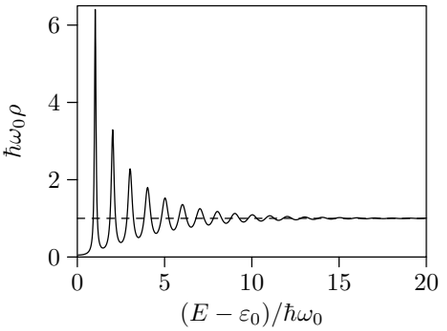


Fig. 1.11. The density of states (1.197) of the damped harmonic oscillator is shown for $\gamma/2\omega_0 = 0.05$. A delta function contribution at $E = \varepsilon_0$ has been omitted. The dashed line marks the average density of states $1/\hbar\omega_0$

frequency becomes

$$\mathcal{Z} = \frac{1}{\hbar\beta\omega_0}\prod_{n=1}^{\infty}\frac{\nu_n^2+\gamma\nu_n}{\nu_n^2+\gamma\nu_n+\omega_0^2}\left(\frac{\hbar\beta\omega_\mathrm{D}}{2\pi}\right)^{-\hbar\beta\gamma/2\pi}. \tag{1.201}$$

On the other hand, we know from our result (1.195) that the ground state energy diverges to leading order with $(\hbar\gamma/2\pi)\ln(\omega_\mathrm{D})$. Therefore, multiplication of the partition function with $\exp(\beta\varepsilon_0)$ leads indeed to an expression with a finite value in the limit of infinite cutoff frequency.

Before taking a look at numerical results, we remark that the partition function contains a pole at $\beta = 0$ with residue $1/\hbar\omega_0$. This pole represents the Laplace transform of the constant $1/\hbar\omega_0$ and therefore is related to the average density of states which takes the value of the undamped case where the energy spacing between adjacent levels is $\hbar\omega_0$. In Fig. 1.11 we present the density of states for weak damping, $\gamma/2\omega_0 = 0.05$. A delta function contribution at $E = \varepsilon_0$ has been omitted. Due to the weak damping we find well defined peaks which are close to the energies expected for an undamped oscillator. With increasing energy the levels become broader. For stronger damping, only the lowest levels can be resolved and a level shift induced by the damping becomes visible.

The behaviour of the level widths shown in Fig. 1.11 is consistent with the result of a perturbative treatment. According to Fermi's golden rule, the width of the n-th level is given by

$$\begin{aligned}\Gamma_n = \frac{2\pi}{\hbar^2}\sum_{j=1}^{\infty}\Big[&\big|\langle n+1,1_j|c_jqx_j|n,0\rangle\big|^2\delta(-\omega_0-\omega_j)\\ &+\big|\langle n-1,1_j|c_jqx_j|n,0\rangle\big|^2\delta(\omega_0-\omega_j)\Big]\end{aligned} \tag{1.202}$$

where we have already taken into account that the matrix element of the dipole-type coupling connects the state n only to its nearest neighbors. Because of

energy conservation the first part of the sum never contributes. With the matrix elements

$$\langle n-1,1_j|qx_j|n,0\rangle = \frac{\hbar}{2(mm_j\omega_0\omega_j)^{1/2}}n^{1/2} \tag{1.203}$$

we thus find for the width of the n-th level in terms of the spectral density of bath oscillators

$$\Gamma_n = \frac{n}{m\omega_0}J(\omega_0)\ . \tag{1.204}$$

For Ohmic damping, $J(\omega) = m\gamma\omega$, we finally get

$$\Gamma_n = n\gamma\ . \tag{1.205}$$

As observed before, the levels broaden with increasing damping strength γ and level number n. We remark that one can demonstrate by a semiclassical analysis of other one-dimensional potentials that it is indeed the level number and not the energy which is decisive for the level width [37].

1.4.3 Position Autocorrelation Function

In the introduction we have mentioned fluctuations as one of the effects arising from the coupling to an environment. Even if a system is in thermal equilibrium with its environment, fluctuations due to the noise term (1.128) will be present. The appropriate quantity to describe this phenomenon are equilibrium correlation functions like the position autocorrelation function

$$C(t) = \langle q(t)q(0)\rangle = \mathrm{Tr}(q(t)q(0)\rho_\beta)\ . \tag{1.206}$$

From this quantity one can derive all other equilibrium correlation functions of the damped harmonic oscillator as is discussed in [38].

We now want to determine this correlation function by first calculating its imaginary time version and start with the Euclidean action

$$\begin{aligned} S^{\mathrm{E}}[q] = \int_0^{\hbar\beta}\mathrm{d}\tau\left(\frac{m}{2}\dot{q}^2+\frac{m}{2}\omega_0^2q^2\right)+\frac{1}{2}\int_0^{\hbar\beta}\mathrm{d}\tau\int_0^{\hbar\beta}\mathrm{d}\sigma k(\tau-\sigma)q(\tau)q(\sigma) \\ +\int_0^{\hbar\beta}\mathrm{d}\tau F(\tau)q(\tau)\ . \end{aligned} \tag{1.207}$$

The second term accounts for the coupling to the environment as we have shown in Sect. 1.3.3. In addition, we have included the third term corresponding to an external force in imaginary time. This constitutes a useful trick which will allow us to determine the correlation function by variation with respect to this force

$$\langle q(\tau)q(\sigma)\rangle = \hbar^2\mathrm{Tr}\left(\frac{\delta}{\delta F(\tau)}\frac{\delta}{\delta F(\sigma)}\rho\right)\Bigg|_{F=0}\ . \tag{1.208}$$

As we know already from Sect. 1.2.7, the force does not appear in the fluctuation part. It is therefore sufficient, to restrict our attention to the classical path.

The classical equation of motion following from variation of the action (1.207)

$$m\ddot{q}(\tau) - \int_0^{\hbar\beta} \mathrm{d}\sigma k(\tau-\sigma)q(\sigma) - m\omega_0^2 q(\tau) = F(\tau) \tag{1.209}$$

is most conveniently solved by Fourier transformation on the interval from 0 to $\hbar\beta$. Introducing the Fourier transforms

$$q(\tau) = \frac{1}{\hbar\beta}\sum_{n=-\infty}^{\infty} q_n \exp(\mathrm{i}\nu_n\tau) \tag{1.210}$$

and

$$F(\tau) = \frac{m}{\hbar\beta}\sum_{n=-\infty}^{\infty} f_n \exp(\mathrm{i}\nu_n\tau) \tag{1.211}$$

and making use of the Fourier representation (1.178) of $k(\tau)$ for Ohmic damping, we find for the Fourier coefficients of the classical solution

$$q_n^{\mathrm{cl}} = -\frac{f_n}{\nu_n^2 + \gamma|\nu_n| + \omega_0^2}\ . \tag{1.212}$$

Inserting this result into the Fourier representation of the action (1.207)

$$S^{\mathrm{E}}[q] = \frac{m}{2\hbar\beta}\sum_{n=-\infty}^{\infty}\left[(\nu_n^2 + \gamma|\nu_n| + \omega_0^2)q_n q_{-n} + f_n q_{-n} + f_{-n}q_n\right] \tag{1.213}$$

yields the classical Euclidean action

$$S_{\mathrm{cl}}^{\mathrm{E}} = -\frac{m}{2\hbar\beta}\sum_{n=-\infty}^{\infty}\frac{f_n f_{-n}}{\nu_n^2 + \gamma|\nu_n| + \omega_0^2} \tag{1.214}$$

or equivalently

$$\begin{aligned} S_{\mathrm{cl}}^{\mathrm{E}} = -\frac{1}{2m\hbar\beta}\sum_{n=-\infty}^{\infty}&\frac{1}{\nu_n^2 + \gamma|\nu_n| + \omega_0^2} \\ &\times\int_0^{\hbar\beta}\mathrm{d}\tau\int_0^{\hbar\beta}\mathrm{d}\sigma F(\tau)F(\sigma)\exp\left(\mathrm{i}\nu_n(\tau-\sigma)\right)\ . \end{aligned} \tag{1.215}$$

Since the external force appears only through this action in the exponent of the equilibrium density matrix, we may easily evaluate the functional derivatives according to (1.208). For the position autocorrelation function in imaginary time we thus find

$$C(\tau) = \frac{1}{m\beta}\sum_{n=-\infty}^{\infty}\frac{\exp(\mathrm{i}\nu_n\tau)}{\nu_n^2 + \gamma|\nu_n| + \omega_0^2}\ . \tag{1.216}$$

Unfortunately, the real time correlation function cannot be obtained simply by replacing the imaginary time τ by $\mathrm{i}t$ where t is a real time. For negative times t,

the sum in (1.216) would not converge. We therefore have to perform an analytic continuation to real times in a more elaborate way.

The idea is to express the sum (1.216) as a contour integral in the complex frequency plane. To this end, we need a function with poles at frequencies $\omega = \mathrm{i}\nu_n, n = -\infty, \ldots, \infty$ with a residuum of one. This requirement is satisfied by $\hbar\beta/[1 - \exp(-\hbar\beta\omega)]$. By integration along the contour shown in Fig. 1.12a we find

$$\begin{aligned}\int_{\mathcal{C}^+} \mathrm{d}\omega \frac{\hbar\beta}{1-\exp(-\hbar\beta\omega)} \frac{\exp(-\omega\tau)}{-\omega^2 + \mathrm{i}\gamma\omega + \omega_0^2} \\ = -\mathrm{i}\frac{\pi}{\omega_0^2} - 2\pi\mathrm{i}\sum_{n=1}^{\infty} \frac{\exp(\mathrm{i}\nu_n\tau)}{\nu_n^2 + \gamma\nu_n + \omega_0^2} \ .\end{aligned} \tag{1.217}$$

Similarly, an integration along the contour shown in Fig. 1.12b leads to

$$\begin{aligned}\int_{\mathcal{C}^-} \mathrm{d}\omega \frac{\hbar\beta}{1-\exp(-\hbar\beta\omega)} \frac{\exp(-\omega\tau)}{-\omega^2 - \mathrm{i}\gamma\omega + \omega_0^2} \\ = \mathrm{i}\frac{\pi}{\omega_0^2} + 2\pi\mathrm{i}\sum_{n=-\infty}^{-1} \frac{\exp(\mathrm{i}\nu_n\tau)}{\nu_n^2 + \gamma\nu_n + \omega_0^2}\end{aligned} \tag{1.218}$$

Subtracting (1.218) from (1.217), the imaginary time correlation function can be expressed as

$$\begin{aligned}\frac{1}{m\beta}\sum_{n=-\infty}^{\infty} \frac{\exp(\mathrm{i}\nu_n\tau)}{\nu_n^2 + \gamma\nu_n + \omega_0^2} \\ = \frac{\hbar}{m\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \frac{\gamma\omega}{(\omega^2-\omega_0^2)^2 + \gamma^2\omega^2} \frac{\exp(-\omega\tau)}{1-\exp(-\hbar\beta\omega)} \ .\end{aligned} \tag{1.219}$$

Now we may pass to real times by the replacement $\tau \to \mathrm{i}t$ to obtain the real time correlation function

$$C(t) = \frac{\hbar}{m\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \frac{\gamma\omega}{(\omega^2-\omega_0^2)^2 + \gamma^2\omega^2} \frac{\exp(-\mathrm{i}\omega t)}{1-\exp(-\hbar\beta\omega)} \ . \tag{1.220}$$

Physical insight into this result can be gained by considering the Fourier transform of this correlation function

$$\tilde{C}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t \exp(\mathrm{i}\omega t) C(t) \tag{1.221}$$

which may be related to the dynamic susceptibility $\tilde{\chi}(\omega)$ of the damped harmonic oscillator. According to the Ehrenfest theorem the equation of motion for the expectation value of the position agrees with the corresponding classical equation of motion. In the presence of an external force $F(t)$, the latter reads

$$m\ddot{q} + m\gamma\dot{q} + m\omega_0^2 q = F(t) \ . \tag{1.222}$$

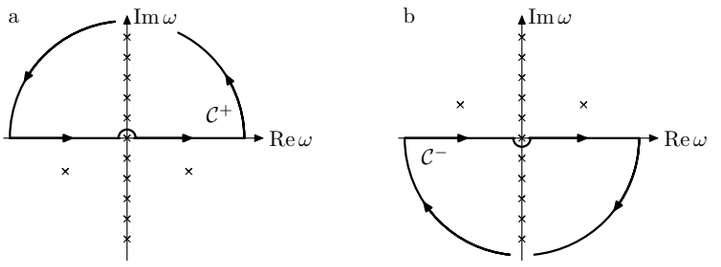


Fig. 1.12. The analytic continuation of the imaginary time correlation function $\langle q(\tau)q(0)\rangle$ to real times makes use of the integration contours depicted in (**a**) and (**b**) to obtain (1.217) and (1.218), respectively

The solution of this equation may be expressed in terms of the response function $\chi(t)$ as

$$q(t) = \int_{-\infty}^{t} \mathrm{d}s \chi(t-s)F(s) \tag{1.223}$$

which by means of a Fourier transformation becomes

$$\tilde{q}(\omega) = \tilde{\chi}(\omega)\tilde{F}(\omega) \; . \tag{1.224}$$

With the equation of motion (1.222) the dynamic susceptibility of the damped harmonic oscillator is then found to read

$$\chi(\omega) = \frac{1}{m}\frac{1}{-\omega^2 - \mathrm{i}\gamma\omega + \omega_0^2} \; . \tag{1.225}$$

Together with (1.220) and (1.221) one finally obtains the relation

$$\tilde{C}(\omega) = \frac{2\hbar}{1-\exp(-\hbar\beta\omega)}\chi''(\omega) \tag{1.226}$$

which represents an example of the so-called fluctuation-dissipation theorem [39]. Here, the position autocorrelation function describes the fluctuations while the imaginary part of the dynamic susceptibility χ'' can be shown to determine the energy dissipated under the influence of the external driving $F(t)$. While the relation (1.226) is exact for linear systems like the damped harmonic oscillator considered here, it still holds for nonlinear systems within linear response theory. There, the response to the external force is considered in leading order perturbation theory [40,41].

It is instructive to consider the real time correlation function (1.220) in more detail. We first decompose the correlation function $C(t)$ into its real and imaginary part, or equivalently, into its symmetric and antisymmetric part

$$C(t) = S(t) + \mathrm{i}A(t) \; , \tag{1.227}$$

with

$$S(t) = \frac{1}{2}\big(\langle q(t)q(0)\rangle + \langle q(0)q(t)\rangle\big) \tag{1.228}$$

and

$$A(t) = -\frac{\mathrm{i}}{2}\big(\langle q(t)q(0)\rangle - \langle q(0)q(t)\rangle\big) \; . \tag{1.229}$$

From (1.220) we find

$$S(t) = \frac{\hbar}{2\pi m}\int_{-\infty}^{\infty}\mathrm{d}\omega \frac{\gamma\omega}{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2}\coth\left(\frac{\hbar\beta\omega}{2}\right)\cos(\omega t) \tag{1.230}$$

and

$$A(t) = -\frac{\hbar}{2\pi m}\int_{-\infty}^{\infty}\mathrm{d}\omega \frac{\gamma\omega}{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2}\sin(\omega t) \; . \tag{1.231}$$

Apart from Planck's constant appearing in the prefactor, the antisymmetric part is purely classical. In fact, one demonstrates within linear response theory the general relation

$$\chi(t) = \frac{\mathrm{i}}{\hbar}\langle[q(t),q(0)]\rangle\Theta(t) = -\frac{2}{\hbar}A(t)\Theta(t) \; . \tag{1.232}$$

Therefore, our statement follows as a consequence of the Ehrenfest theorem which ensures that the response function $\chi(t)$ of the damped harmonic oscillator is classical.

More interesting is the symmetric part (1.230) of the correlation function $C(t)$. There exist two different types of time scales determined by the poles of the integrand in (1.230). One set of four poles at frequencies

$$\omega = \pm\left(\bar{\omega} \pm \mathrm{i}\frac{\gamma}{2}\right) \tag{1.233}$$

corresponds to the characteristic frequencies of a damped harmonic oscillator with the oscillation frequency

$$\bar{\omega} = \left(\omega_0^2 - \frac{\gamma^2}{4}\right)^{1/2} \tag{1.234}$$

shifted by the damping. In addition, there exists an infinite number of poles at imaginary frequencies $\mathrm{i}\nu_n, n = -\infty, \ldots, \infty$ depending on the temperature of the heat bath via the Matsubara frequencies defined in (1.107). With these poles, one can evaluate the integral (1.230) by means of a contour integration to obtain

$$\begin{aligned} S(t) = {} & \frac{\hbar}{2m\bar{\omega}}\exp\left(-\frac{\gamma|t|}{2}\right)\frac{[\sinh(\hbar\beta\bar{\omega})\cos(\bar{\omega}t)+\sin(\hbar\beta\gamma/2)\sin(\bar{\omega}|t|)]}{\cosh(\hbar\beta\bar{\omega})-\cos(\hbar\beta\gamma/2)} \\ & - \frac{2\gamma}{m\beta}\sum_{n=1}^{\infty}\frac{\nu_n\exp(-\nu_n|t|)}{(\nu_n^2+\omega_0^2)^2-\gamma^2\nu_n^2} \; . \end{aligned} \tag{1.235}$$

The sum in the second line becomes important at low temperatures $k_{\mathrm{B}}T \ll \hbar\gamma/4\pi$. In order to discuss this quantum effect, we focus in the following discussion on the case of zero temperature. Then, the poles of the hyperbolic cotangent

in (1.230) become dense and form a cut on the imaginary axis which has consequences for the long-time behaviour of the correlation function. In the limit $\beta \to \infty$ the sum in the second line of (1.235) turns into an integral. Noting that the long-time behaviour is determined by the behaviour of the integrand at small arguments we find

$$S(t) \sim -\frac{\hbar\gamma}{\pi m}\int_0^\infty \mathrm{d}x \frac{x\exp(-x|t|)}{\omega_0^4} = -\frac{\hbar\gamma}{\pi m\omega_0^4}\frac{1}{t^2}\ . \tag{1.236}$$

Instead of the usual exponential decay we thus find an algebraic decay.

In the limit of vanishing damping, the imaginary part of the dynamic susceptibility appearing in the integrand of (1.230) turns into a sum of two delta functions located at ω_0 and $-\omega_0$. For weak but finite damping, the delta functions broaden into Lorentzians

$$\frac{\gamma\omega}{(\omega^2-\omega_0^2)^2+\gamma^2\omega^2} = \frac{\gamma}{4\bar{\omega}}\left(\frac{1}{(\omega-\bar{\omega})^2+\gamma^2/4} - \frac{1}{(\omega+\bar{\omega})^2+\gamma^2/4}\right) \tag{1.237}$$

corresponding to the four poles (1.233) and one can assume that only the integrand in the neighbourhood of these poles is relevant. Within the so-called Markov approximation, one then replaces the contributions to the integrand which depend only weakly on frequency by their values at $\bar{\omega}$. As we will see, in contrast to (1.236) the correlation function $S(t)$ at zero temperature then no longer decays algebraically. It is interesting to discuss the reason for this discrepancy.

To this end we go back to the integral representation (1.230) of the correlation function $S(t)$. In a first step, we apply the so-called rotating wave approximation (RWA) which consists in neglecting the Lorentzian located at $-\bar{\omega}$, i.e. the second term in (1.237) [42]. Physically speaking, we neglect processes where the system is excited into an energetically higher state by loosing energy to the driving force or where the system decays to a lower state by absorbing energy. For finite temperatures, we now have

$$S_{\mathrm{RWA}}(t) = \frac{\hbar}{8\pi m\bar{\omega}}\int_{-\infty}^\infty \mathrm{d}\omega \frac{\gamma}{(\omega-\bar{\omega})^2+\gamma^2/4}\coth\left(\frac{\hbar\beta\omega}{2}\right)\cos(\omega t)\ . \tag{1.238}$$

Within the Markov approximation, we replace the hyperbolic cotangent by its value at $\omega = \bar{\omega}$. In the limit of zero temperature this leads to

$$\begin{aligned} S_{\mathrm{RWA,Markov}}(t) &= \frac{\hbar}{8\pi m\bar{\omega}}\int_{-\infty}^\infty \mathrm{d}\omega \frac{\gamma}{(\omega-\bar{\omega})^2+\gamma^2/4}\cos(\omega t) \\ &= \frac{\hbar}{4m\bar{\omega}}\cos(\bar{\omega}t)\exp\Big(-\frac{\gamma}{2}t\Big)\ . \end{aligned} \tag{1.239}$$

We thus find an oscillation with a frequency shifted due to the environmental coupling and an exponential decay in contrast to the algebraic decay (1.236).

This difference can be understood by taking a closer look at the Markov approximation. In Fig. 1.13 the Lorentzian and the hyperbolic cotangent appearing in (1.238) are schematically shown as full and dashed line, respectively.

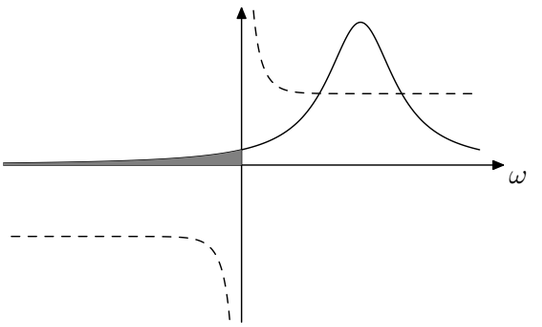


Fig. 1.13. The full and dashed lines represent the Lorentzian and hyperbolic cotangent, respectively, which contribute to the integrand in (1.238)

In order to obtain (1.239), we have approximated the hyperbolic cotangent by a constant. However, in doing this, we have replaced an antisymmetric function by a symmetric one which can only yield a non-vanishing result together with the rotating wave approximation made above. As a consequence, the area shaded in grey has been taken with the wrong sign. The idea was that this difference should be small and arising from frequencies far from $\bar{\omega}$. However, due to the large extent of a Lorentzian, it is of order γ and, in addition, it replaces an exponential decay by an algebraic decay.

At zero temperature the difference between (1.238) and (1.239) becomes

$$\begin{aligned} \Delta &= S_{\mathrm{RWA}}(t) - S_{\mathrm{RWA,Markov}}(t) \\ &= -2\frac{\hbar}{8\pi m}\int_{-\infty}^{0} \mathrm{d}\omega \frac{\gamma}{(\omega-\bar{\omega})^2+\gamma^2/4}\cos(\omega t) \end{aligned} \tag{1.240}$$

which may be expressed in terms of integral sine and integral cosine functions. For our purpose it is sufficient to note that for long times, the difference indeed decays algebraically as

$$\Delta = -\frac{\hbar\gamma}{2\pi m\bar{\omega}_0^4}\frac{1}{t^2}\,. \tag{1.241}$$

We remark that the factor of two relative to the result (1.236) arises because no rotating wave approximation has been made in deriving the latter.

In the previous discussion, we have been concerned with an effect of order γ which in the spirit of the Markov approximation should be considered as small. However, the Markov approximation changes the long-time behaviour of the equilibrium correlation function $S(t)$ qualitatively and the absence of an algebraic decay, even with a small prefactor, may be relevant. For truly weak coupling, the damping constant γ should be the smallest frequency scale. Apart from the usual weak coupling condition $\gamma \ll \omega_0$ we also have to require $\gamma \ll k_{\mathrm{B}}T/\hbar$ which at low temperatures may become difficult to satisfy.

Acknowledgment

This chapter is based on courses taught at the Max-Planck-Institut für Physik komplexer Systeme in Dresden and at the Université Louis Pasteur Strasbourg. The author is grateful to the organizers of the school on "Coherent evolution in noisy environments", in particular Andreas Buchleitner (Dresden), as well as Rodolfo Jalabert (ULP Strasbourg) for having provided the opportunity to lecture on path integrals and quantum dissipation. The participants of the two courses through their questions contributed to these lecture notes. It is a pleasure to acknowledge inspiring conversations with Giovanna Morigi which have motivated some of the discussions in particular in the last section. Furthermore, I am indebted to Peter Hänggi for several useful comments on an earlier version of the manuscript and for drawing my attention to some less known references.

References

1. U. Weiss: *Quantum Dissipative Systems* (World Scientific 1999)
2. T. Dittrich, P. Hänggi, G.-L. Ingold, B. Kramer, G. Schön, W. Zwerger: *Quantum transport and dissipation* (Wiley-VCH 1998)
3. H. Kleinert: *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics* (World Scientific 1995)
4. R. P. Feynman: Rev. Mod. Phys. **20**, 367 (1948)
5. For a few historical remarks see e.g. D. Derbes: Am. J. Phys. **64**, 881 (1996)
6. P. A. M. Dirac: Phys. Zs. Sowjetunion **3**, 64 (1933)
7. see e.g. E. Nelson, J. Math. Phys. **5**, 332 (1964), Appendix B
8. F. Langouche, D. Roekaerts, E. Tirapegui: *Functional Integration and Semiclassical Expansions*, Mathematics and Its Applications Vol. 10 (D. Reidel 1982)
9. W. Janke, H. Kleinert: Lett. Nuovo Cim. **25**, 297 (1979)
10. M. Goodman: Am. J. Phys. **49**, 843 (1981)
11. A. Auerbach, L. S. Schulman: J. Phys. A **30**, 5993 (1997)
12. I. S. Gradshteyn, I. M. Rhyzik: *Table of Integrals, Series, and Products* (Academic Press 1980)
13. V. P. Maslov, M. V. Fedoriuk: *Semi-classical approximation in quantum mechanics*. Mathematical Physics and Applied Mathematics Vol. 7, ed. by M. Flato et al. (D. Reidel 1981)
14. P. A. Horváthy: Int. J. Theor. Phys. **18**, 245 (1979)
15. D. Rohrlich: Phys. Lett. A **128**, 307 (1988)
16. M. S. Marinov: Phys. Rep. **60**, 1 (1980), Chap. 3.3
17. see e.g. M. Brack, R. K. Bhaduri: *Semiclassical Physics*, Frontiers in Physics Vol. 96 (Addison-Wesley 1997)
18. J. H. van Vleck: Proc. Natl. Acad. Sci. USA **14**, 178 (1928)
19. C. Morette: Phys. Rev. **81**, 848 (1951)
20. W. Pauli: *Selected Topics in Field Quantization*, Pauli Lectures on Physics Vol. 6, ed. by C. P. Enz (Dover 2000)
21. N. Wiener: J. Math. Phys. M.I.T. **2**, 131 (1923)
22. A. O. Caldeira, A. J. Leggett: Phys. Rev. Lett. **46**, 211 (1981)
23. A. O. Caldeira, A. J. Leggett: Ann. Phys. (N.Y.) **149**, 374 (1983)

24. V. B. Magalinskiĭ: Zh. Eksp. Teor. Fiz. **36**, 1942 (1959) [Sov. Phys. JETP **9**, 1381 (1959)]
25. I. R. Senitzky: Phys. Rev. **119**, 670 (1960); **124**, 642 (1961)
26. G. W. Ford, M. Kac, P. Mazur: J. Math. Phys. **6**, 504 (1965)
27. P. Ullersma: Physica **32**, 27, 56, 74, 90 (1966)
28. R. Zwanzig: J. Stat. Phys. **9**, 215 (1973)
29. P. C. Hemmer, L. C. Maximon, H. Wergeland: Phys. Rev. **111**, 689 (1958)
30. P. Hänggi: 'Generalized Langevin Equations: A Useful Tool for the Perplexed Modeller of Nonequilibrium Fluctuations?'. In: *Lecture Notes in Physics Vol. 484*, ed. by L. Schimansky-Geier, T. Pöschel (Springer, Berlin 1997) pp. 15–22
31. V. Hakim, V. Ambegaokar: Phys. Rev. A **32**, 423 (1985)
32. H. Grabert, P. Schramm, G.-L. Ingold: Phys. Rev. Lett. **58**, 1285 (1987)
33. R. P. Feynman, F. L. Vernon, Jr.: Ann. Phys. (N.Y.) **24**, 118 (1963)
34. H. Grabert, P. Schramm, G.-L. Ingold: Phys. Rep. **168**, 115 (1988)
35. A. Hanke, W. Zwerger: Phys. Rev. E **52**, 6875 (1995)
36. M. Abramowitz, I. A. Stegun (eds.): *Handbook of mathematical functions* (Dover 1972)
37. G.-L. Ingold, R. A. Jalabert, K. Richter: Am. J. Phys. **69**, 201 (2001)
38. H. Grabert, U. Weiss, P. Talkner: Z. Phys. B **55**, 87 (1984)
39. H. B. Callen, T. A. Welton: Phys. Rev. **83**, 34 (1951)
40. R. Kubo: J. Phys. Soc. Japan **12**, 570 (1957)
41. R. Kubo: Rep. Progr. Phys. **29**, 255 (1966)
42. R. Loudon: *The Quantum Theory of Light* (Oxford University Press 1983)

2 Five Lectures on Dissipative Master Equations

Berthold-Georg Englert[1,2] and Giovanna Morigi[1,3]

[1] Max-Planck-Institut für Quantenoptik,
Hans-Kopfermann-Straße 1, 85748 Garching, Germany
[2] Department of Mathematics and Department of Physics, Texas A&M University,
College Station, TX 77843-4242, U. S. A.
[3] Abteilung Quantenphysik, Universität Ulm,
Albert-Einstein-Allee 11, 89081 Ulm, Germany

Introductory Remarks

The damped harmonic oscillator is arguably the simplest open quantum system worth studying. It is also of great practical importance because it is an essential ingredient in the theoretical description of many quantum-optical experiments. One can assume rather safely that the quantum master equation of the simple harmonic oscillator would not be studied so extensively if it did not play such a central role in the quantum theory of lasers and the masers. Not surprisingly, then, all major textbook accounts of theoretical quantum optics [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15] contain a fair amount of detail about damped harmonic oscillators. Fock state representations or phase space functions of some sort are invariably employed in these treatments.

The algebraic methods on which we will focus here are quite different. They should be regarded as a supplement of, not as a replacement for, the traditional approaches. As always, every method has its advantages and its drawbacks: a particular problem can be technically demanding in one approach, but quite simple in another. This is, of course, also true for the algebraic method. We will illustrate its technical power by a few typical examples for which the standard approaches would be quite awkward.

2.1 First Lecture: Basics

The evolution of a simple damped harmonic oscillator is governed by the master equation

$$\frac{\partial}{\partial t}\rho_t = \mathrm{i}\omega[\rho_t, a^\dagger a] - \frac{1}{2}A(\nu+1)\left(a^\dagger a\rho_t - 2a\rho_t a^\dagger + \rho_t a^\dagger a\right) - \frac{1}{2}A\nu\left(aa^\dagger\rho_t - 2a^\dagger\rho_t a + \rho_t aa^\dagger\right) , \tag{2.1}$$

where $a^\dagger, a$ are the ladder operators of the oscillator; ω is its natural (circular) frequency; A is the energy decay rate; and ν is the number of thermal excitations

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 55–106, 2002.

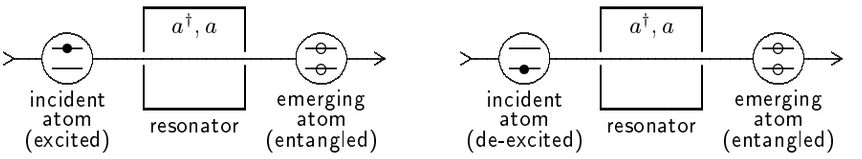


Fig. 2.1. A two-level atom traverses a high-quality cavity, coupling resonantly to a privileged photon mode of the cavity. Prior to the interaction, there is some initial photon state in the resonator and the atom is either in the upper state of the pertinent transition (on the left) or in the lower state (on the right). After the interaction, the transition degree of the atom and the photon degree of the cavity are entangled

in the steady state that the statistical operator $\rho_t \equiv \rho_t\big(a^\dagger, a\big)$ approaches for very late times t. We will have much to say about the properties of the solutions of (2.1), but first we would like to give a physical derivation of this equation.

2.1.1 Physical Derivation of the Master Equation

For this purpose we consider the following model. The oscillator is a mode of the quantized radiation field of an ideal resonator, so that excitations of this mode (*vulgo* photons) would stay in the resonator forever. In reality they have a finite lifetime, of course, and we describe this fact by letting the photons interact with atoms that pass through the resonator. As is depicted in Fig. 2.1, these atoms are also of the simplest kind conceivable: they only have two levels which – another simplification – are separated in energy by $\hbar\omega$, the energy per photon in the privileged resonator mode. Incident atoms in the upper level (symbolically: $\overset{\bullet}{\underline{}}$) will have a chance to deposit energy into the resonator, while those in the lower level ($\overset{}{\underline{\bullet}}$) will tend to remove energy.

The evolution of the interacting atom-photon system is governed by the Hamilton operator

$$H = \hbar\omega a^\dagger a + \hbar\omega\sigma^\dagger\sigma - \hbar g(\sigma a^\dagger + \sigma^\dagger a) \,, \tag{2.2}$$

which goes with the name "resonant Jaynes–Cummings interaction in the rotating-wave approximation" in the quantum-optical literature. It applies as long as the atom is inside the resonator and is replaced by

$$H_{\rm free} = \hbar\omega a^\dagger a + \hbar\omega\sigma^\dagger\sigma \tag{2.3}$$

before and after the period of interaction. Here $\sigma^\dagger$ and σ are the atomic ladder operators,

$$\sigma^\dagger = |\overset{\bullet}{\underline{}}\rangle\langle\overset{}{\underline{\bullet}}| \mathrel{\widehat{=}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \,, \qquad \sigma = |\overset{}{\underline{\bullet}}\rangle\langle\overset{\bullet}{\underline{}}| \mathrel{\widehat{=}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \,, \tag{2.4}$$

and g is the so-called Rabi frequency, the measure of the interaction strength. Note that $\sigma^\dagger\sigma$ and $\sigma\sigma^\dagger$ project to the upper and lower atomic states,

$$\sigma^\dagger\sigma = |\overset{\bullet}{\underline{}}\rangle\langle\overset{\bullet}{\underline{}}| \mathrel{\widehat{=}} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \,, \qquad \sigma\sigma^\dagger = |\overset{}{\underline{\bullet}}\rangle\langle\overset{}{\underline{\bullet}}| \mathrel{\widehat{=}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \,, \tag{2.5}$$

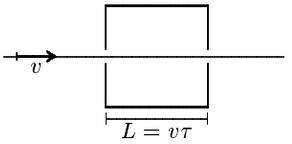


Fig. 2.2. An atom takes time $\tau = L/v$ to traverse a cavity of length L at speed v

respectively.

The interaction term in (2.2) is a multiple of the coupling operator

$$\gamma = a\sigma^\dagger + a^\dagger\sigma \tag{2.6}$$

and H_{free} is essentially the square of γ since

$$\gamma^2 = a^\dagger a + \sigma^\dagger\sigma \ . \tag{2.7}$$

So (2.2) and (2.3) can be rewritten as

$$H = \hbar\omega\gamma^2 - \hbar g\gamma \ , \qquad H_{\text{free}} = \hbar\omega\gamma^2 \ . \tag{2.8}$$

That H commutes with H_{free} and both are just simple functions of γ will enable us to solve the equations of motion quite explicitly without much effort.

We denote by ρ_t the statistical operator describing the combined atom-cavity system. It is a function of the dynamical variables $a^\dagger$, a, $\sigma^\dagger$, σ and has also a parametric dependence on t, indicated by the subscript,

$$\rho_t = \rho_t\big(a^\dagger(t), a(t), \sigma^\dagger(t), \sigma(t)\big) \ . \tag{2.9}$$

Since a statistical operator has no total time dependence, Heisenberg's equation of motion,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\rho_t = \frac{\partial}{\partial t}\rho_t - \frac{\mathrm{i}}{\hbar}\big[\rho_t, H\big] \ , \tag{2.10}$$

becomes von Neumann's equation for the parametric time dependence,

$$\frac{\partial}{\partial t}\rho_t = \frac{\mathrm{i}}{\hbar}[\rho_t, H] \ . \tag{2.11}$$

Now suppose that t is the instant at which the atom enters the cavity; then it emerges at time $t + \tau$, and we have

$$\rho_{t+\tau} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H\tau}\rho_t\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}H\tau} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_{\text{free}}\tau}\left[\mathrm{e}^{\mathrm{i}\phi\gamma}\rho_t\,\mathrm{e}^{-\mathrm{i}\phi\gamma}\right]\mathrm{e}^{\frac{\mathrm{i}}{\hbar}H_{\text{free}}\tau} \tag{2.12}$$

after we use (2.8) and introduce the abbreviation $\phi = g\tau$. This phase ϕ is the accumulated Rabi angle and, for atoms moving classically through the cavity of length L with constant velocity v, we have $\tau = L/v$ and $\phi = gL/v$; see Fig. 2.2.

Clearly, the $[\dots]$ term in (2.12) accounts for the net effect of the interaction, that part of the evolution that happens in addition to the free evolution generated by H_{free}. We have

$$\mathrm{e}^{\mathrm{i}\phi\gamma} = \cos(\phi\gamma) + \mathrm{i}\sin(\phi\gamma) = \cos\left(\phi\sqrt{\gamma^2}\right) + \mathrm{i}\gamma\frac{\sin\left(\phi\sqrt{\gamma^2}\right)}{\sqrt{\gamma^2}} \ , \tag{2.13}$$

which is just saying that the cosine is an even function and the sine is odd. Further we note that the identities

$$\begin{aligned} F\left(\gamma^2\right) &= \sigma^\dagger\sigma F\left(aa^\dagger\right) + \sigma\sigma^\dagger F\left(a^\dagger a\right) \ , \\ \gamma F\left(\gamma^2\right) &= \sigma a^\dagger F\left(aa^\dagger\right) + F\left(aa^\dagger\right) a\sigma^\dagger \end{aligned} \tag{2.14}$$

hold for all functions $F\big(\gamma^2\big)$. They are immediate implications of familiar relations such as $af(a^\dagger a) = f(aa^\dagger)a$, $\sigma f(\sigma^\dagger\sigma) = \sigma f(1)$, and $\sigma^\dagger f(\sigma^\dagger\sigma) = \sigma^\dagger f(0)$. We use (2.13) and (2.14) to arrive at

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\phi\gamma} = \sigma^\dagger\sigma\cos\left(\phi\sqrt{aa^\dagger}\right) + \sigma\sigma^\dagger\cos\left(\phi\sqrt{a^\dagger a}\right) \\ + \mathrm{i}\sigma a^\dagger\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}} + \mathrm{i}\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}}a\sigma^\dagger \ . \end{aligned} \tag{2.15}$$

In terms of the 2×2 matrix representation for the σ's that is suggested in (2.4) and (2.5), this has the compact form

$$\mathrm{e}^{\mathrm{i}\phi\gamma} \mathrel{\widehat{=}} \begin{pmatrix} C & \mathrm{i}S^\dagger \\ \mathrm{i}S & \widetilde{C} \end{pmatrix} \tag{2.16}$$

with the photon operators

$$C = \cos\left(\phi\sqrt{aa^\dagger}\right) \ , \qquad \widetilde{C} = \cos\left(\phi\sqrt{a^\dagger a}\right) \ , \qquad S = a^\dagger\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}} \ , \tag{2.17}$$

and the adjoint of (2.16) reads

$$\mathrm{e}^{-\mathrm{i}\phi\gamma} \mathrel{\widehat{=}} \begin{pmatrix} C & -\mathrm{i}S^\dagger \\ -\mathrm{i}S & \widetilde{C} \end{pmatrix} \ . \tag{2.18}$$

We use these results for calculating the net effect of the interaction with one atom on the statistical operator $\rho^{(\mathrm{ph})}$ of the photon state. Initially the total state $\rho_t = \rho_t^{(\mathrm{ph})}\rho_t^{(\mathrm{at})}$ is not entangled, it is a product of the statistical operators referring respectively to the photons by themselves and the atom by itself. At the final instant $t+\tau$, we get $\rho_{t+\tau}^{(\mathrm{ph})}$ by tracing over the two-level atom,

$$\rho_{t+\tau}^{(\mathrm{ph})} = \mathrm{tr}_{\mathrm{at}}\left\{\rho_{t+\tau}\right\} = \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\ \mathrm{tr}_{\mathrm{at}}\left\{\mathrm{e}^{\mathrm{i}\phi\gamma}\rho_t^{(\mathrm{ph})}\rho_t^{(\mathrm{at})}\,\mathrm{e}^{-\mathrm{i}\phi\gamma}\right\}\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a} \ . \tag{2.19}$$

To proceed further we need to specify the initial atomic state $\rho_t^{(\mathrm{at})}$, and for the present purpose the two situations of Fig. 2.1 will do.

On the left of Fig. 2.1 we have $\overset{\bullet}{-}$ atoms arriving,

$$\rho_t^{(\mathrm{at})} = |\overset{\bullet}{-}\rangle\langle\overset{\bullet}{-}| = \sigma^\dagger\sigma \mathrel{\widehat{=}} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0) \ , \tag{2.20}$$

and (2.19) tells us that

$$\begin{aligned}\rho_{t+\tau}^{(\mathrm{ph})} &= \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\ \mathrm{tr}_{2\times 2}\left\{\begin{pmatrix} C \\ \mathrm{i}S\end{pmatrix}\rho_t^{(\mathrm{ph})}(C,-\mathrm{i}S^\dagger)\right\}\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a} \\ &= \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\left(C\rho_t^{(\mathrm{ph})}C + S\rho_t^{(\mathrm{ph})}S^\dagger\right)\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a}\ . \end{aligned} \tag{2.21}$$

Likewise, in the situation on the right-hand side of Fig. 2.1 we have

$$\rho_t^{(\mathrm{at})} = |\overline{\bullet}\rangle\langle\overline{\bullet}| = \sigma\sigma^\dagger \mathrel{\widehat{=}} \begin{pmatrix} 0 & 0 \\ 0 & 1\end{pmatrix} = \begin{pmatrix} 0 \\ 1\end{pmatrix}(0,1) \tag{2.22}$$

and get

$$\begin{aligned}\rho_{t+\tau}^{(\mathrm{ph})} &= \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\ \mathrm{tr}_{2\times 2}\left\{\begin{pmatrix} \mathrm{i}S^\dagger \\ \widetilde{C}\end{pmatrix}\rho_t^{(\mathrm{ph})}(-\mathrm{i}S,\widetilde{C})\right\}\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a} \\ &= \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\left(\widetilde{C}\rho_t^{(\mathrm{ph})}\widetilde{C} + S^\dagger\rho_t^{(\mathrm{ph})}S\right)\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a}\ . \end{aligned} \tag{2.23}$$

We remember our goal of modeling the coupling of the photons to a reservoir, and therefore we want to identify the effect of very many atoms traversing the cavity (one by one) but with each atom coupled very weakly to the photons. Weak atom-photon interaction means a small value of ϕ so that only the terms of lowest order in ϕ will be relevant. Since the $\phi = 0$ version of both (2.21) and (2.23), that is:

$$\rho_{t+\tau}^{(\mathrm{ph})} = \mathrm{e}^{-\mathrm{i}\omega\tau a^\dagger a}\rho_t^{(\mathrm{ph})}\,\mathrm{e}^{\mathrm{i}\omega\tau a^\dagger a}\ , \tag{2.24}$$

is just the free evolution of $\rho_t^{(\mathrm{ph})}$, the additional change in $\rho_t^{(\mathrm{ph})}$ that results from a single atom is

$$\begin{aligned}\Delta_1\rho_t^{(\mathrm{ph})} &= C\rho_t^{(\mathrm{ph})}C + S\rho_t^{(\mathrm{ph})}S^\dagger - \rho_t^{(\mathrm{ph})} \\ &= \cos\left(\phi\sqrt{aa^\dagger}\right)\rho_t^{(\mathrm{ph})}\cos\left(\phi\sqrt{aa^\dagger}\right) + a^\dagger\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}}\rho_t^{(\mathrm{ph})}\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}}a \\ &\quad - \rho_t^{(\mathrm{ph})} \\ &= -\frac{1}{2}\phi^2\left(aa^\dagger\rho_t^{(\mathrm{ph})} - 2a^\dagger\rho_t^{(\mathrm{ph})}a + \rho_t^{(\mathrm{ph})}aa^\dagger\right) + \mathcal{O}(\phi^4) \end{aligned} \tag{2.25}$$

for a $\underline{\bullet}$ atom arriving, and

$$\begin{aligned}\Delta_2\rho_t^{(\mathrm{ph})} &= \widetilde{C}\rho_t^{(\mathrm{ph})}\widetilde{C} + S^\dagger\rho_t^{(\mathrm{ph})}S - \rho_t^{(\mathrm{ph})} \\ &= \cos\left(\phi\sqrt{a^\dagger a}\right)\rho_t^{(\mathrm{ph})}\cos\left(\phi\sqrt{a^\dagger a}\right) + \frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}}a\rho_t^{(\mathrm{ph})}a^\dagger\frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}} \\ &\quad - \rho_t^{(\mathrm{ph})} \\ &= -\frac{1}{2}\phi^2\left(a^\dagger a\rho_t^{(\mathrm{ph})} - 2a\rho_t^{(\mathrm{ph})}a^\dagger + \rho_t^{(\mathrm{ph})}a^\dagger a\right) + \mathcal{O}(\phi^4) \end{aligned} \tag{2.26}$$

for a $\overline{\bullet}$ atom. So, for weak atom-photon interaction the relevant terms in (2.25) and (2.26) are the ones proportional to ϕ^2.

Atoms arriving at statistically independent times (Poissonian statistics for the arrival times) will thus induce a rate of change of $\rho_t^{(\mathrm{ph})}$ that is given by

$$\begin{aligned}\frac{\partial}{\partial t}\rho_t^{(\mathrm{ph})}\Big|_{\mathrm{weak}} &= r_1\Delta_1\rho_t^{(\mathrm{ph})} + r_2\Delta_2\rho_t^{(\mathrm{ph})} \\ &= -\frac{1}{2}r_1\phi^2\left(aa^\dagger\rho_t^{(\mathrm{ph})} - 2a^\dagger\rho_t^{(\mathrm{ph})}a + \rho_t^{(\mathrm{ph})}aa^\dagger\right) \\ &\quad -\frac{1}{2}r_2\phi^2\left(a^\dagger a\rho_t^{(\mathrm{ph})} - 2a\rho_t^{(\mathrm{ph})}a^\dagger + \rho_t^{(\mathrm{ph})}a^\dagger a\right) \end{aligned} \tag{2.27}$$

where r_1 and r_2 are the arrival rates for the $\underline{\bullet}$ atoms and the $\overline{\bullet}$ atoms, respectively. This is to say that during a period of duration T there will arrive on average r_1T atoms in state $\underline{\bullet}$ and r_2T atoms in state $\overline{\bullet}$.

Since the weak interaction with many atoms is supposed to simulate the coupling to a thermal bath (temperature Θ), these rates must be related to each other by a Maxwell–Boltzmann factor,

$$\frac{r_1}{r_2} = \exp\left(-\frac{\hbar\omega}{k_{\mathrm{B}}\Theta}\right) = \frac{\nu}{\nu+1} \tag{2.28}$$

where $\nu > 0$ is a convenient parameterization of the temperature. Also for matters of convenience, we introduce a rate parameter A by writing $r_1\phi^2 = A\nu$, $r_2\phi^2 = A(\nu+1)$ and arrive at

$$\begin{aligned}\frac{\partial}{\partial t}\rho_t &= \frac{\partial}{\partial t}\rho_t\Big|_{\mathrm{free}} + \frac{\partial}{\partial t}\rho_t\Big|_{\mathrm{weak}} \\ &= \mathrm{i}\omega[\rho_t, a^\dagger a] - \frac{1}{2}A(\nu+1)\left(a^\dagger a\rho_t - 2a\rho_t a^\dagger + \rho_t a^\dagger a\right) \\ &\qquad -\frac{1}{2}A\nu\left(aa^\dagger\rho_t - 2a^\dagger\rho_t a + \rho_t aa^\dagger\right) \\ &\equiv \mathcal{L}\rho_t\ , \end{aligned} \tag{2.29}$$

where the replacement $\rho_t^{(\mathrm{ph})} \to \rho_t$ is made to simplify the notation from here on. It should be clear that the $\mathcal{O}(\phi^4)$ terms of (2.25) and (2.26) are really negligible in the limiting situation of $r_1, r_2 \gg A$ and $\phi^2 \ll 1$ with finite values for the products $r_1\phi^2$ and $r_2\phi^2$.

Equation (2.29) is, of course, the master equation (2.1) that we had wished to derive by some physical arguments or, at least, make plausible. It is, by the way, essentially the only master equation that is consistent with some natural mathematical requirements of simplicity and symmetry (see, e.g., [16,17,18]).

From now on, we'll accept (2.29) as a candidate for describing a simple damped harmonic oscillator and study its implications. These implications as a whole are the ultimate justification for our conviction that very many crucial properties of damped oscillators are very well modeled by (2.29).

Before turning to these implications, however, we should not fail to mention the obvious. The Liouville operator $\mathcal{L}$ of (2.29) is a linear operator: the identities

$$\mathcal{L}(\lambda\rho) = \lambda\mathcal{L}\rho\ , \quad \mathcal{L}(\rho_1+\rho_2) = \mathcal{L}\rho_1 + \mathcal{L}\rho_2 \tag{2.30}$$

hold for all operators ρ, ρ_1, and ρ_2 and all numbers λ.

2.1.2 Some Simple Implications

As a basic check of consistency let us first make sure that (2.29) is not in conflict with the normalization of ρ_t to unit total probability, that is: $\mathrm{tr}\,\{\rho_t\} = 1$ for all t. Indeed, remembering the cyclic property of the trace, one easily verifies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{tr}\,\{\rho_t\} = \mathrm{tr}\left\{\frac{\partial}{\partial t}\rho_t\right\} = \mathrm{tr}\,\{\mathcal{L}\rho_t\} = 0\ , \tag{2.31}$$

as it should be. Much more difficult to answer is the question if (2.29) preserves the positivity of ρ_t; we will remark on that at the end of the third lecture (see Sect. 2.3.3 on p. 76).

Next, as a first application, we determine the time dependence of the expectation values of the ladder operator $a^\dagger$, a and the number operator $a^\dagger a$. Again, the cyclic property of the trace is the tool to be used, and we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big\langle a^\dagger\big\rangle_t &= \frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{tr}\left\{a^\dagger\rho_t\right\} = \mathrm{tr}\left\{a^\dagger\frac{\partial}{\partial t}\rho_t\right\} = (\mathrm{i}\omega - \tfrac{1}{2}A)\big\langle a^\dagger\big\rangle_t\ ,\\
\frac{\mathrm{d}}{\mathrm{d}t}\big\langle a\big\rangle_t &= (-\mathrm{i}\omega - \tfrac{1}{2}A)\big\langle a\big\rangle_t\ ,\\
\frac{\mathrm{d}}{\mathrm{d}t}\big\langle a^\dagger a\big\rangle_t &= -A\big(\big\langle a^\dagger a\big\rangle_t - \nu\big)\ ,
\end{aligned} \tag{2.32}$$

which are solved by

$$\begin{aligned}
\langle a^\dagger\rangle_t &= \langle a^\dagger\rangle_0\,\mathrm{e}^{-At/2}\,\mathrm{e}^{\mathrm{i}\omega t}\ ,\\
\langle a\rangle_t &= \langle a\rangle_0\,\mathrm{e}^{-At/2}\,\mathrm{e}^{-\mathrm{i}\omega t}\ ,\\
\langle a^\dagger a\rangle_t &= \nu + \big(\langle a^\dagger a\rangle_0 - \nu\big)\,\mathrm{e}^{-At}\ ,
\end{aligned} \tag{2.33}$$

respectively. Their long-time behavior,

$$t\to\infty:\quad \langle a^\dagger\rangle_t \to 0\ , \quad \langle a\rangle_t\to 0\ , \quad \langle a^\dagger a\rangle_t\to\nu\ , \tag{2.34}$$

seems to indicate that the evolution comes to a halt eventually.

2.1.3 Steady State

If this is indeed the case, then the master equation (2.29) must have a steady state $\rho^{(\mathrm{ss})}$. As we see in (2.32), it is impossible for ω and A to compensate for

each other and, therefore, $\rho^{(\mathrm{ss})}$ must commute with the number operator $a^\dagger a$, and as a consequence it must be a function of $a^\dagger a$ and cannot depend on $a^\dagger$ and a individually. Upon writing $f(a^\dagger a)$ for this function, we have

$$0 = \frac{\partial}{\partial t}\rho^{(\mathrm{ss})} = \mathcal{L}f(a^\dagger a) = -A(\nu+1)\left[a^\dagger a f(a^\dagger a) - a f(a^\dagger a)a^\dagger\right] - A\nu\left[a a^\dagger f(a^\dagger a) - a^\dagger f(a^\dagger a)a\right] , \tag{2.35}$$

and this implies the three-term recurrence relation

$$(a^\dagger a+1)\left[(\nu+1)f(a^\dagger a+1) - \nu f(a^\dagger a)\right] = a^\dagger a\left[(\nu+1)f(a^\dagger a) - \nu f(a^\dagger a-1)\right] , \tag{2.36}$$

which, incidentally, is a statement about detailed balance (see [19] for further details). In this equation, the left-hand side is obtained from the right-hand side by the replacement $a^\dagger a \to a^\dagger a+1$. Accordingly, the common value of both sides can be determined by evaluating the expression for any value that $a^\dagger a$ may have, that is: for any of its eigenvalues $(a^\dagger a)' = 0, 1, 2, \ldots$. We pick $(a^\dagger a)' = 0$ and find that either side of (2.36) must vanish. The resulting two-term recursion,

$$(\nu+1)f(a^\dagger a) = \nu f(a^\dagger a - 1) \tag{2.37}$$

is immediately solved by $f(a^\dagger a) = f(0)[\nu/(\nu+1)]^{a^\dagger a}$ and, after determining the value of $f(0)$ by normalization, we arrive at

$$\rho^{(\mathrm{ss})} = \frac{1}{\nu+1}\left(\frac{\nu}{\nu+1}\right)^{a^\dagger a} . \tag{2.38}$$

This steady state is in fact a thermal state, as we see after re-introducing the temperature Θ of (2.28),

$$\rho^{(\mathrm{ss})} = \left[1 - \exp\left(-\frac{\hbar\omega}{k_\mathrm{B}\Theta}\right)\right]\exp\left(-\frac{\hbar\omega}{k_\mathrm{B}\Theta}a^\dagger a\right) . \tag{2.39}$$

Indeed, together with (2.34) this tells us that, as stated at (2.1), "ν is the number of thermal excitations in the steady state". And the physical significance of A – it "is the energy decay rate" – is evident in (2.32). We might add that $\frac{1}{2}A$ is the decay rate of the oscillator's amplitude $\langle a\rangle$, which is proportional to the strength of the electromagnetic field in the optical model.

As the derivation shows, the steady state of (2.29) is unique, unless $A = 0$. Indeed, if $A = 0$ but $\omega \neq 0$, $\mathcal{L}\rho = 0$ is solved by all $\rho = f(a^\dagger a)$ irrespective of the actual form of the function f. We will take $A > 0$ for granted from here on.

2.1.4 Action to the Left

In (2.32) we obtained differential equations for expectation values from the equation of motion obeyed by the statistical operator, the master equation of (2.29).

This can be done systematically. We begin with the expectation value of some observable X and its time derivative,

$$\big\langle X\big\rangle_t = \mathrm{tr}\,\{X\rho_t\}\ , \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big\langle X\big\rangle_t = \mathrm{tr}\left\{X\frac{\partial}{\partial t}\rho_t\right\}\ , \tag{2.40}$$

and then use (2.29) to establish

$$\frac{\mathrm{d}}{\mathrm{d}t}\big\langle X\big\rangle_t = \mathrm{tr}\,\{X\mathcal{L}\rho_t\} = \big\langle X\mathcal{L}\big\rangle_t\ , \tag{2.41}$$

where the last equation defines the meaning of $X\mathcal{L}$, that is: the action of $\mathcal{L}$ to the left. The cyclic property of the trace is crucial once more in establishing the explicit expression

$$\begin{aligned} X\mathcal{L} = \mathrm{i}\omega[a^\dagger a, X] &- \frac{1}{2}A(\nu+1)\left(Xa^\dagger a - 2a^\dagger Xa + a^\dagger aX\right) \\ &- \frac{1}{2}A\nu\left(Xaa^\dagger - 2aXa^\dagger + aa^\dagger X\right)\ . \end{aligned} \tag{2.42}$$

When applied to $a^\dagger$, a, and $a^\dagger a$, this reproduces (2.32), of course.

How about (2.31)? It is also contained, namely as the statement

$$\mathrm{tr}\,\{\mathbb{1}\mathcal{L}\rho_t\} = 0 \quad \text{for all } \rho_t, \text{ or } \quad \mathbb{1}\mathcal{L} = 0\ , \tag{2.43}$$

which is a statement about the identity operator $\mathbb{1}$.

Homework Assignments

1 Take the explicit forms of $\mathrm{e}^{\mathrm{i}\phi\gamma}$ and $\mathrm{e}^{-\mathrm{i}\phi\gamma}$ in (2.16)–(2.18) and verify that

$$\mathrm{e}^{\mathrm{i}\phi\gamma}\,\mathrm{e}^{-\mathrm{i}\phi\gamma} = 1\ , \qquad \mathrm{e}^{-\mathrm{i}\phi\gamma}\,\mathrm{e}^{\mathrm{i}\phi\gamma} = 1\ . \tag{2.44}$$

2 According to (2.35), the steady state (2.38) is a right eigenvector of $\mathcal{L}$ with eigenvalue 0, $\mathcal{L}\rho^{(\mathrm{ss})} = 0$. What is the corresponding left eigenvector $\check{\rho}^{(\mathrm{ss})}$ such that $\check{\rho}^{(\mathrm{ss})}\mathcal{L} = 0$?

3 Reconsider (2.32) and (2.33). Show that these equations identify some other eigenvalues of $\mathcal{L}$ and their left eigenvectors.

4 Use the ansatz

$$\rho_t = \lambda(t)\big[1-\lambda(t)\big]^{a^\dagger a} \tag{2.45}$$

in the master equation (2.29), where $\lambda(t) \to \lambda = 1/(1+\nu)$ for $t \to \infty$. Derive a differential equation for the numerical function $\lambda(t)$, and solve it for arbitrary $\lambda(0)$. [If necessary, impose restrictions on $\lambda(0)$.] Then recognize that the solution reveals to you some eigenvalues of $\mathcal{L}$. Optional: Identify the corresponding right eigenvectors of $\mathcal{L}$.

2.2 Second Lecture: Eigenvalues and Eigenvectors of $\mathcal{L}$

2.2.1 A Simple Case First

When taking care of homework assignment 4, the reader used (2.35) to establish

$$\mathcal{L}\rho_t = -A\frac{1-(\nu+1)\lambda}{1-\lambda}\big[\lambda a^\dagger a-(1-\lambda)\big]\rho_t \tag{2.46}$$

for ρ_t of (2.45) and found

$$\frac{\partial}{\partial t}\rho_t = -\frac{1}{\lambda(1-\lambda)}\frac{\mathrm{d}\lambda}{\mathrm{d}t}\big[\lambda a^\dagger a-(1-\lambda)\big]\rho_t \tag{2.47}$$

by differentiation. Accordingly, (2.45) solves (2.29) if $\lambda(t)$ obeys

$$-\frac{1}{\lambda^2}\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{\lambda} = -A\left(\frac{1}{\lambda}-(\nu+1)\right)\ , \tag{2.48}$$

which is solved by

$$\lambda(t) = \frac{1}{(\nu+1)-\big[(\nu+1)-1/\lambda(0)\big]\,\mathrm{e}^{-At}} \tag{2.49}$$

where the restriction $\lambda(0)>0$ is sufficient to avoid ill-defined values of $\lambda(t)$ at later times, and $\lambda(0)\leq 1$ ensures a positive ρ_t throughout.

With (2.49) in (2.45) we have

$$\rho_t = \sum_{n=0}^{\infty}\mathrm{e}^{-nAt}\rho_n^{(0)}\ , \tag{2.50}$$

where the $\rho_n^{(0)}$'s are some functions of $a^\dagger a$. In particular, $\rho_0^{(0)}=\rho^{(\mathrm{ss})}$ is the steady state (2.38) that is reached for $t\to\infty$ when $\lambda(t)\to 1/(\nu+1)$. Since (2.50) is a solution of the master equation (2.29) by construction, it follows that

$$\sum_{n=0}^{\infty}\mathrm{e}^{-nAt}(-nA)\rho_n^{(0)} = \sum_{n=0}^{\infty}\mathrm{e}^{-nAt}\mathcal{L}\rho_n^{(0)} \tag{2.51}$$

holds for all $t>0$ and, therefore, $\rho_n^{(0)}$ is a right eigenvector of $\mathcal{L}$ with eigenvalue $-nA$,

$$\mathcal{L}\rho_n^{(0)} = -nA\rho_n^{(0)}\ . \tag{2.52}$$

As defined in (2.50) with ρ_t of (2.45) and $\lambda(t)$ of (2.49) on the left-hand side, the $\rho_n^{(0)}$'s depend on the particular value for $\lambda(0)$, a dependence of no relevance. We get rid of it by introducing a more appropriate expansion parameter x in accordance with

$$\frac{1}{\lambda(t)} = (\nu+1)(1+x) \quad\text{or}\quad x = \left[\frac{1}{(\nu+1)\lambda(0)}-1\right]\mathrm{e}^{-At}\ , \tag{2.53}$$

so that counting powers of e^{-At} is done by counting powers of x. Then

$$\frac{1}{(\nu+1)(1+x)}\left(1-\frac{1}{(\nu+1)(1+x)}\right)^{a^\dagger a}=\sum_{n=0}^{\infty}x^n\rho_n^{(0)} \tag{2.54}$$

is a generating function for the $\rho_n^{(0)}$'s with the spurious $\lambda(0)$ dependence removed.

The left-hand side of (2.54) can be expanded in powers of x for any value of $\nu\geq 0$, but we will be content with a look at the $\nu=0$ case and use a different method in Sect. 2.2.2 to handle the general situation. For $\nu=0$, the power series (2.54) is simplicity itself,

$$\begin{aligned}\nu=0:\quad \sum_{n=0}^{\infty}x^n\rho_n^{(0)}&=\frac{1}{1+x}\left(\frac{x}{1+x}\right)^{a^\dagger a}\\&=x^{a^\dagger a}\sum_{m=0}^{\infty}\binom{a^\dagger a+m}{m}(-x)^m\\&=\sum_{m=0}^{\infty}(-1)^m\binom{a^\dagger a+m}{a^\dagger a}x^{a^\dagger a+m}\,,\end{aligned} \tag{2.55}$$

so that

$$\nu=0:\quad \rho_n^{(0)}=(-1)^n-a^\dagger a\binom{n}{a^\dagger a}\,,\quad \mathcal{L}\rho_n^{(0)}=-nA\rho_n^{(0)}\,. \tag{2.56}$$

It is a matter of inspection to verify that

$$\begin{aligned}\rho_0^{(0)}&=\delta_{a^\dagger a,0}\,,\\ \rho_1^{(0)}&=\delta_{a^\dagger a,1}-\delta_{a^\dagger a,0}\,,\\ \rho_2^{(0)}&=\delta_{a^\dagger a,2}-2\delta_{a^\dagger a,1}+\delta_{a^\dagger a,0}\,,\\ &\vdots\end{aligned} \tag{2.57}$$

are the $\nu=0$ right eigenvectors of $\mathcal{L}$ to eigenvalues $0,-A,-2A,\ldots$.

We obtain the corresponding left eigenvectors from

$$\nu=0:\quad (1+y)^{a^\dagger a}=\sum_{m=0}^{\infty}y^m\check{\rho}_m^{(0)} \tag{2.58}$$

after verifying that this is a generating function indeed. For $\nu=0$, (2.42) says

$$\begin{aligned}(1+y)^{a^\dagger a}\mathcal{L}&=-Aa^\dagger a(1+y)^{a^\dagger a}+Aa^\dagger(1+y)^{a^\dagger a}a\\&=-Aya^\dagger a(1+y)^{a^\dagger a}-1\\&=-Ay\frac{\partial}{\partial y}(1+y)^{a^\dagger a}\end{aligned} \tag{2.59}$$

and the eigenvector equation

$$\check{\rho}_m^{(0)}\mathcal{L} = -mA\check{\rho}_m^{(0)} \tag{2.60}$$

gives

$$\sum_{m=0}^{\infty} y^m \check{\rho}_m^{(0)}\mathcal{L} = \sum_{m=0}^{\infty} y^m(-mA)\check{\rho}_m^{(0)} = -Ay\frac{\partial}{\partial y}\sum_{m=0}^{\infty} y^m\check{\rho}_m^{(0)} \ , \tag{2.61}$$

and now (2.59) and (2.61) establish (2.58). So we find

$$\nu = 0 : \quad \check{\rho}_m^{(0)} = \binom{a^\dagger a}{m} \ , \tag{2.62}$$

of which the first few are

$$\begin{aligned} \check{\rho}_0^{(0)} &= 1 \ , \\ \check{\rho}_1^{(0)} &= a^\dagger a \ , \\ \check{\rho}_2^{(0)} &= \frac{1}{2}a^\dagger a(a^\dagger a - 1) \ . \end{aligned} \tag{2.63}$$

For $n = 0$ and $n = 1$ this just repeats what was learned in (2.43) and (2.32) (recall homework assignments 2 and 3).

As dual eigenvector sets, the $\rho_n^{(0)}$'s and $\check{\rho}_m^{(0)}$'s must be orthogonal if $n \neq m$, which is here a statement about the trace of their product. It is simplest to deal with them as sets, and we use the generating functions to establish

$$\begin{aligned} \sum_{m,n=0}^{\infty} y^m \ \mathrm{tr}\left\{\check{\rho}_m^{(0)}\rho_n^{(0)}\right\}x^n &= \ \mathrm{tr}\left\{(1+y)^{a^\dagger a}\frac{1}{1+x}\left(\frac{x}{1+x}\right)^{a^\dagger a}\right\} \\ &= \frac{1}{1+x}\sum_{N=0}^{\infty}\left((1+y)\frac{x}{1+x}\right)^N \\ &= \frac{1}{1-xy} = \sum_{n=0}^{\infty} x^n y^n \\ &= \sum_{m,n=0}^{\infty} y^m\delta_{m,n}x^n \ , \end{aligned} \tag{2.64}$$

from which

$$\mathrm{tr}\left\{\check{\rho}_m^{(0)}\rho_n^{(0)}\right\} = \delta_{m,n} \tag{2.65}$$

follows immediately. This states the orthogonality of the $\nu = 0$ eigenvectors and also reveals the sense in which we have normalized them.

In the third lecture we will convince ourselves of the completeness of the eigenvector sets. Let us take this later insight for granted. Then we can write any given initial statistical operator $\rho_{t=0} = f(a^\dagger a)$ as a sum of the $\rho_n^{(0)}$,

$$\rho_{t=0} = \sum_{n=0}^{\infty}\alpha_n^{(0)}\rho_n^{(0)} \ , \tag{2.66}$$

and solve the master equation (2.29) by

$$\rho_t = \sum_{n=0}^{\infty} \alpha_n^{(0)} \, \mathrm{e}^{-nAt} \rho_n^{(0)} \; . \tag{2.67}$$

As a consequence of the orthogonality relation (2.65), we get the coefficients $\alpha_n^{(0)}$ as

$$\alpha_n^{(0)} = \mathrm{tr} \left\{ \check{\rho}_n^{(0)} \rho_{t=0} \right\} \; . \tag{2.68}$$

For $\nu = 0$, in particular, they are

$$\alpha_0^{(0)} = 1 \; , \quad \alpha_1^{(0)} = \left\langle a^\dagger a \right\rangle_{t=0} \; , \quad \alpha_2^{(0)} = \frac{1}{2} \left\langle a^\dagger a (a^\dagger a - 1) \right\rangle_{t=0} \; , \ldots \; , \tag{2.69}$$

which tells us that (2.66) and (2.67) are expansions in moments of the number operator. Put differently, the identity

$$f(a^\dagger a) = \sum_{n=0}^{\infty} \rho_n^{(0)} \; \mathrm{tr} \left\{ \check{\rho}_n^{(0)} f(a^\dagger a) \right\} \tag{2.70}$$

holds for any function $f(a^\dagger a)$ that has finite moments. For most of the others, one can exchange the roles of $\rho_n^{(0)}$ and $\check{\rho}_n^{(0)}$,

$$f(a^\dagger a) = \sum_{n=0}^{\infty} \mathrm{tr} \left\{ f(a^\dagger a) \rho_n^{(0)} \right\} \check{\rho}_n^{(0)} \; . \tag{2.71}$$

A useful rule of thumb is to employ expansion (2.70) for functions that have the basic characteristics of statistical operators [the traces in (2.70) are finite], and use (2.71) if $f(a^\dagger a)$ is of the kind that is typical for observables [such as $a^\dagger a$ for which the $n = 0$ trace in (2.70) is infinite, for example].

2.2.2 The General Case

Let us observe that some of the expressions in Sect. 2.2.1 are of a somewhat simpler structure when written in normally ordered form – all $a^\dagger$ operators to the left of all a operators – as exemplified by

$$\begin{aligned} \lambda(1-\lambda)^{a^\dagger a} &= \lambda \; : \mathrm{e}^{-\lambda a^\dagger a} : \, , \\ (1+y)^{a^\dagger a} &= \; : \mathrm{e}^{y a^\dagger a} : \, , \\ \nu = 0 \; : \quad \check{\rho}_m^{(0)} &= \binom{a^\dagger a}{m} = \frac{1}{m!} \; : (a^\dagger a)^m : = \frac{{a^\dagger}^m a^m}{m!} \, . \end{aligned} \tag{2.72}$$

We regard this as an invitation to generalize the ansatz (2.45) to

$$\rho_t = : \frac{1}{\kappa(t)} \, \mathrm{e}^{-\left[a^\dagger - \alpha^*(t)\right]\left[a - \alpha(t)\right]/\kappa(t)} : \tag{2.73}$$

where the switching from λ to $\kappa = 1/\lambda$ is strongly suggested by (2.48). Note that for the following it is not required that $\alpha^*(t)$ is the complex conjugate of $\alpha(t)$, it is more systematic to regard them as independent variables. We pay due attention to the ordering and obtain

$$\begin{aligned}\frac{\partial}{\partial t}\rho_t = &-\frac{1}{\kappa}\frac{\mathrm{d}\kappa}{\mathrm{d}t}\rho_t + \frac{1}{\kappa^2}\frac{\mathrm{d}\kappa}{\mathrm{d}t}(a^\dagger-\alpha^*)\rho_t(a-\alpha)\\ &+\frac{1}{\kappa}\frac{\mathrm{d}\alpha}{\mathrm{d}t}(a^\dagger-\alpha^*)\rho_t + \frac{1}{\kappa}\frac{\mathrm{d}\alpha^*}{\mathrm{d}t}\rho_t(a-\alpha)\end{aligned} \tag{2.74}$$

for the parametric time derivative of ρ_t. The evaluation of $\mathcal{L}\rho_t$ is equally straightforward once we note that $a^\dagger$'s on the right and a's on the left of ρ_t are moved to their natural side with the aid of these rules:

$$\begin{aligned}\rho_t a^\dagger &= a^\dagger\rho_t + [\rho_t, a^\dagger] = a^\dagger\rho_t + \frac{\partial}{\partial a}\rho_t = a^\dagger\rho_t - \frac{1}{\kappa}(a^\dagger-\alpha^*)\rho_t\ ,\\ a\rho_t &= \rho_t a + [a,\rho_t] = \rho_t a + \frac{\partial}{\partial a^\dagger}\rho_t = \rho_t a - \frac{1}{\kappa}\rho_t(a-\alpha)\end{aligned} \tag{2.75}$$

of which

$$[\rho_t, a^\dagger a] = -\frac{1}{\kappa}(a^\dagger-\alpha^*)\rho_t a + \frac{1}{\kappa}a^\dagger\rho_t(a-\alpha) \tag{2.76}$$

is an immediate application. Upon equating the numerical coefficients of both ρ_t and $(a^\dagger-\alpha^*)\rho_t(a-\alpha)$ we then get a single equation for $\kappa(t)$,

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t} = -A\big[\kappa-(\nu+1)\big]\ , \tag{2.77}$$

and the coefficients of $(a^\dagger-\alpha^*)\rho_t$ and $\rho_t(a-\alpha)$ supply equations for $\alpha(t)$ and $\alpha^*(t)$,

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \left(-\mathrm{i}\omega-\frac{1}{2}A\right)\alpha\ ,\quad \frac{\mathrm{d}\alpha^*}{\mathrm{d}t} = \left(\mathrm{i}\omega-\frac{1}{2}A\right)\alpha^*\ . \tag{2.78}$$

We have, in fact, met these differential equations before, namely (2.77) in (2.48) and (2.78) in (2.32). Their solutions

$$\begin{aligned}\kappa(t) &= \nu+1-(\nu+1-\kappa_0)\,\mathrm{e}^{-At}\ ,\\ \alpha(t) &= \alpha_0\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{-At/2}\ ,\\ \alpha^*(t) &= \alpha_0^*\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{e}^{-At/2}\ ,\end{aligned} \tag{2.79}$$

where κ_0, α_0, α_0^* are the arbitrary initial values at $t=0$ [not to be confused with the time-independent coefficients $\alpha_n^{(0)}$ of (2.66)–(2.69)], tell us that the time dependence of

$$\rho_t = \mathrm{e}^{\mathcal{L}t}\rho_{t=0} \tag{2.80}$$

contains powers of e^{-At} combined with powers of $\mathrm{e}^{\mp\mathrm{i}\omega t - At/2}$. Therefore, the values

$$\lambda_n^{(k)} = -\mathrm{i}k\omega - \big(n+\tfrac{1}{2}|k|\big)A \quad \text{with } n=0,1,2,\ldots \text{ and } k=0,\pm1,\pm2,\ldots \tag{2.81}$$

must be among the eigenvalues of $\mathcal{L}$. In fact, these are all eigenvalues, and the expansion of the generating function (2.73)

$$\rho_t = \sum_{n=0}^{\infty} \sum_{k=-\infty}^{\infty} b_n^{(k)} \, \mathrm{e}^{-\mathrm{i}k\omega t - \left(n + \frac{1}{2}|k|\right)At} \rho_n^{(k)} \tag{2.82}$$

yields all right eigenvectors of $\mathcal{L}$,

$$\mathcal{L}\rho_n^{(k)} = \left[-\mathrm{i}k\omega - \left(n + \tfrac{1}{2}|k|\right)A\right] \rho_n^{(k)} \, . \tag{2.83}$$

We will justify the assertion that these are *all* in the third lecture. Right now we just report the explicit expressions that one obtains for the eigenvectors $\rho_n^{(k)}$ and the coefficients $b_n^{(k)}$. They are [20,21]

$$\rho_n^{(k)} = \frac{(-1)^n}{(\nu+1)^{|k|+1}} a^{\dagger \, \frac{1}{2}(|k|+k)} \; : \mathrm{L}_n^{(|k|)}\left(\frac{a^\dagger a}{\nu+1}\right) \mathrm{e}^{-\frac{a^\dagger a}{\nu+1}} : \; a^{\frac{1}{2}(|k|-k)} \tag{2.84}$$

and

$$b_n^{(k)} = \frac{n!}{(n+|k|)!} \left(\frac{\kappa_0}{\nu+1} - 1\right)^n \alpha_0{}^{\frac{1}{2}(|k|+k)} \mathrm{L}_n^{(|k|)}\left(\frac{\alpha_0^* \alpha_0}{\nu+1-\kappa_0}\right) \alpha_0^{*\,\frac{1}{2}(|k|-k)} \, , \tag{2.85}$$

where the $\mathrm{L}_n^{(|k|)}$'s are Laguerre polynomials. Note that all memory about the initial values of κ, α, and α^* is stored in $b_n^{(k)}$, as it should be. The right eigenvectors $\rho_n^{(k)}$ of (2.84) and the left eigenvectors $\check{\rho}_n^{(k)}$ of (2.87) constitute the two so-called *damping bases* [20] associated with the Liouville operator $\mathcal{L}$ of (2.29).

Homework Assignments

5 Verify that (2.84) reproduces (2.56) for $k = 0$ and $\nu = 0$.

6 Show that (2.82) with (2.84) and (2.85) is correct. What is $b_n^{(k)}$ for $\kappa_0 = \nu+1$, $\alpha_0 = (\nu+1)\beta$, and $\alpha_0^* = (\nu+1)\beta^*$? [You need some familiarity with Bessel functions, Laguerre polynomials, and relations between them; consult the Appendix if necessary.]

7 Consider

$$U = \mathrm{e}^{\alpha a^\dagger} \mathrm{e}^{\alpha^* a} \mathrm{e}^{\lambda} = \mathrm{e}^{\alpha^* a} \mathrm{e}^{\alpha a^\dagger} \mathrm{e}^{\lambda - \alpha^*\alpha} = \mathrm{e}^{\alpha a^\dagger + \alpha^* a} \mathrm{e}^{\lambda - \frac{1}{2}\alpha^*\alpha} \tag{2.86}$$

with $\alpha^*(t)$, $\alpha(t)$, $\lambda(t)$ such that $\partial U/\partial t = U\mathcal{L}$. Find the differential equations obeyed by α^*, α, λ and solve them.

8 Show that this also establishes the eigenvalues (2.81) of $\mathcal{L}$.

9 Extract the left eigenvectors $\check{\rho}_n^{(k)}$ of $\mathcal{L}$. Normalize them such that

$$\check{\rho}_n^{(k)} = \left(\frac{-\nu}{1+\nu}\right)^n \frac{n!}{(n+|k|)!} a^{\dagger \, \frac{1}{2}(|k|-k)} \; : \mathrm{L}_n^{(|k|)}\left(\frac{a^\dagger a}{\nu}\right) : \; a^{\frac{1}{2}(|k|+k)} \, . \tag{2.87}$$

What do you get in the limit $\nu \to 0$?

10 State $\check{\rho}_n^{(k)}$ explicitly for $k = 0$, $n = 0, 1, 2$ and for $k = \pm 1, \pm 2$, $n = 0$. Compare with (2.63).

11 Use the two generating functions (2.73) and (2.86) to demonstrate

$$\mathrm{tr}\left\{\check{\rho}_m^{(k)}\rho_n^{(k')}\right\} = \delta_{m,n}\delta_{k,k'} \ . \tag{2.88}$$

2.3 Third Lecture: Completeness of the Damping Bases

2.3.1 Phase Space Functions

As a preparation we first consider the standard phase space functions $f(Q', P')$ associated with an operator $F(Q, P)$ where Q, P are a Heisenberg pair (colloquially: position Q and momentum P). This is to say that they obey the commutation relation

$$[Q, P] = \mathrm{i} \tag{2.89}$$

and have complete, orthonormal sets of eigenkets and eigenbras,

$$\begin{aligned} Q|Q'\rangle &= |Q'\rangle Q' \ , & P|P'\rangle &= |P'\rangle P' \ , \\ \langle Q'|Q &= Q'\langle Q'| \ , & \langle P'|P &= P'\langle P'| \ , \\ \langle Q'|Q''\rangle &= \delta(Q'-Q'') \ , & \langle P'|P''\rangle &= \delta(P'-P'') \ , \\ |Q'\rangle\langle Q'| &= \delta(Q-Q') \ , & |P'\rangle\langle P'| &= \delta(P-P') \ , \\ \int \mathrm{d}Q' \, |Q'\rangle\langle Q'| &= 1 \ , & \int \mathrm{d}P' \, |P'\rangle\langle P'| &= 1 \ , \end{aligned} \tag{2.90}$$

where Q', Q'' and P', P'' denote eigenvalues. The familiar plane waves

$$\langle Q'|P'\rangle = \frac{\mathrm{e}^{\mathrm{i}Q'P'}}{\sqrt{2\pi}} \tag{2.91}$$

relate the eigenvector sets to each other.

By using both completeness relations, we can put any $F = F(Q, P)$ into its Q, P-ordered form,

$$\begin{aligned} F(Q,P) &= \int \mathrm{d}Q' \, \mathrm{d}P' \, |Q'\rangle\langle Q'|F|P'\rangle\langle P'| \\ &= \int \mathrm{d}Q' \, \mathrm{d}P' \, |Q'\rangle\langle Q'| \left[\frac{\langle Q'|F|P'\rangle}{\langle Q'|P'\rangle}\right] |P'\rangle\langle P'| \\ &= \int \mathrm{d}Q' \, \mathrm{d}P' \, \delta(Q-Q') f(Q',P') \delta(P-P') \\ &= f(Q,P)\Big|_{Q,P-\mathrm{ordered}} \equiv f(Q;P) \ , \end{aligned} \tag{2.92}$$

where the last step defines the meaning of the semicolon in $f(Q; P)$. Thus, the procedure is this: evaluate the normalized matrix element

$$f(Q', P') = \frac{\langle Q'|F|P'\rangle}{\langle Q'|P'\rangle} = \mathrm{tr}\left\{F\frac{|P'\rangle\langle Q'|}{\langle Q'|P'\rangle}\right\} \ , \tag{2.93}$$

then replace $Q' \to Q$, $P' \to P$ with due attention to their order in products – all Q's must stand to the left of all P's – and so obtain $F = f(Q;P)$, the Q,P-ordered form of F. The P,Q-ordered version of F is found by an analogous procedure with the roles of Q and P interchanged.

The fraction that appears in the trace formula of (2.93) is equal to its square,

$$\left[\frac{|P'\rangle\langle Q'|}{\langle Q'|P'\rangle}\right]^2 = \frac{|P'\rangle\langle Q'|}{\langle Q'|P'\rangle} \tag{2.94}$$

but, not being Hermitian, it is not a projector. It has, however, much in common with projectors, and this is emphasized by using

$$\frac{1}{\langle Q'|P'\rangle} = 2\pi\langle P'|Q'\rangle \tag{2.95}$$

to turn it into

$$\frac{|P'\rangle\langle Q'|}{\langle Q'|P'\rangle} = 2\pi|P'\rangle\langle P'|Q'\rangle\langle Q'| = 2\pi\delta(P-P')\delta(Q-Q')\ , \tag{2.96}$$

which is essentially the product of two projectors. Then,

$$f(Q',P') = \text{ tr }\{F\,2\pi\delta(P-P')\delta(Q-Q')\} \tag{2.97}$$

is yet another way of presenting $f(Q',P')$.

In (2.92) and (2.97) we recognize two basis sets of operators,

$$\begin{aligned} B(Q',P') &= 2\pi\delta(Q-Q')\delta(P-P')\ , \\ \widetilde{B}(Q',P') &= 2\pi\delta(P-P')\delta(Q-Q')\ , \end{aligned} \tag{2.98}$$

labeled by the phase space variables Q' and P'. Their dual roles are exhibited by the compositions of (2.92) and (2.97),

$$F(Q,P) = \int \frac{\mathrm{d}Q'\,\mathrm{d}P'}{2\pi}\, B(Q',P')\ \text{tr}\left\{F\,\widetilde{B}(Q',P')\right\} \tag{2.99}$$

and

$$f(Q',P') = \text{ tr}\left\{\widetilde{B}(Q',P')\int \frac{\mathrm{d}Q''\,\mathrm{d}P''}{2\pi}\, f(Q'',P'')B(Q'',P'')\right\}\ , \tag{2.100}$$

and therefore

$$\text{tr}\, B(Q',P')\widetilde{B}(Q'',P'') = 2\pi\delta(Q'-Q'')\delta(P'-P'') \tag{2.101}$$

states both their orthogonality and their completeness.

We note that the displacements

$$Q \to Q-Q'\ , \qquad P \to P-P' \tag{2.102}$$

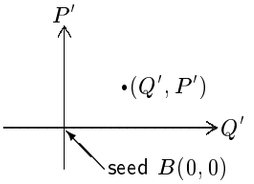


Fig. 2.3. The unitary displacements $Q \to Q - Q'$, $P \to P - P'$ map out all of phase space, turning the basis seed $B(0,0)$ into all other operators $B(Q', P')$ of the basis

that map out all of phase space (see Fig. 2.3) are unitary operations, and so is the interchange

$$Q \to P \;, \quad P \to -Q \;. \tag{2.103}$$

As a consequence, $B(Q', P')$ and $\widetilde{B}(Q', P')$ are unitarily equivalent to their respective basis seeds

$$B(0,0) = 2\pi\delta(Q)\delta(P) \quad \text{and} \quad \widetilde{B}(0,0) = 2\pi\delta(P)\delta(Q) \;, \tag{2.104}$$

and these seeds themselves are unitarily equivalent and are also adjoints of each other,

$$B(0,0)^\dagger = \widetilde{B}(0,0) \;, \quad \widetilde{B}(0,0)^\dagger = B(0,0) \;, \tag{2.105}$$

Here, $B(0,0)$ is stated as a Q,P-ordered operator and $\widetilde{B}(0,0)$ is P,Q-ordered. The reverse orderings are also available, as we illustrate for $B(0,0)$:

$$\frac{\langle P'|B(0,0)|Q'\rangle}{\langle P'|Q'\rangle} = 2\pi \frac{\langle P'|Q''\rangle\langle Q''|P''\rangle\langle P''|Q'\rangle}{\langle P'|Q'\rangle}\bigg|_{Q''=0,\ P''=0} = \mathrm{e}^{\mathrm{i}P'Q'} \;, \tag{2.106}$$

giving

$$B(0,0) = \mathrm{e}^{\mathrm{i}P;\,Q} = \sum_{k=0}^{\infty} \frac{\mathrm{i}^k}{k!} P^k Q^k \;, \tag{2.107}$$

where we meet a typical ordered exponential operator function. We get the Q,P-ordered version of $\widetilde{B}(0,0)$,

$$\widetilde{B}(0,0) = \mathrm{e}^{-\mathrm{i}Q;\,P} \;, \tag{2.108}$$

by taking the adjoint.

These observations can be generalized in a simple and straightforward manner [22]. The operators

$$\lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q} = \bar\lambda\,\mathrm{e}^{-\mathrm{i}\bar\lambda Q;\,P} \;, \quad \lambda\,\mathrm{e}^{-\mathrm{i}\lambda Q;\,P} = \bar\lambda\,\mathrm{e}^{\mathrm{i}\bar\lambda P;\,Q}$$
$$\text{with } \lambda, \bar\lambda \geq 1 \text{ and } \lambda\bar\lambda = \lambda + \bar\lambda \tag{2.109}$$

are seeds of a dual pair of bases for each choice of $\lambda, \bar\lambda$ because the orthogonality-completeness relation

$$\mathrm{tr}\left\{\lambda\,\mathrm{e}^{\mathrm{i}\lambda(P-P');\,(Q-Q')}\lambda\,\mathrm{e}^{-\mathrm{i}\lambda(Q-Q'');\,(P-P'')}\right\} = 2\pi\delta(Q'-Q'')\delta(P'-P'') \tag{2.110}$$

holds generally, not just for $\lambda = 1, \bar{\lambda} = \infty$ as we've seen in (2.101). We demonstrate the case by using the two completeness relations of (2.90) to turn the operators into numbers, and then recognize two Fourier representation of Dirac's δ function:

$$\begin{aligned}
\mathrm{tr}\left\{\cdots\right\} &= \int \frac{\mathrm{d}\bar{Q}\,\mathrm{d}\bar{P}}{2\pi} \lambda\, \mathrm{e}^{\mathrm{i}\lambda(\bar{P}-P')(\bar{Q}-Q')}{}_{\lambda}\, \mathrm{e}^{-\mathrm{i}\lambda(\bar{Q}-Q'')(\bar{P}-P'')} \\
&= 2\pi\, \mathrm{e}^{\mathrm{i}\lambda(Q'P'-Q''P'')} \int \frac{\mathrm{d}\bar{P}}{2\pi} \lambda\, \mathrm{e}^{-\mathrm{i}\lambda\bar{P}(Q'-Q'')} \int \frac{\mathrm{d}\bar{Q}}{2\pi} \lambda\, \mathrm{e}^{-\mathrm{i}\lambda\bar{Q}(P'-P'')} \\
&= 2\pi\, \mathrm{e}^{\mathrm{i}\lambda(Q'P'-Q''P'')} \delta(Q'-Q'')\delta(P'-P'') \\
&= 2\pi\delta(Q'-Q'')\delta(P'-P'')\;, \qquad (2.111)
\end{aligned}$$

indeed. The equivalence of the λ and $\bar{\lambda}$ versions of (2.109) is the subject matter of homework assignment 12.

The basis seeds (2.109) can be characterized by the similarity transformations they generate,

$$\begin{aligned}
Q\,\lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q} &= \lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q}\,(1-\lambda)Q \quad \text{or} \quad Q \to (1-\lambda)Q\;, \\
P\,\lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q} &= \lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q}\,(1-\bar{\lambda})P \quad \text{or} \quad P \to (1-\bar{\lambda})P\;, \qquad (2.112)
\end{aligned}$$

which are scaling transformations essentially (but not quite because $1-\lambda = 1/(1-\bar{\lambda}) < 0$ and the cases $\lambda = 1$ or $\bar{\lambda} = 1$ are particular). One verifies (2.112) with the aid of identities such as

$$P\,\lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q} = \frac{1}{\mathrm{i}}\frac{\partial}{\partial Q}\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q} = \left[P,\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q}\right]\;. \qquad (2.113)$$

Equations (2.112) by themselves determine the seed only up to an over-all factor, and this ambiguity is removed by imposing the normalization to unit trace,

$$\mathrm{tr}\left\{\lambda\,\mathrm{e}^{\mathrm{i}\lambda P;\,Q}\right\} = \int \frac{\mathrm{d}\bar{Q}\,\mathrm{d}\bar{P}}{2\pi}\lambda\,\mathrm{e}^{\mathrm{i}\lambda\bar{P}\bar{Q}} = 1\;. \qquad (2.114)$$

The cases $\lambda = 1, \bar{\lambda} \to \infty$ and $\lambda \to \infty, \bar{\lambda} = 1$ are just the bases associated with the Q,P-ordered and P,Q-ordered phase space functions that we discussed above. Of particular interest is also the symmetric case of $\lambda = \bar{\lambda} = 2$ which has the unique property that the two bases are really just one: The operator basis underlying Wigner's phase space function. Its seed (not seeds!) is Hermitian, since $\lambda = \bar{\lambda} = 2$ in (2.109) implies

$$2\,\mathrm{e}^{\mathrm{i}2P;\,Q} = 2\,\mathrm{e}^{-\mathrm{i}2Q;\,P} = \left[2\,\mathrm{e}^{\mathrm{i}2P;\,Q}\right]^{\dagger}\;. \qquad (2.115)$$

It then follows, for example, that $F = F^{\dagger}$ has a real Wigner function; in fact, all of the well known properties of the much studied Wigner functions can be derived rather directly from the properties of this seed (see [22] for details).

For our immediate purpose we just need to know the following. When $\lambda = \bar{\lambda} = 2$, the transformation (2.112) is the inversion

$$Q \to -Q\,, \qquad P \to -P\,. \tag{2.116}$$

For $a^\dagger = 2^{-1/2}(Q - \mathrm{i}P)$, $a = 2^{-1/2}(Q + \mathrm{i}P)$ this means that the number operator $a^\dagger a = \frac{1}{2}(Q^2 + P^2 - 1)$ is invariant. Put differently, the Wigner seed (2.115) commutes with $a^\dagger a$, it is a function of $a^\dagger a$: $2\,\mathrm{e}^{\mathrm{i}2P;\,Q} = f(a^\dagger a)$. For a, the inversion (2.116) requires $af(a^\dagger a) = -f(a^\dagger a)a$, and this combines with $af(a^\dagger a) = f(a^\dagger a + 1)a$ to tell us that $f(a^\dagger a + 1) = -f(a^\dagger a)$. We conclude that

$$2\,\mathrm{e}^{\mathrm{i}2P;\,Q} = 2(-1)^{a^\dagger a} =: 2\,\mathrm{e}^{-2a^\dagger a}\,: \tag{2.117}$$

after using the normalization (2.114) to determine the prefactor of 2,

$$\mathrm{tr}\left\{2(-1)^{a^\dagger a}\right\} = 2\sum_{n=0}^{\infty}(-1)^n = 2\sum_{n=0}^{\infty}(-x)^n\Big|_{1>x\to 1} = \frac{2}{1+x}\Big|_{x\to 1} = 1\,. \tag{2.118}$$

The last, normally ordered, version in (2.117) of the Wigner seed shows the $\lambda = 2$ case of (2.72).

Perhaps the first to note the intimate connection between the Wigner function and the inversion (2.116) and thus to recognize that the Wigner seed is (twice) the parity operator $(-1)^{a^\dagger a}$ was Royer [23]. In the equivalent language of the Weyl quantization scheme the analogous observation was made a bit earlier by Grossmann [24]. A systematic study from the viewpoint of operator bases is given in [22].

We use the latter form (2.117) to write an operator $F(a^\dagger, a)$ in terms of its Wigner function $f(z^*, z)$,

$$F(a^\dagger, a) = \int \frac{\mathrm{d}Q'\,\mathrm{d}P'}{2\pi} f(z^*, z)\ :2\,\mathrm{e}^{-2\left(a^\dagger - z^*\right)\left(a - z\right)}:\,, \tag{2.119}$$

where $z^* = 2^{-1/2}(Q' - \mathrm{i}P')$ and $z = 2^{-1/2}(Q' + \mathrm{i}P')$ are understood, and

$$f(z^*, z) = \mathrm{tr}\left\{F(a^\dagger, a)\ :2\,\mathrm{e}^{-2\left(a^\dagger - z^*\right)\left(a - z\right)}:\right\} \tag{2.120}$$

reminds us of how we get the phase space function by tracing the products with the operators of the dual basis (which, we repeat, is identical to the expansion basis in the $\lambda = \bar{\lambda} = 2$ case of the Wigner basis).

2.3.2 Completeness of the Eigenvectors of $\mathcal{L}$

The stage is now set for a demonstration of the completeness of the damping bases of Sect. 2.2.2. We will deal explicitly with the right eigenvectors $\rho_n^{(k)}$ of

the Liouville operator $\mathcal{L}$ of (2.29) as obtained in (2.84) by expanding the generating function (2.73). Consider some arbitrary initial state $\rho_{t=0}$ and its Wigner function representation

$$\rho_{t=0} = \int \frac{\mathrm{d}Q'\,\mathrm{d}P'}{2\pi} \rho(z^*, z) \;:2\,\mathrm{e}^{-2\left(a^\dagger - z^*\right)\left(a - z\right)}:\,, \tag{2.121}$$

where $\rho(z^*, z)$ is the Wigner phase space function of $\rho_{t=0}$. Then, according to Sect. 2.2.2, we have at any later time

$$\rho_t = \mathrm{e}^{\mathcal{L}t}\rho_{t=0} = \int \frac{\mathrm{d}Q'\,\mathrm{d}P'}{2\pi} \rho(z^*, z) \;:\frac{1}{\kappa(t)}\,\mathrm{e}^{-\left[a^\dagger - \alpha^*(t)\right]\left[a - \alpha(t)\right]/\kappa(t)}: \tag{2.122}$$

with

$$\kappa_0 = \tfrac{1}{2}\,, \quad \alpha_0 = z\,, \quad \alpha_0^* = z^* \tag{2.123}$$

in (2.79). In conjunction with (2.82), (2.84), and (2.85) this gives

$$\rho_t = \sum_{n=0}^{\infty}\sum_{k=-\infty}^{\infty} \beta_n^{(k)}\,\mathrm{e}^{-\mathrm{i}k\omega t - \left(n + \frac{1}{2}|k|\right)At}\rho_n^{(k)} \tag{2.124}$$

with

$$\begin{aligned}\beta_n^{(k)} &= \frac{n!}{(n+|k|)!}\left(-\frac{\nu + \frac{1}{2}}{\nu + 1}\right)^n \\ &\quad\times \int \frac{\mathrm{d}Q'\,\mathrm{d}P'}{2\pi} \rho(z^*, z) z^{\frac{1}{2}(|k|+k)} \mathrm{L}_n^{(|k|)}\left(\frac{z^* z}{\nu + \frac{1}{2}}\right) z^{*\frac{1}{2}(|k|-k)}.\end{aligned} \tag{2.125}$$

But this is just to say that any given ρ_t has an expansion in terms of the $\rho_n^{(k)}$ for all t – any ρ_t can be expanded in the right damping basis. Equation (2.124) also confirms that the exponentials $\exp\left(-\mathrm{i}k\omega t - \left(n + \frac{1}{2}|k|\right)At\right) = \exp\left(\lambda_n^{(k)}t\right)$ constitute all possible time dependences that ρ_t might have. Clearly, then, the $\lambda_n^{(k)}$'s of (2.81) are *all* eigenvalues of $\mathcal{L}$, indeed, and the right eigenvectors $\rho_n^{(k)}$ of (2.84) are complete. The completeness of the left eigenvectors $\check{\rho}_n^{(k)}$ of (2.87) can be shown similarly, or can be inferred from (2.88).

Note that this argument does not use any of the particular properties that $\rho_{t=0}$ might have as a statistical operator. All that is really required is that its Wigner function $\rho(z^*, z)$ exists, and this is almost no requirement at all, because only operators that are singularly pathological may not possess a Wigner function. Such exceptions are of no interest to the physicist.

More critical are those operators $\rho_{t=0}$ for which the expansion is of a more formal character because the resulting coefficients $\beta_n^{(k)}$ are distributions, rather than numerical functions, of the parameters that are implicit in $\rho(z^*, z)$. In this situation the recommended procedure is to expand in terms of the left eigenvectors $\check{\rho}_n^{(k)}$ instead of the right eigenvectors $\rho_n^{(k)}$.

The demonstration of completeness given here relies on machinery developed in [20,21,22]. An alternative approach can be found in [25], where the case $\omega = 0$, $\nu = 0$ is treated, but it should be possible to use the method for the general case as well.

Note also that the time dependence in (2.122) is solely carried by the operator basis, not by the phase space function. This is reminiscent of – and in fact closely related to – the interaction-picture formalism of unitary quantum evolutions. Owing to the non-unitary terms in $\mathcal{L}$, those proportional to the decay rate A, the evolution of ρ_t is not unitary in (2.122). Of course, one could also have a description in which the operator basis does not change in time, but the phase space function does. It then obeys a partial differential equation of the Fokker–Planck type. Concerning these matters, the reader should consult the standard quantum optics literature as well as special focus books such as [26].

2.3.3 Positivity Conservation

Let us now return to the question that we left in limbo in Sect. 2.1.2: Does the master equation (2.29) preserve the positivity of ρ_t?

Suppose that $\rho_{t=0}$ is not a general operator but really a statistical operator. Then the coefficients $\beta_n^{(k)}$ in (2.124) are such that the right-hand side is non-negative for $t = 0$ and has unit trace. Since $\rho_0^{(0)}$ is the steady state $\rho^{(\mathrm{ss})}$ of (2.38) and $\check{\rho}_0^{(0)} = \mathbb{1}$ is the identity, we have

$$\mathrm{tr}\left\{\rho_n^{(k)}\right\} = \mathrm{tr}\left\{\check{\rho}_0^{(0)}\rho_n^{(k)}\right\} = \delta_{n,0}\delta_{k,0} \tag{2.126}$$

as a consequence of (2.88), and

$$\beta_n^{(k)} = \mathrm{tr}\left\{\check{\rho}_n^{(k)}\rho_{t=0}\right\} \tag{2.127}$$

implies $\beta_0^{(0)} = \mathrm{tr}\,\{\rho_{t=0}\} = 1$. So, all terms in (2.124) are traceless with the sole exception of the time independent $n = 0, k = 0$ term, which has unit trace. This demonstrates once more that the trace of ρ_t is conserved, as we noted in Sect. 2.1.2 already.

The time independent $n = 0, k = 0$ term is also the only one in (2.124) that is non-negative by itself. By contrast, the expectation values of $\rho_n^{(k)}$ are both positive and negative if $n > 0, k = 0$ and even complex if $k \neq 0$. Now, since $\rho_{t=0} \geq 0$, the $n = 0, k = 0$ term clearly dominates all others for $t = 0$, and then it dominates them even more for $t > 0$ because the weight of $\rho_0^{(0)} = \rho^{(\mathrm{ss})}$ is constant in time while the other $\rho_n^{(k)}$'s have weight factors that decrease exponentially with t. We can therefore safely infer that the master equation (2.29) conserves the positivity of ρ_t.

2.3.4 Lindblad Form of Liouville Operators

The Liouville operator of (2.29) can be rewritten as

$$\begin{aligned}\mathcal{L}\rho = \mathrm{i}\omega[\rho, a^\dagger a] + \frac{1}{2}A(\nu+1)\left([a,\rho a^\dagger] + [a\rho, a^\dagger]\right) \\ + \frac{1}{2}A\nu\left([a^\dagger,\rho a] + [a^\dagger\rho, a]\right)\end{aligned} \tag{2.128}$$

and then it is an example of the so-called Lindblad form of Liouville operators. The general Lindblad form is

$$\mathcal{L}\rho = \frac{\mathrm{i}}{\hbar}[\rho, H] + \sum_j \Big([V_j^\dagger, \rho V_j] + [V_j^\dagger\rho, V_j]\Big) \tag{2.129}$$

where $H = H^\dagger$ is the Hermitian Hamilton operator that generates the unitary part of the evolution. Clearly, any $\mathcal{L}$ of this form will conserve the trace, but the requirement of trace conservation would also be met if the sum were subtracted in (2.129) rather than added. As Lindblad showed [27], however, this option is only apparent: all terms of the form $[V_j^\dagger, \rho V_j] + [V_j^\dagger\rho, V_j]$ in a Liouville operator must come with a positive weight. He further demonstrated that all $\mathcal{L}$'s of the form (2.129) surely conserve the positivity of ρ provided that all the V_j's are bounded. This fact has become known as the *Lindblad theorem*. (Theorem 12 in Sect. 4.9 is a mathematically precise statement of the theorem for action to the left.) In the case that some V_j is not bounded, positivity may be conserved or not. A proof of the Lindblad theorem is far beyond the scope of these lectures. Reference [28] is perhaps a good starting point for the reader who wishes to learn more about these matters, and [29] is a recent monograph that deals with various mathematical aspects. We must in fact recognize that (2.128) is obtained with

$$V_1 = \sqrt{\tfrac{1}{2}A(\nu+1)}\,a^\dagger\,, \quad V_2 = \sqrt{\tfrac{1}{2}A\nu}\,a \tag{2.130}$$

in (2.129) and these are actually *not* bounded so that the Lindblad theorem does not apply. Fortunately, we have other arguments at hand, namely the ones of Sect. 2.3.3, to show that the master equation (2.29) conserves positivity.

But while we are at it, let's just give a little demonstration of what goes wrong when a master equation is not of the Lindblad form. Take

$$\frac{\partial}{\partial t}\rho_t = VV^\dagger\rho_t - 2V^\dagger\rho_t V + \rho_t VV^\dagger\,, \tag{2.131}$$

for example, and assume that the initial state is pure, $\rho_{t=0} = |\psi_0\rangle\langle\psi_0|$. Then we have at $t = \mathrm{d}t > 0$

$$\rho_{t=\mathrm{d}t} = \rho_{t=0} + \mathrm{d}t\frac{\partial\rho_t}{\partial t}\bigg|_{t=0} = |\psi_0\rangle\langle\psi_0| + \mathrm{d}t\big(|\psi_2\rangle\langle\psi_0| - 2|\psi_1\rangle\langle\psi_1| + |\psi_0\rangle\langle\psi_2|\big)\,, \tag{2.132}$$

where $|\psi_1\rangle = V^\dagger|\psi_0\rangle$ and $|\psi_2\rangle = VV^\dagger|\psi_0\rangle$. Choose $|\psi_0\rangle$ such that $\langle\psi_1|\psi_0\rangle = 0$ and $\langle\psi_1|\psi_2\rangle = 0$, and calculate the probability for $|\psi_1\rangle$ at time $t = \mathrm{d}t$,

$$\frac{\langle\psi_1|\rho_{t=\mathrm{d}t}|\psi_1\rangle}{\langle\psi_1|\psi_1\rangle} = -2\mathrm{d}t < 0\,! \tag{2.133}$$

Positivity is violated already after an infinitesimal time step and, therefore, (2.131) is not really a master equation.

Homework Assignments

12 Show the equivalence of the λ and $\bar{\lambda}$ versions of (2.109).
13 If $F = f_1(Q;P) = f_2(P;Q)$, what is the relation between $f_1(Q',P')$ and $f_2(P',Q')$?
14 Consider $\Gamma = \frac{1}{2}(QP+PQ)$, the generator of scaling transformations. For real numbers μ, write $\mathrm{e}^{\mathrm{i}\mu\Gamma}$ in Q,P-ordered and P,Q-ordered form.
15 Find the Wigner function of the number operator $a^\dagger a$.
16 Construct an explicit (and simple!) example for (2.131)–(2.133), that is: specify V and $|\psi_0\rangle$.

2.4 Fourth Lecture: Quantum-Optical Applications

Single atoms are routinely passed through high-quality resonators in experiments performed in Garching and Paris, much like it is depicted in Fig. 2.1; see the review articles [30,**?**,**?**,**?**,34] and the references cited in them. In a set-up typical of the Garching experiments, the atoms deposit energy into the resonator and so compensate for the losses that result from dissipation. The intervals between the atoms are usually so large that an atom is long gone before the next one comes, so that at any time at most one atom is inside the cavity, all other cases being extremely rare. And, therefore, the fitting name "one-atom maser" or "micromaser" has been coined for this system.

The properties of the steady state of the radiation field that is established in the one-atom maser are determined by the values of several parameters of which the photon decay rate A, the thermal photon number ν, and the atomic arrival rate r are the most important ones. Rare exceptions aside, an atom is entangled with the photon mode after emerging from the cavity and it becomes entangled with the next atom after that has traversed the resonator. As a consequence, measurements on the exiting atoms reveal intriguing correlations which are the primary source of information of the photon state inside the cavity. A wealth of phenomena has been studied in these experiments over the last 10–15 years, and we look forward to seeing many more exciting results in the future.

2.4.1 Periodically Driven Damped Oscillator

To get an idea of the theoretical description of such experiments and to show a simple, yet typical, application of the damping bases of Sect. 2.2 without,

however, getting into too much realistic detail, we will consider the following somewhat idealized scenario. The atoms come at the regularly spaced instants $t = \ldots, -2T, -T, 0, T, 2T, \ldots$ and each atom effects a quasi-instantaneous change of the field of the cavity mode ($\equiv$ the oscillator) that is specified by a "kick operator" $\mathcal{K}$ which is such that

$$\rho_{jT-0} \to \rho_{jT+0} = (1 + \mathcal{K})\rho_{jT-0} \tag{2.134}$$

states how ρ_t changes abruptly at $t = jT$ ($j = 0, \pm 1, \pm 2, \ldots$). Physically speaking, this abruptness just means that the time spent by an atom inside the resonator is very short on the scale set by the decay constant A.

The evolution between the kicks follows the master equation (2.29). We have

$$\frac{\partial}{\partial t}\rho_t = \mathcal{L}\rho_t + \mathcal{K}\sum_j \delta(t - jT)\rho_{t-0} \tag{2.135}$$

as a formal master equation that incorporates the kicks (2.134). The $\mathcal{L}\rho_t$ term accounts for the decay of the photon field in the resonator. In Sect. 2.1.1 we derived the form of $\mathcal{L}$ by pretending that the dissipation is the result of a very weak interaction with very many atoms. But, of course, this model must not be regarded as the true physical mechanism. Actually, the loss of electromagnetic energy is partly caused by leakage through the cavity openings (through which the atoms enter and leave) and partly by the ohmic resistance from which the currents suffer that are induced on the surface of the conducting walls. The ohmic losses are kept very small by fabricating the cavity from superconducting metal (niobium below its critical temperature).

Thus, the term $\mathcal{L}\rho_t$ in (2.135) has nothing to do with the atoms that the experimenter passes through the resonator in a micromaser experiment. These atoms interact strongly with the photons and give rise to the kick operator $\mathcal{K}$. Figure 2.4 shows what to expect under the circumstances to which (2.135) refers. The mean number of excitations, $\langle a^\dagger a \rangle_t$, decays between the kicks (solid —— line) in accordance with (2.32) and changes abruptly when a kick happens (vertical dashed $- - -$ lines at $At = 0.4, 0.8, 1.2, \ldots$). After an initial period, which lasts about a dozen kicks in Fig. 2.4, the cyclically steady state $\rho_t^{(\mathrm{css})}$ is reached whose defining property is that it is the periodic solution of (2.135),

$$\rho_{t+T}^{(\mathrm{css})} = \rho_t^{(\mathrm{css})} \,. \tag{2.136}$$

Its value just before a kick is determined by

$$\rho_{t=-0}^{(\mathrm{css})} = \mathrm{e}^{\mathcal{L}T}\rho_{t=+0}^{(\mathrm{css})} = \mathrm{e}^{\mathcal{L}T}(1 + \mathcal{K})\rho_{t=-0}^{(\mathrm{css})} \,, \tag{2.137}$$

and we have

$$\rho_t^{(\mathrm{css})} = \mathrm{e}^{\mathcal{L}t}\rho_{t=+0}^{(\mathrm{css})} = \mathrm{e}^{-\mathcal{L}(T-t)}\rho_{t=-0}^{(\mathrm{css})} \tag{2.138}$$

for $0 < t < T$, that is: between two successive kicks.

In passing, it is worth noting that a recent experiment [35], in which photon states of a definite photon number (Fock states) were prepared in a micromaser,

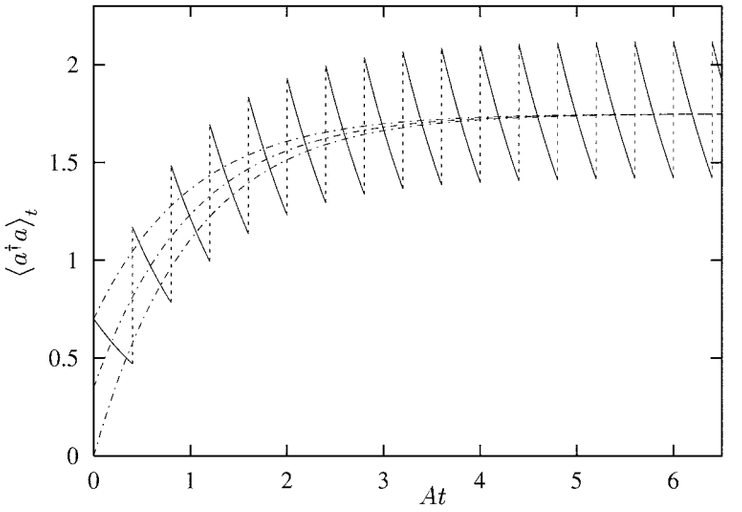


Fig. 2.4. Mean number of excitations of a periodically kicked oscillator; see text

used a periodic scheme for pumping and probing. The theoretical analysis [36] benefitted from damping-bases techniques.

The fine detail that we see in Fig. 2.4 is usually not of primary interest, partly because experiments tend to not resolve it. For example, if one asks how long it takes to reach the cyclically steady state, all one needs to know is the time-averaged behavior of the smooth dash-dotted $-\cdot-\cdot-$ lines in Fig. 2.4. In the cyclically steady state, the meaning of "time-averaged" is hardly ambiguous, we simply have

$$\bar{\rho}^{(\mathrm{ss})} = \frac{1}{T}\int_0^T \mathrm{d}t\, \rho_t^{(\mathrm{css})} \,, \tag{2.139}$$

where it does not matter over which time interval we average as long as it covers one or more periods of $\rho_t^{(\mathrm{css})}$. The periodicity (2.136) of $\rho_t^{(\mathrm{css})}$ implies that $\partial\rho_t^{(\mathrm{css})}/\partial t$ is zero on average and, therefore, we obtain

$$\mathcal{L}\bar{\rho}^{(\mathrm{ss})} + \frac{1}{T}\mathcal{K}\rho_{t=-0}^{(\mathrm{css})} = 0 \tag{2.140}$$

when time averaging (2.135). When combined with what we get upon using (2.138) in (2.139),

$$\bar{\rho}^{(\mathrm{ss})} = \frac{1-\mathrm{e}^{-\mathcal{L}T}}{\mathcal{L}T}\rho_{t=-0}^{(\mathrm{css})} \,, \tag{2.141}$$

it yields the equation that determines $\bar{\rho}^{(\mathrm{ss})}$,

$$\mathcal{L}\bar{\rho}^{(\mathrm{ss})} + \mathcal{K}\frac{\mathcal{L}}{1-\mathrm{e}^{-\mathcal{L}T}}\bar{\rho}^{(\mathrm{ss})} = 0 \,. \tag{2.142}$$

For $\mathcal{K} = 0$, it is of course solved by $\rho^{(\mathrm{ss})} = \rho_0^{(0)}$ of (2.38).

A master equation for the time-averaged evolution, that is: an equation obeyed by the time-averaged statistical operator $\bar{\rho}_t$, cannot be *derived* from (2.135) for the same reasons for which one cannot derive the macroscopic Maxwell equations from the microscopic ones. But they can be *inferred* with physical arguments that are more than just reasonably convincing. The task is actually easier here because we have to deal with temporal averages only whereas one also needs spatial averages in the case of electromagnetism.

Imagine, then, that a linear time average is taken of (2.135),

$$\frac{\partial}{\partial t}\bar{\rho}_t = \mathcal{L}\bar{\rho}_t + \mathcal{K}\,(\textbf{?})\,\bar{\rho}_t \;, \tag{2.143}$$

where $(\textbf{?})\,\bar{\rho}_t$ is the ill-determined average of the summation in (2.135) that accounts for the periodic kicks (2.134). We require, of course, that $\bar{\rho}^{(\mathrm{ss})}$ is the steady state of (2.143). In view of (2.142), this requirement settles the issue [37],

$$\frac{\partial}{\partial t}\bar{\rho}_t = \mathcal{L}\bar{\rho}_t + \mathcal{K}\frac{\mathcal{L}}{1-\,\mathrm{e}^{-\mathcal{L}T}}\bar{\rho}_t \;, \tag{2.144}$$

which we now accept as the master equation that describes the time-averaged evolution. This is another case where an equation is ultimately justified by its consequences.

We run a simple, but important consistency check on (2.144). If the spacing T between the atoms decreases, $T \to 0$, and also the effect of a single atom, $\mathcal{K} = p\mathcal{M}$ with $p \to 0$, such that their ratio $r = p/T$ is constant, then the situation should be equivalent to that of Poissonian arrival statistics with rate r and each atom effecting a kick $\mathcal{M}$. Indeed, (2.144) turns into

$$\frac{\partial}{\partial t}\bar{\rho}_t = \mathcal{L}\bar{\rho}_t + r\mathcal{M}\bar{\rho}_t \;, \tag{2.145}$$

as it should, because this is the familiar Scully–Lamb equation that is known to apply in the case of Poissonian statistics. Thus, with $\mathcal{K} = rT\mathcal{M}$ in (2.144) we obtain a master equation,

$$\frac{\partial}{\partial t}\bar{\rho}_t = \mathcal{L}\bar{\rho}_t + r\mathcal{M}\frac{\mathcal{L}T}{1-\,\mathrm{e}^{-\mathcal{L}T}}\bar{\rho}_t \;, \tag{2.146}$$

that interpolates between the Poissonian Scully–Lamb limit of $T = 0$ and that of highly regular arrival times, $T = 1/r$.

The damping bases associated with $\mathcal{L}$ are the crucial tool for handling (2.144). We write

$$\bar{\rho}_t = \sum_{n=0}^{\infty}\sum_{k=-\infty}^{\infty} \bar{\alpha}_n^{(k)}(t)\rho_n^{(k)} \tag{2.147}$$

and obtain differential equations for the numerical coefficients $\bar{\alpha}_n^{(k)}$,

$$\bar{\alpha}_n^{(k)}(t) = \mathrm{tr}\left\{\check{\rho}_n^{(k)}\bar{\rho}_t\right\} \;, \tag{2.148}$$

by exploiting

$$f(\mathcal{L})\rho_n^{(k)} = f\big(\lambda_n^{(k)}\big)\rho_n^{(k)} \ , \qquad \lambda_n^{(k)} = -\mathrm{i}k\omega - (n+\tfrac{1}{2}|k|)A \ , \tag{2.149}$$

which holds for any function $f(\mathcal{L})$ simply because $\rho_n^{(k)}$ is the right eigenvector of $\mathcal{L}$ to eigenvalue $\lambda_n^{(k)}$. In this way, (2.144) implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\alpha}_n^{(k)} = \lambda_n^{(k)}\bar{\alpha}_n^{(k)} + \sum_{n',k'} \mathcal{K}_{n,n'}^{(k,k')} \frac{\lambda_{n'}^{(k')}}{1-\,\mathrm{e}^{-\lambda_{n'}^{(k')}T}} \bar{\alpha}_{n'}^{(k')} \tag{2.150}$$

where

$$\mathcal{K}_{n,n'}^{(k,k')} = \mathrm{tr}\left\{\check{\rho}_n^{(k)}\mathcal{K}\rho_{n'}^{(k')}\right\} \tag{2.151}$$

is the matrix representation of the kick operator $\mathcal{K}$ in the damping bases. The seemingly troublesome ratio in (2.144) is not a big deal anymore in (2.150), where

$$\left.\frac{\lambda}{1-\,\mathrm{e}^{-\lambda T}}\right|_{\lambda\to\lambda_0^{(0)}=0} = \frac{1}{T} \ , \tag{2.152}$$

of course.

Let us illustrate this for the particularly simple kick operator specified by

$$\mathcal{K}\rho = p\left(a^\dagger \frac{1}{\sqrt{aa^\dagger}}\rho\frac{1}{\sqrt{aa^\dagger}}a - \rho\right) \qquad \text{with} \quad 0\le p\le 1 \ , \tag{2.153}$$

which describes the over-idealized situation in which an atom adds one photon with probability p and does nothing with probability $1-p$. Here,

$$\mathcal{K}_{n,n'}^{(k,k')} = \delta_{k,k'}\mathcal{K}_{n,n'}^{(k,k)} \tag{2.154}$$

so that the evolution does not mix $\bar{\alpha}_n^{(k)}$'s of different k values. As a further simplification it is therefore permissible to just consider the $k=0$ terms. For $\nu=0$ (homework assignment 17 deals with $\nu>0$), the generating functions (2.55) and (2.58) give

$$\begin{aligned}\sum_{m,n=0}^{\infty} y^m\mathcal{K}_{m,n}^{(0,0)}x^n &= \mathrm{tr}\left\{(1+y)^{a^\dagger a}\,\mathcal{K}\frac{1}{1+x}\left(\frac{x}{1+x}\right)^{a^\dagger a}\right\} \\ &= \frac{py}{1-xy} = \sum_{n=0}^{\infty} py^{n+1}x^n \ ,\end{aligned} \tag{2.155}$$

where we let $\mathcal{K}$ act to the left,

$$(1+y)^{a^\dagger a}\mathcal{K} = py(1+y)^{a^\dagger a} \ , \tag{2.156}$$

and recall the trace evaluation of (2.64). We find

$$\mathcal{K}_{m,n}^{(0,0)} = p\delta_{m,n+1} \ , \tag{2.157}$$

and the equation for $\bar{\alpha}_n^{(0)}(t)$ then has the explicit form

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\alpha}_n^{(0)} = -nA\bar{\alpha}_n^{(0)} + p\frac{(n-1)A}{\mathrm{e}^{(n-1)AT}-1}\bar{\alpha}_{n-1}^{(0)} \,. \tag{2.158}$$

This differential recurrence relation is solved successively by

$$\begin{aligned}
\bar{\alpha}_0^{(0)}(t) &= \mathrm{tr}\,\{\bar{\rho}_t\} = 1 \,, \\
\bar{\alpha}_1^{(0)}(t) &= \langle a^\dagger a\rangle_t = \langle a^\dagger a\rangle_\infty + \left(\langle a^\dagger a\rangle_0 - \langle a^\dagger a\rangle_\infty\right)\mathrm{e}^{-At} \,, \\
\bar{\alpha}_2^{(0)}(t) &= \frac{1}{2}\langle {a^\dagger}^2 a^2\rangle_t = \frac{1}{2}\langle {a^\dagger}^2 a^2\rangle_\infty + \frac{1}{2}\left(\langle {a^\dagger}^2 a^2\rangle_0 - \langle {a^\dagger}^2 a^2\rangle_\infty\right)\mathrm{e}^{-2At} \\
&\quad + \langle {a^\dagger}^2 a^2\rangle_\infty \left(\frac{\langle a^\dagger a\rangle_0}{\langle a^\dagger a\rangle_\infty} - 1\right)\left(\mathrm{e}^{-At} - \mathrm{e}^{-2At}\right)
\end{aligned} \tag{2.159}$$

and so forth, where

$$\langle a^\dagger a\rangle_\infty = \frac{p}{AT} \,, \quad \langle {a^\dagger}^2 a^2\rangle_\infty = \frac{p}{AT}\frac{p}{\mathrm{e}^{AT}-1} \,, \quad \ldots \tag{2.160}$$

are the expectations values in the time-averaged steady state $\bar{\rho}^{(\mathrm{ss})}$.

Actually, Fig. 2.4 just shows this $\langle a^\dagger a\rangle_t$ for $p = 0.7$ and $AT = 0.4$, so that $\langle a^\dagger a\rangle_\infty = 1.75$ and the approach to this asymptotic value is plotted for $\langle a^\dagger a\rangle_0 = 0, 0.35, 0.7$ by the three dash-dotted $-\cdot-\cdot-$ curves. The time-averaged value of $\langle a^\dagger a\rangle_t$ is not well defined at $t = 0$, the instant of the first kick, any value in the range $0 \cdots 0.7$ can be justified equally well. The memory of this arbitrary initial value is always lost quickly.

2.4.2 Conditional and Unconditional Evolution

Let us now be more realistic about the effect an atom has on the photon state. In fact, we have worked that out already in Sect. 2.1 for the case of atoms incident in state $\overset{\bullet}{-}$ or in state $\underset{\bullet}{-}$ and a resonant Jaynes–Cummings coupling between the photons and the atoms. For the theoretical description of many one-atom maser experiments this is actually quite accurate, and it will surely do for the purpose of these lectures.

Since the atoms should deposit energy into the resonator as efficiently as possible, they are prepared in the $\overset{\bullet}{-}$ state. The net effect of a single atom is then available in (2.25), which we now present conveniently as

$$\mathcal{M}\rho_t = \mathcal{A}\rho_t + \mathcal{B}\rho_t - \rho_t = (\mathcal{A} + \mathcal{B} - 1)\rho_t \tag{2.161}$$

with

$$\begin{aligned}
\mathcal{A}\rho_t &= a^\dagger \frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}}\rho_t \frac{\sin\left(\phi\sqrt{aa^\dagger}\right)}{\sqrt{aa^\dagger}} a \,, \\
\mathcal{B}\rho_t &= \cos\left(\phi\sqrt{aa^\dagger}\right)\rho_t \cos\left(\phi\sqrt{aa^\dagger}\right) \,.
\end{aligned} \tag{2.162}$$

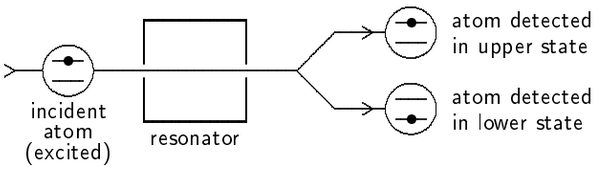


Fig. 2.5. The final state of the exiting atom is detected: Is it $\overset{\bullet}{-}$ or $\underset{\bullet}{-}$?

As the derivation in Sect. 2.1 shows, the term $\mathcal{A}\rho_t$ corresponds to the atom emerging in state $\underset{\bullet}{-}$, and likewise $\mathcal{B}\rho_t$ refers to $\overset{\bullet}{-}$. Accordingly, the respective probabilities for the final atom states $\underset{\bullet}{-}$ and $\overset{\bullet}{-}$ are

$$\text{prob}(\overset{\bullet}{-} \to \underset{\bullet}{-}) = \text{tr}\,\{\mathcal{A}\rho_t\}\ , \qquad \text{prob}(\overset{\bullet}{-} \to \overset{\bullet}{-}) = \text{tr}\,\{\mathcal{B}\rho_t\}\ , \tag{2.163}$$

and the probabilities $p_{\underset{\bullet}{-}}$, $p_{\overset{\bullet}{-}}$ that the state-selective detection of Fig. 2.5 finds the atom in $\underset{\bullet}{-}$ or $\overset{\bullet}{-}$ are

$$p_{\underset{\bullet}{-}} = \eta_{\underset{\bullet}{-}}\ \text{tr}\,\{\mathcal{A}\rho_t\}\ , \qquad p_{\overset{\bullet}{-}} = \eta_{\overset{\bullet}{-}}\ \text{tr}\,\{\mathcal{B}\rho_t\}\ , \tag{2.164}$$

respectively, where $\eta_{\underset{\bullet}{-}}, \eta_{\overset{\bullet}{-}}$ are the detection efficiencies.

The effect on the photon state of an atom traversing the cavity at time t can, therefore, be written as

$$\rho_t \to \eta_{\underset{\bullet}{-}}\mathcal{A}\rho_t + \eta_{\overset{\bullet}{-}}\mathcal{B}\rho_t + \big[(1-\eta_{\underset{\bullet}{-}})\mathcal{A}\rho_t + (1-\eta_{\overset{\bullet}{-}})\mathcal{B}\rho_t\big] \tag{2.165}$$

where the three terms correspond to detecting the atom in state $\underset{\bullet}{-}$, detecting it in state $\overset{\bullet}{-}$, and not detecting it at all. The probability for the latter case is

$$\text{prob(no click)} = \text{tr}\,\big\{(1-\eta_{\underset{\bullet}{-}})\mathcal{A}\rho_t + (1-\eta_{\overset{\bullet}{-}})\mathcal{B}\rho_t\big\} = 1 - \text{tr}\,\{\mathcal{C}\rho_t\}\ , \tag{2.166}$$

where we recognize that

$$\text{prob}(\overset{\bullet}{-} \to \underset{\bullet}{-}) + \text{prob}(\overset{\bullet}{-} \to \overset{\bullet}{-}) = \text{tr}\,\{(\mathcal{A}+\mathcal{B})\rho_t\} = 1 \tag{2.167}$$

and introduce the *click operator* $\mathcal{C}$,

$$\mathcal{C} = \eta_{\underset{\bullet}{-}}\mathcal{A} + \eta_{\overset{\bullet}{-}}\mathcal{B}\ . \tag{2.168}$$

We take for granted that the atoms arrive with rate r at statistically independent instants (Poissonian arrival statistics once more). The change of ρ_t brought about by a single *undetected* atom is

$$\Delta\rho_t\Big|_{\text{undetected atom}} = \frac{(\mathcal{A}+\mathcal{B}-\mathcal{C})\rho_t}{1-\text{tr}\,\{\mathcal{C}\rho_t\}} - \rho_t\ , \tag{2.169}$$

where the numerator is just the third term of (2.165) and the denominator is its trace. We multiply this with the probability that there is an atom between t and

$t + \mathrm{d}t$, which is $r\mathrm{d}t$, and with the probability that the atom escapes detection, which is given in (2.166) and equal to the denominator in (2.169), and so get

$$\mathrm{d}t \frac{\partial \rho_t}{\partial t}\Big|_{\text{undetected atoms}} = r\mathrm{d}t\big[(\mathcal{A}+\mathcal{B}-\mathcal{C})\rho_t - \rho_t + \mathrm{tr}\,\{\mathcal{C}\rho_t\}\,\rho_t\big]\ . \tag{2.170}$$

We combine it with (2.29),

$$\frac{\partial \rho_t}{\partial t}\Big|_{\text{photon decay}} = \mathcal{L}\rho_t\ , \tag{2.171}$$

to arrive at

$$\frac{\partial}{\partial t}\rho_t = \big[\mathcal{L} + r(\mathcal{A}+\mathcal{B}-1)\big]\rho_t - r\big[\mathcal{C} - \mathrm{tr}\,\{\mathcal{C}\rho_t\}\big]\rho_t\ , \tag{2.172}$$

the master equation that applies *between detection events*. Owing to the term that involves the click probability $\mathrm{tr}\,\{\mathcal{C}\rho_t\}$, this is a *nonlinear* master equation unless $\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} \equiv \eta$ when $\mathrm{tr}\,\{\mathcal{C}\rho_t\} = \eta$ for all ρ_t. Fortunately, the nonlinearity is of a very mild form, since we can write (2.172) as

$$\frac{\partial}{\partial t}\rho_t = \mathcal{L}_\eta\rho_t - \mathrm{tr}\,\{\mathcal{L}_\eta\rho_t\}\,\rho_t \tag{2.173}$$

with the *linear* operator $\mathcal{L}_\eta$ given by

$$\mathcal{L}_\eta = \mathcal{L} + r(\mathcal{A}+\mathcal{B}-1) - r\mathcal{C} = \mathcal{L} + r\mathcal{M} - r\mathcal{C}\ , \tag{2.174}$$

and then solve it by

$$\rho_t = \frac{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}}{\mathrm{tr}\left\{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}}\ . \tag{2.175}$$

In effect, we can just ignore the second term in (2.173), evolve ρ_t linearly from $\rho_{t=0}$ to $\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}$, and then normalize this to unit trace. The normalization is necessary because $\mathcal{L}_\eta$ by itself does not conserve the trace, except for $\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} = 0$. As we will see in the fifth lecture, the normalizing denominator of (2.175) has a simple and important physical significance; see Sect. 2.5.2.

Note that there is a great difference between the situation in which the atoms are not observed (you don't listen) and the situation in which they are not detected (you listen but you don't hear anything). The evolution of the photon field with unobserved atoms is the $\mathcal{C} = 0$ version of (2.172) that obtains for $\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} = 0$, which is just the Scully–Lamb equation (2.145) with $\mathcal{M}$ of (2.161). It describes the *unconditional* evolution of the photon state. By contrast, if the atoms are under observation but escape detection, the nonlinear master equation (2.172) or (2.173) applies. It describes the *conditional* evolution of the photon state, conditioned by the constraint that there are no detection events although detection is attempted.

The difference between conditional and unconditional evolution is perhaps best illustrated in the extreme circumstance of perfectly efficient detectors, $\eta_{\overline{\bullet}} =$

$\eta_{\dot{-}} = 1$, when no atom escapes detection. Then (2.172) turns into (2.171) – as it should because "between detection events" is tantamount to "between atoms" if every atom is detected. More generally, if the detection efficiency is the same for $\bar{\cdot}$ and $\dot{-}$, $\eta_{\bar{\cdot}} = \eta_{\dot{-}} = \eta$, so that each atom is detected with probability η, we have

$$\frac{\partial}{\partial t}\rho_t = \big[\mathcal{L} + (1-\eta)r\mathcal{M}\big]\rho_t \tag{2.176}$$

for the evolution between detection events. This is the Scully–Lamb equation with the actual rate r replaced by the effective rate $(1-\eta)r$, the rate of undetected atoms.

2.4.3 Physical Significance of Statistical Operators

The master equation (2.172) is the generalization of (2.176) that takes into account that $\bar{\cdot}$ atoms are not detected with the same efficiency as $\dot{-}$ atoms. This asymmetry may originate in actually different detection devices or – and this is in fact the more important situation – it is a consequence of the question we are asking.

Assume, for example, that the $\bar{\cdot}$ detector clicked at $t = 0$ and you want to know how probable it is that the *next* click of the $\bar{\cdot}$ detector occurs between t and $t + \mathrm{d}t$. Clicks of the $\dot{-}$ detector are of no interest to you whatsoever. In an experiment you would just ignore them because they have no bearing on your question. The same deliberate ignorance enters the theoretical treatment: you employ (2.172) with $\eta_{\dot{-}} = 0$ in the click operator (2.168) irrespective of the actual efficiency of the $\dot{-}$ detector used in the experiment. Likewise if your question was about the next click of the $\dot{-}$ detector you would have to put $\eta_{\bar{\cdot}} = 0$ in (2.168).

All of this is well in accord with the physical significance of the statistical operator ρ_t: it serves the sole purpose of enabling us to make correct predictions about measurement at time t, in particular about the probability that a certain outcome is obtained if a measurement is performed. Such probabilities are always conditioned, they naturally depend on the constraints to which they are subjected. Therefore, it is quite possible that two persons have different statistical operators for the same physical object because they take different conditions into account.

Let us illustrate this point by a detection scheme that is simpler, and more immediately transparent, than the standard one-atom maser experiment specified by $\mathcal{A}$ and $\mathcal{B}$ of (2.162). Instead we take

$$\begin{aligned}\mathcal{A}\rho_t &= \frac{1+(-1)^{a^\dagger a}}{2}\rho_t\frac{1+(-1)^{a^\dagger a}}{2}\,,\\ \mathcal{B}\rho_t &= \frac{1-(-1)^{a^\dagger a}}{2}\rho_t\frac{1-(-1)^{a^\dagger a}}{2}\,,\end{aligned} \tag{2.177}$$

which can be realized by suitably prepared two-level atoms that have a nonresonant interaction with the photon field and a suitable manipulation prior to

detecting $\overset{\bullet}{-}$ or $\underset{\bullet}{+}$ [38]. What is measured in such an experiment is the value of $(-1)^{a^\dagger a}$, the parity of the photon state. Detecting the atom in state $\overset{\bullet}{-}$ indicates even parity, $(-1)^{a^\dagger a} = 1$, and a $\underset{\bullet}{+}$ click indicates odd parity, $(-1)^{a^\dagger a} = -1$.

Now consider the four cases of Figs. 2.6 and 2.7. They refer to parity measurements on a one-atom maser that is not pumped (no resonant $\underset{\bullet}{+}$ atoms are sent through) with $\nu = 2$ (an atypically large number of thermal photons for a micromaser experiment) and $r/A = 10$. The plots show the period $t = 0 \cdots 100/r$ of the simulated experiment. In Fig. 2.6 we see the parity expectation value as a function of t, and in Fig. 2.7 we have the expectation value of the photon number. The solid —— lines are the actual values, the vertical dashed $- - -$ lines guide the eye through state-reduction jumps, and the horizontal dash-dotted $- \cdot - \cdot -$ lines indicate the steady state values

$$\left\langle (-1)^{a^\dagger a} \right\rangle^{(\mathrm{ss})} = \frac{1}{5} \,, \quad \left\langle a^\dagger a \right\rangle^{(\mathrm{ss})} = 2 \,. \tag{2.178}$$

On average, 100 atoms traverse the resonator in this time span, the actual number is 108 here, of which 67 emerge in state $\overset{\bullet}{-}$ (even parity) and 41 in state $\underset{\bullet}{+}$ (odd parity). The final state is $\overset{\bullet}{-}$ for 60% of the atoms on average, and $\underset{\bullet}{+}$ for 40%. The values chosen for the detection efficiencies are $\eta_{\overset{\bullet}{-}} = 10\%$ and $\eta_{\underset{\bullet}{+}} = 15\%$ so that each detector should register 6 atoms on average in a period of this duration. In fact, 7 $\overset{\bullet}{-}$ clicks occurred (when rt=38.51, 44.80, 49.52, 53.07, 72.05, 76.41, and 76.75) and 5 $\underset{\bullet}{+}$ clicks (when rt=3.88, 85.81, 86.09, 94.12, and 94.90).

Experimenter Bob pays no attention to the even-parity clicks of the $\overset{\bullet}{-}$ detector, he is either not aware of them or has reasons to ignore them deliberately. Therefore he uses the nonlinear master equation (2.173) with $\eta_{\overset{\bullet}{-}} = 0$ and $\eta_{\underset{\bullet}{+}} = 0.15$ for the evolution between two successive clicks of the $\underset{\bullet}{+}$ detector, and performs the state reduction

$$\rho_t \to \frac{\mathcal{B}\rho_t}{\mathrm{tr}\,\{\mathcal{B}\rho_t\}} \tag{2.179}$$

whenever a $\underset{\bullet}{+}$ click happens. For example, to find the statistical operator $\rho_{t=60/r}$ and then the expectation values

$$\text{Bob, } rt = 60 : \quad \left\langle (-1)^{a^\dagger a} \right\rangle = 0.1920 \,, \quad \left\langle a^\dagger a \right\rangle = 1.787 \,, \tag{2.180}$$

he applies (2.179) to the $\underset{\bullet}{+}$ click at $rt = 3.88$, the only one between $rt = 0$ and $rt = 60$. Bob's $\rho_{t=0}$ is the steady state $\rho^{(\mathrm{ss})}$ of the Scully–Lamb equation, consistent with his knowledge that the experiment has been running long enough to have lost memory of its early history. For (2.177), $\rho^{(\mathrm{ss})}$ is actually the thermal state (2.38), here with $\nu = 2$. Bob's detailed accounts are reported in Figs. 2.6b and 2.7b.

Similarly, Figs. 2.6c and 2.7c show what Chuck has to say who pays no attention to $\underset{\bullet}{+}$ clicks, but keeps a record of $\overset{\bullet}{-}$ clicks. He uses (2.173) with

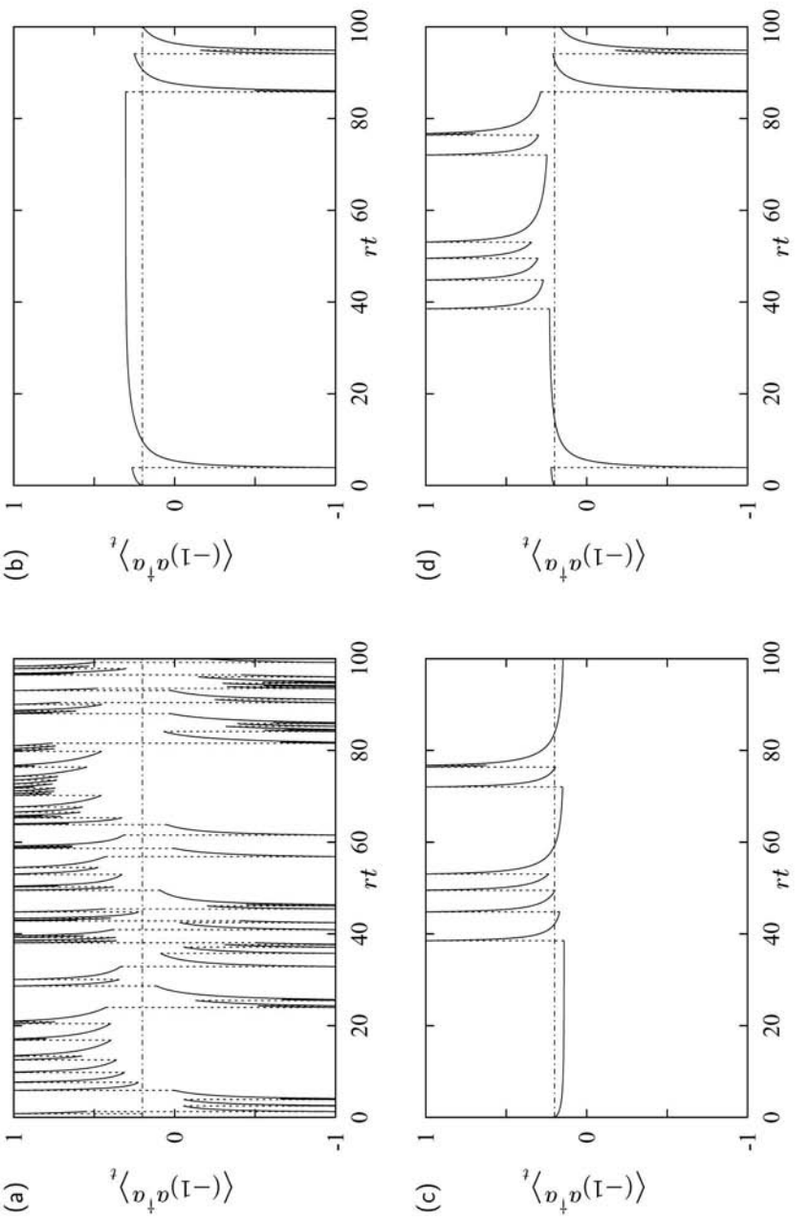


Fig. 2.6. Different, but equally consistent, statistical predictions about the same physical system: Parity expectation value when parity measurements are performed on an unpumped resonator; see text

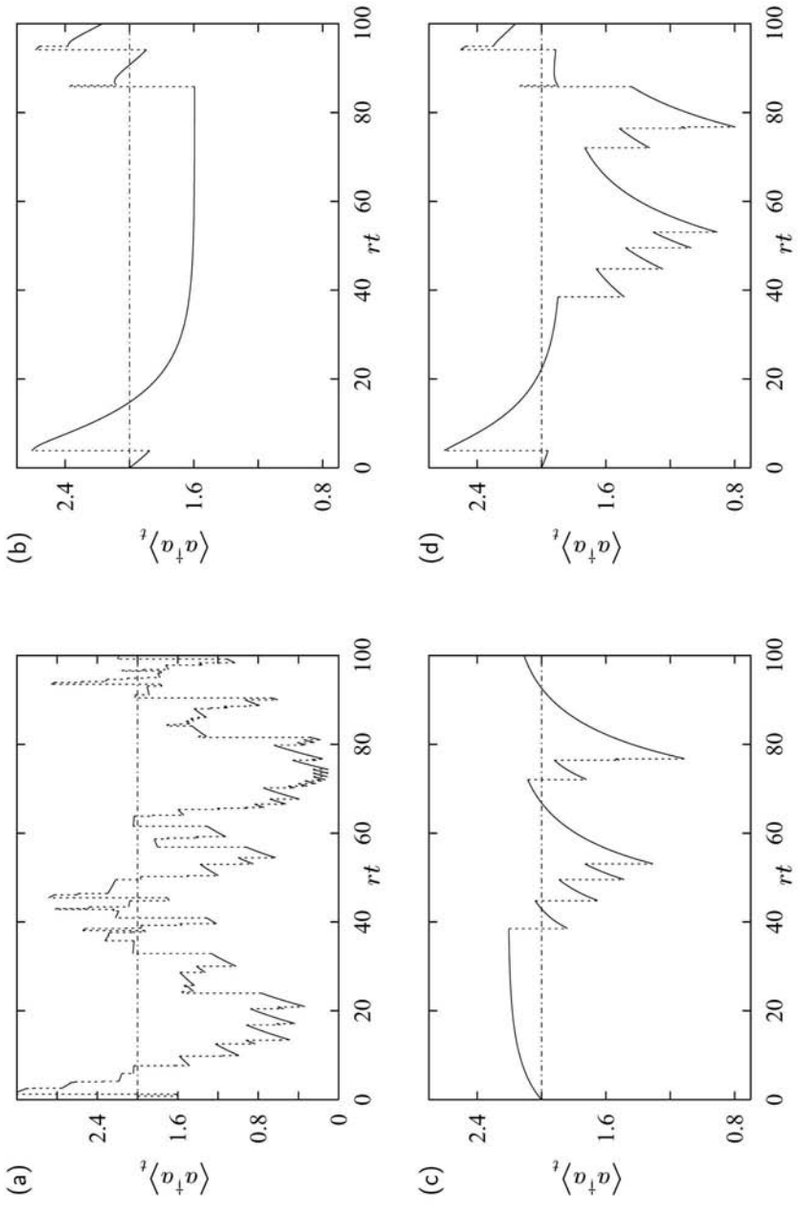


Fig. 2.7. Different, but equally consistent, statistical predictions about the same physical system: Photon number expectation value when parity measurements are performed on an unpumped resonator; see text

$\eta_{\overline{\bullet}} = 0.1$ and $\eta_{\underline{\bullet}} = 0$ for the evolution between two successive $\overline{\bullet}$ clicks and performs the state reduction

$$\rho_t \to \frac{\mathcal{A}\rho_t}{\mathrm{tr}\,\{\mathcal{A}\rho_t\}} \tag{2.181}$$

for each $\overline{\bullet}$ click. To establish

$$\text{Chuck, } rt = 60: \quad \left\langle (-1)^{a^\dagger a}\right\rangle = 0.3042\,, \quad \left\langle a^\dagger a\right\rangle = 1.599 \tag{2.182}$$

he has to do this for the four clicks prior to $t = 60/r$. Chuck uses the same $\rho_{t=0}$ as Bob, of course, because both have the same information about the preparation.

And then there is Doris who pays full attention to all detector clicks. She uses (2.173) with $\eta_{\overline{\bullet}} = 0.1$ and $\eta_{\underline{\bullet}} = 0.15$ between successive clicks, does (2.181) for $\overline{\bullet}$ clicks and (2.179) for $\underline{\bullet}$ clicks, and arrives at

$$\text{Doris, } rt = 60: \quad \left\langle (-1)^{a^\dagger a}\right\rangle = 0.2995\,, \quad \left\langle a^\dagger a\right\rangle = 1.390\,. \tag{2.183}$$

Her account is shown in Figs. 2.6d and 2.7d.

For the sake of completeness, we also have Figs. 2.6a and 2.7a, where state reductions are performed for each atom, whether detected or not, and the $\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} = 0$ version of (2.173) – the "between atoms" equation (2.171) – applies between the reductions. What is obtained in this manner is of no consequence, however, because it incorporates data that are never actually available.

Why, then, do we show Figs. 2.6a and 2.7a at all? Because one might think that they report the "true state of affairs" so that

$$\text{all atoms, } rt = 60: \quad \left\langle (-1)^{a^\dagger a}\right\rangle = 0.4818\,, \quad \left\langle a^\dagger a\right\rangle = 1.189 \tag{2.184}$$

would be the "true expectation values" of $(-1)^{a^\dagger a}$ and $a^\dagger a$ at $t = 60/r$. And then one would conclude that the accounts given by Bob, Chuck, and Doris are wrong in some sense. In fact, all three give *correct*, though differing accounts, and the various predictions for $t = 60/r$ in (2.180), (2.182), and (2.183) are all statistically correct. For, if you repeat the experiment very often you will find that Bob's expectation values are confirmed by the data, and so are Chuck's, and so are Doris's.

But, of course, when extracting $\left\langle a^\dagger a\right\rangle_{t=60/r}$, say, from the data of the very many runs, you must take different subensembles for checking Bob's predictions, or Chuck's, or Doris's. Bob's prediction (2.180) refers to the subensemble characterized by a single $\underline{\bullet}$ click at $rt = 3.88$ and no other $\underline{\bullet}$ click between $rt = 0$ and $rt = 60$, but any number of $\overline{\bullet}$ clicks. Likewise, Chuck's prediction (2.182) is about the subensemble that has $\overline{\bullet}$ clicks at rt=38.51, 44.80, 49.52, and 53.07, no other $\overline{\bullet}$ clicks and any number of $\underline{\bullet}$ clicks. And Doris's subensemble is specified by having this one $\underline{\bullet}$ click, these four $\overline{\bullet}$ clicks, and no other clicks of either kind. By contrast, no experimentally identifiable ensemble corresponds to Figs. 2.6a and 2.7a; they represent sheer imagination, not phenomenological reality.

In summary, although it is the same physical object (the privileged mode of the resonator) that Bob, Chuck, and Doris make predictions about, they regard it as a representative of different ensembles, which are respectively characterized by the information taken into account. This illustrates the basic fact that a statistical operator is just an encoding of what we know about the system. Depending on the question we are asking, we may even have to ignore some of the information deliberately. The appropriate ρ_t is the one that pays due attention to the pertinent conditions under which we wish to make statistical predictions. Different conditions simply require different statistical operators. That there are several ρ_t's for the same physical object is then not bewildering, but rather expected.

The ensembles that Bob, Chuck, and Doris make statements about need not be, and usually are not, real ensembles created by many repeated runs of the experiment (or, since the system is ergodic – see [39,40] and Sect. 4.10 – perhaps a single run of very long duration). Rather, they are Gibbs ensembles in the standard meaning of statistical mechanics: imagined ensembles characterized by the respective constraints and consistent weights. The constraints are duly taken into account by the nonlinear master equation (2.173).

We should also mention that Bob's ρ_t is consistent with Chuck's and Doris's by construction. If we didn't know how they arrive at their respective statistical operators we might wonder how we could verify that the three ρ_t's do not contradict each other. Consult [41,42,43] if you find the question interesting.

Homework Assignments

17 Show that

$$\sum_{m=0}^{\infty} y^m \check{\rho}_m^{(0)} = \frac{1+\nu}{1+\nu+\nu y}\left(1+\frac{y}{1+\nu+\nu y}\right)^{a^\dagger a} \tag{2.185}$$

is the $\nu > 0$ generalization of (2.58), then use this and (2.54) to find the $\nu > 0$ version of (2.157).

18 State the $\nu > 0$ version of (2.158) and solve the equations for $n = 0, 1, 2$.

19 Take the Scully–Lamb limit of (2.158), that is: $T \to 0$ after putting $p = rT$, and find the steady state values of all $\bar{\alpha}_n^{(0)}$'s.

20 Evaluate $\bar{\rho}^{(\mathrm{ss})} = \sum_{n=0}^{\infty} \bar{\alpha}_n^{(0)} \rho_n^{(0)}$ for these $\bar{\alpha}_n^{(0)}$'s. You should obtain

$$\bar{\rho}^{(\mathrm{ss})} = \frac{1}{\nu+1}\left(\frac{\nu}{\nu+1}\right)^{a^\dagger a} \mathrm{L}_{a^\dagger a}^{(0)}\left(-\frac{r}{\nu A}\right) \mathrm{e}^{-r/A} \,. \tag{2.186}$$

In which sense is this yet another generating function for the $\rho_n^{(0)}$'s? What do you get for $\nu \to 0$?

21 Verify (2.167) for $\mathcal{A}$ and $\mathcal{B}$ of (2.162).

22 Show that

$$\mathcal{M}f(a^\dagger a) = \left[\sin\left(\phi\sqrt{a^\dagger a}\right)\right]^2 f(a^\dagger a - 1) - \left[\sin\left(\phi\sqrt{a^\dagger a + 1}\right)\right]^2 f(a^\dagger a) \tag{2.187}$$

for $\mathcal{M}$ of (2.161), then combine this with (2.35) to find the steady state of the Scully–Lamb equation (2.145). You should get

$$\rho^{(\mathrm{ss})} = N \prod_{n=1}^{a^\dagger a} \left[\frac{\nu}{\nu+1} + \frac{r/A}{\nu+1}\left(\frac{\sin\left(\phi\sqrt{n}\,\right)}{\sqrt{n}}\right)^2\right] \tag{2.188}$$

as the steady state of the one-atom maser. What is the physical significance of the normalization factor N? How do you determine it?

23 Differentiate ρ_t of (2.175) to verify that it solves (2.173).

24 In addition to Bob, Chuck, and Doris, there is also Alice who pays attention to all detector clicks but doesn't care which detector fires. Which version of (2.173) does she employ, and how does she go about state reduction?

25 (a) Since you found Sect. 2.4.3 very puzzling, read it again.
(b) If you are still puzzled, repeat (a), else proceed with (c).
(c) Convince your favorite skeptical colleague that there can be different, but equally consistent, statistical operators for the same object.

2.5 Fifth Lecture: Statistics of Detected Atoms

A one-atom maser is operated under steady-state conditions. What is the probability to detect a $\overline{\bullet}$ atom between t and $t + \mathrm{d}t$? It is

$$r\mathrm{d}t\,\eta_{\overline{\bullet}}\ \mathrm{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\}\ , \tag{2.189}$$

the product of the probability $r\mathrm{d}t$ of having an atom in this time interval ($r > 0$ is taken for granted from here on) and the probability $\eta_{\overline{\bullet}}\ \mathrm{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\}$ that the $\overline{\bullet}$ detector clicks if there is an atom. Similarly,

$$r\mathrm{d}t\,\eta_{\underline{\bullet}}\ \mathrm{tr}\left\{\mathcal{B}\rho^{(\mathrm{ss})}\right\} \tag{2.190}$$

is the probability for a $\underline{\bullet}$ click between t and $t + \mathrm{d}t$. What multiplies $r\mathrm{d}t$ here are the traces of the first and the second term of (2.165), respectively. The third term gave us the no-click probability (2.166).

The a priori probabilities (2.189) and (2.190) make no reference to other detection events. But the detector clicks are not statistically independent. For the example of the parity measurements of Sect. 2.4.3, you can see this very clearly in Fig. 2.6d, where the even-parity clicks and the odd-parity clicks come in bunches. This bunching is easily understood: a $\overline{\bullet}$ click is accompanied by the state reduction (2.181) so that ρ_t is an even-parity state immediately after a $\overline{\bullet}$ click and, therefore, the next atom is much more likely to encounter even parity than odd parity. The detection of the first atom conditions the probabilities for the second atom, and for all subsequent ones as well.

2.5.1 Correlation Functions

Perhaps the simplest question one can ask in this context is: Given that a $\overline{\bullet}$ atom was detected at $t = 0$, what is now the probability to detect a $\underline{\bullet}$ atom between t and $t + \mathrm{d}t$? Since detection events that occur earlier than t are not relevant here, we must use (2.173) with $\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} = 0$ to propagate $\rho_{t=0}$ to ρ_t,

$$\rho_t = \mathrm{e}^{\mathcal{L}_0 t} \rho_{t=0} \tag{2.191}$$

where

$$\mathcal{L}_0 = \mathcal{L}_\eta \Big|_{\eta_{\overline{\bullet}} = \eta_{\underline{\bullet}} = 0} = \mathcal{L} + r(\mathcal{A} + \mathcal{B} - 1) \tag{2.192}$$

is the Liouville operator of the Scully–Lamb equation that determines the steady state,

$$\mathcal{L}_0 \rho^{(\mathrm{ss})} = 0 \,, \qquad \lim_{t \to \infty} \mathrm{e}^{\mathcal{L}_0 t} \rho_{t=0} = \rho^{(\mathrm{ss})} \,. \tag{2.193}$$

The denominator of (2.175) equals unity here because $\mathcal{L}_0$ conserves the trace,

$$\mathrm{tr} \left\{ \mathrm{e}^{\mathcal{L}_0 t} \rho_{t=0} \right\} = \mathrm{tr} \left\{ \rho_{t=0} \right\} = 1 \,. \tag{2.194}$$

With

$$\rho_{t=0} = \frac{\mathcal{A} \rho^{(\mathrm{ss})}}{\mathrm{tr} \left\{ \mathcal{A} \rho^{(\mathrm{ss})} \right\}} \tag{2.195}$$

accounting for the initial $\overline{\bullet}$ click at $t = 0$ (which, under steady-state conditions, is really any instant), the probability for a $\underline{\bullet}$ click at $t \cdots t + \mathrm{d}t$ is then

$$r \mathrm{d}t \, \eta_{\underline{\bullet}} \, \mathrm{tr} \left\{ \mathcal{B} \, \mathrm{e}^{\mathcal{L}_0 t} \rho_{t=0} \right\} = r \mathrm{d}t \, \frac{\mathrm{tr} \left\{ \mathcal{B} \, \mathrm{e}^{\mathcal{L}_0 t} \mathcal{A} \rho^{(\mathrm{ss})} \right\}}{\mathrm{tr} \left\{ \mathcal{A} \rho^{(\mathrm{ss})} \right\}} \,. \tag{2.196}$$

If t is very late, the memory of the initial state gets lost and (2.196) becomes equal to the a priori probability (2.190). The ratio of the two probabilities,

$$G_{\overline{\bullet}\underline{\bullet}}(t) = \frac{\mathrm{tr} \left\{ \mathcal{B} \, \mathrm{e}^{\mathcal{L}_0 t} \mathcal{A} \rho^{(\mathrm{ss})} \right\}}{\mathrm{tr} \left\{ \mathcal{B} \rho^{(\mathrm{ss})} \right\} \, \mathrm{tr} \left\{ \mathcal{A} \rho^{(\mathrm{ss})} \right\}} \,, \tag{2.197}$$

is the correlation function for $\underline{\bullet}$ clicks after $\overline{\bullet}$ clicks. There are no correlations if $G_{\overline{\bullet}\underline{\bullet}} = 1$, positive correlations if $G_{\overline{\bullet}\underline{\bullet}} > 1$, negative correlations if $G_{\overline{\bullet}\underline{\bullet}} < 1$. The bunching of Fig. 2.6d shows strong negative $\overline{\bullet}\underline{\bullet}$ correlations (odd after even) at short times.

Likewise, if we ask about $\overline{\bullet}$ clicks after $\underline{\bullet}$ clicks, we get

$$G_{\underline{\bullet}\overline{\bullet}}(t) = \frac{\mathrm{tr} \left\{ \mathcal{A} \, \mathrm{e}^{\mathcal{L}_0 t} \mathcal{B} \rho^{(\mathrm{ss})} \right\}}{\mathrm{tr} \left\{ \mathcal{A} \rho^{(\mathrm{ss})} \right\} \, \mathrm{tr} \left\{ \mathcal{B} \rho^{(\mathrm{ss})} \right\}} \,, \tag{2.198}$$

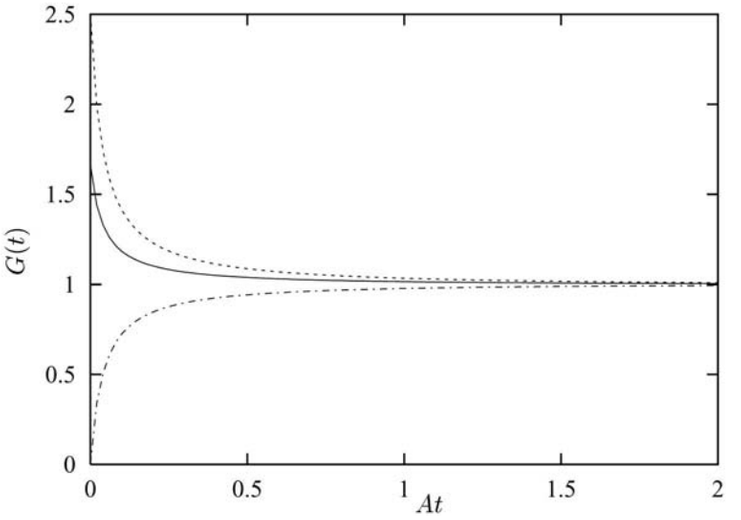


Fig. 2.8. Correlation functions for parity measurements on an unpumped resonator. The solid —— line shows the even-even correlation function $G_{\overline{\bullet}\,\overline{\bullet}}(t)$, the dashed $- - -$ line the odd-odd correlation function $G_{\underline{\bullet}\,\underline{\bullet}}(t)$, and the dash-dotted $- \cdot - \cdot -$ line the cross correlations $G_{\overline{\bullet}\,\underline{\bullet}}(t) = G_{\underline{\bullet}\,\overline{\bullet}}(t)$. The plot is for $\nu = 2$ in the expressions given in (2.204) and (2.205)

and we have

$$G_{\overline{\bullet}\,\overline{\bullet}}(t) = \frac{\operatorname{tr}\left\{\mathcal{A}\,\mathrm{e}^{\mathcal{L}_0 t}\mathcal{A}\rho^{(\mathrm{ss})}\right\}}{\left[\operatorname{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\}\right]^2}\,, \qquad G_{\underline{\bullet}\,\underline{\bullet}}(t) = \frac{\operatorname{tr}\left\{\mathcal{B}\,\mathrm{e}^{\mathcal{L}_0 t}\mathcal{B}\rho^{(\mathrm{ss})}\right\}}{\left[\operatorname{tr}\left\{\mathcal{B}\rho^{(\mathrm{ss})}\right\}\right]^2}\,, \tag{2.199}$$

for the correlation functions of clicks of the same kind. Note that the detection efficiencies $\eta_{\overline{\bullet}}, \eta_{\underline{\bullet}}$ do not enter the correlation functions (2.197)–(2.199). This is a result of normalizing the conditional probabilities to the unconditional ones. But see homework assignment 26.

As an example we take once more the parity measurements of Figs. 2.6 and 2.7, for which $\mathcal{A}, \mathcal{B}$ are given in (2.177) so that $\mathcal{L}_0 = \mathcal{L}$. Then $\rho^{(\mathrm{ss})}$ is the thermal state of (2.38) and the a priori probabilities are

$$\begin{aligned}\operatorname{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\} &= \frac{1}{2} + \frac{1}{2}\left\langle(-1)^{a^\dagger a}\right\rangle^{(\mathrm{ss})} = \frac{\nu+1}{2\nu+1}\,,\\ \operatorname{tr}\left\{\mathcal{B}\rho^{(\mathrm{ss})}\right\} &= \frac{1}{2} - \frac{1}{2}\left\langle(-1)^{a^\dagger a}\right\rangle^{(\mathrm{ss})} = \frac{\nu}{2\nu+1}\,.\end{aligned} \tag{2.200}$$

For $\nu = 0$, there are no odd-parity clicks of the $\underline{\bullet}$ detector and, therefore, we will only consider the case of $\nu > 0$. Since

$$\left.\begin{matrix}\mathcal{A}\rho^{(\mathrm{ss})}\\ \mathcal{B}\rho^{(\mathrm{ss})}\end{matrix}\right\} = \frac{1}{2}\rho^{(\mathrm{ss})} \pm \frac{1}{2}\frac{1}{\nu+1}\left(-\frac{\nu}{\nu+1}\right)^{a^\dagger a}\,, \tag{2.201}$$

we have

$$\mathrm{e}^{\mathcal{L}_0 t}(\mathcal{A}+\mathcal{B})\rho^{(\mathrm{ss})} = \rho^{(\mathrm{ss})} \tag{2.202}$$

and

$$\mathrm{e}^{\mathcal{L}_0 t}(\mathcal{A}-\mathcal{B})\rho^{(\mathrm{ss})} = \frac{1}{2\nu+1}\lambda(t)\big[1-\lambda(t)\big]^{a^\dagger a}\,, \tag{2.203}$$

where $\lambda(t)$ is given in (2.49) with $\lambda(0) = (2\nu+1)/(\nu+1)$. The numerators of (2.197)–(2.199) are then easily evaluated and we obtain

$$G_{\bullet\,\overline{\bullet}}(t) = G_{\overline{\bullet}\,\bullet}(t) = 1 - \Big[1+(2\nu+1)^2(\,\mathrm{e}^{At}-1)\Big]^{-1} \tag{2.204}$$

and

$$\begin{aligned} G_{\overline{\bullet}\,\overline{\bullet}}(t) &= 1 + \frac{\nu}{\nu+1}\Big[1+(2\nu+1)^2(\,\mathrm{e}^{At}-1)\Big]^{-1}\,, \\ G_{\bullet\,\bullet}(t) &= 1 + \frac{\nu+1}{\nu}\Big[1+(2\nu+1)^2(\,\mathrm{e}^{At}-1)\Big]^{-1}\,. \end{aligned} \tag{2.205}$$

The cross correlations of (2.204) vanish at $t = 0$ (no even-parity click immediately after an odd-parity click and vice versa) and increase monotonically toward 1. The two same-click correlation functions of (2.205) are always larger than 1 and decrease monotonically from their $t = 0$ values

$$G_{\overline{\bullet}\,\overline{\bullet}}(t=0) = \frac{2\nu+1}{\nu+1}\,, \qquad G_{\bullet\,\bullet}(t=0) = \frac{2\nu+1}{\nu}\,, \tag{2.206}$$

which exceed unity and so confirm the bunching observed in Fig. 2.6d. For $\nu = 2$, the parameter value of Figs. 2.6 and 2.7, the correlation functions are plotted in Fig. 2.8.

Data of actual measurements of correlation functions for atoms emerging from a real-life micromaser are reported in [44], for example. Theoretical values for some related quantities, such as the mean number of successive detector clicks of the same kind, agree very well with the experimental findings.

2.5.2 Waiting Time Statistics

Here is a different question: A $\bullet$ click happened at $t = 0$, what is the probability that the next $\overline{\bullet}$ clicks occurs between t and $t + \mathrm{d}t$? In marked contrast to the question asked before (2.191), we are now not interested in any later click, but in the *next* click, and this just says that there are no other $\overline{\bullet}$ clicks before t. Since we ignore deliberately all $\bullet$ clicks at intermediate times, we have to use (2.173) with $\eta_{\bullet} = 0$ in the click operator $\mathcal{C}$ of (2.168).

We introduce the following quantities:

$$\begin{aligned} r\mathrm{d}t\, p_{\overline{\bullet}}(t) &= \text{probability for a } \overline{\bullet} \text{ click at } t\cdots t+\mathrm{d}t\,, \\ p_{\mathrm{no}\overline{\bullet}}(t) &= \text{probability for no } \overline{\bullet} \text{ click before } t\,, \\ \mathrm{d}t\, P_{\mathrm{next}\overline{\bullet}}(t) &= \text{probability for the next } \overline{\bullet} \text{ click to happen at } t\cdots t+\mathrm{d}t\,. \end{aligned} \tag{2.207}$$

Since $1 - r\mathrm{d}t\, p_{\overline{\bullet}}(t)$ is then the probability that there is no $\overline{\bullet}$ click between t and $t + \mathrm{d}t$, we have

$$p_{\mathrm{no}\overline{\bullet}}(t + \mathrm{d}t) = p_{\mathrm{no}\overline{\bullet}}(t)\big[1 - r\mathrm{d}t\, p_{\overline{\bullet}}(t)\big]$$

$$\text{or}\quad p_{\overline{\bullet}}(t) = -\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}\ln p_{\mathrm{no}\overline{\bullet}}(t)\;, \tag{2.208}$$

and

$$p_{\mathrm{no}\overline{\bullet}}(t + \mathrm{d}t) = p_{\mathrm{no}\overline{\bullet}}(t) - \mathrm{d}t\, P_{\mathrm{next}\overline{\bullet}}(t)$$

$$\text{or}\quad P_{\mathrm{next}\overline{\bullet}}(t) = -\frac{\mathrm{d}}{\mathrm{d}t}p_{\mathrm{no}\overline{\bullet}}(t) \tag{2.209}$$

is another immediate consequence of the significance given to these quantities.

We know $p_{\overline{\bullet}}(t)$ from (2.166) and (2.175),

$$p_{\overline{\bullet}}(t) = \frac{\mathrm{tr}\left\{\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}}{\mathrm{tr}\left\{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}} \tag{2.210}$$

where

$$\mathcal{C} = \eta_{\overline{\bullet}}\mathcal{A} \qquad \text{and} \qquad \rho_{t=0} = \frac{\mathcal{B}\rho^{(\mathrm{ss})}}{\mathrm{tr}\left\{\mathcal{B}\rho^{(\mathrm{ss})}\right\}} \tag{2.211}$$

in the present context. As required by (2.208), the right-hand side of (2.210) must be a logarithmic derivative and, indeed, it is because the identity

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{tr}\left\{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\} &= \mathrm{tr}\left\{\mathcal{L}_\eta\,\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\\ &= \mathrm{tr}\left\{(\mathcal{L}_0 - r\mathcal{C})\,\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\\ &= -r\,\mathrm{tr}\left\{\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\;,\end{aligned} \tag{2.212}$$

which uses the trace-conserving property of $\mathcal{L}_0$, implies

$$p_{\overline{\bullet}}(t) = -\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}\ln\,\mathrm{tr}\left\{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\;. \tag{2.213}$$

It follows that

$$p_{\mathrm{no}\overline{\bullet}}(t) = \mathrm{tr}\left\{\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\;, \tag{2.214}$$

and then (2.209), (2.212) give

$$P_{\mathrm{next}\overline{\bullet}}(t) = r\,\mathrm{tr}\left\{\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t}\rho_{t=0}\right\}\;. \tag{2.215}$$

The "important physical significance" of the denominator in (2.173) that was left in limbo in Sect. 2.4.2 is finally revealed in (2.214): it is the probability that

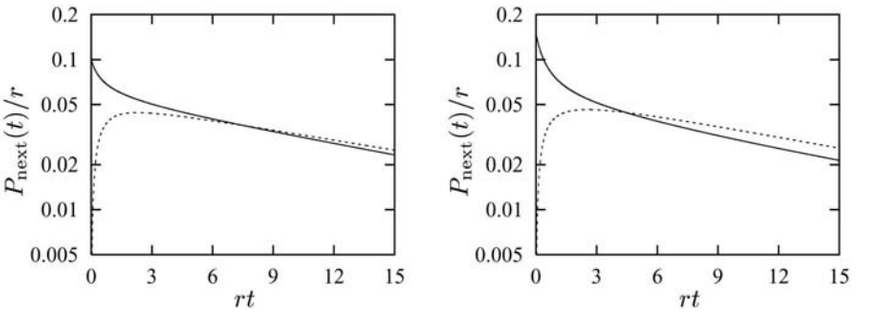


Fig. 2.9. Waiting time distributions for parity measurements on an unpumped resonator. The solid —— lines show $P_{\rm next}(t)/r$ for $\overline{\bullet}$ clicks after $\overline{\bullet}$ clicks (left) and $\underline{\bullet}$ clicks after $\underline{\bullet}$ clicks (right) as functions of rt. The dashed − − − lines refer to $\underline{\bullet}$ clicks after $\overline{\bullet}$ clicks (left) and $\overline{\bullet}$ clicks after $\underline{\bullet}$ clicks (right). In these logarithmic plots, straight lines would correspond to the Poissonian statistics of uncorrelated clicks

no atom is detected before t. Since an atom is surely detected if we just wait long enough, the limit

$$p_{\rm no\,\overline{\bullet}}(t) \to 0 \quad \text{as } t \to \infty \tag{2.216}$$

is a necessary property of $p_{\rm no\,\overline{\bullet}}(t)$. As a consequence, all eigenvalues of $\mathcal{L}_\eta$ must have a negative real part.

Putting all things together we obtain

$$P_{\rm next\,\overline{\bullet}}(t) = r\eta_{\overline{\bullet}}\,\frac{\mathrm{tr}\left\{\mathcal{A}\,\mathrm{e}^{(\mathcal{L}_0 - r\eta_{\underline{\bullet}}\mathcal{A})t}\mathcal{B}\rho^{(\rm ss)}\right\}}{\mathrm{tr}\left\{\mathcal{B}\rho^{(\rm ss)}\right\}} \tag{2.217}$$

for the waiting time distribution for the next $\overline{\bullet}$ click after a $\underline{\bullet}$ click. Analogous expressions apply for the next $\underline{\bullet}$ click after a $\overline{\bullet}$ click, the next $\overline{\bullet}$ click after a $\overline{\bullet}$ click, and so forth. As a basic check of consistency we consider the situation in which $\mathcal{A}$ and $\mathcal{B}$ are just multiples of the identity,

$$\mathcal{A}\rho_t = q\rho_t\,, \qquad \mathcal{B}\rho_t = (1-q)\rho_t \qquad \text{with} \quad 0 < q < 1\,. \tag{2.218}$$

Then the detector clicks are not correlated at all and the waiting time distribution (2.217) should be Poissonian,

$$P_{\rm next\,\overline{\bullet}}(t) = r\eta_{\overline{\bullet}}\,q\,\mathrm{e}^{-r\eta_{\overline{\bullet}}qt}\,, \tag{2.219}$$

and this is indeed what we get from (2.217) for (2.218).

Figure 2.9 shows the waiting time distributions to the parity measurements of Figs. 2.6–2.8. Other examples are presented in some figures of [39].

2.5.3 Counting Statistics

Yet another question is this: What is the probability $w_n(t)$ for detecting n atoms in state $\overline{\bullet}$ during a period of duration t? We pay no attention to $\underline{\bullet}$ clicks and,

therefore, have $\eta_{\underline{\bullet}} = 0$ in the nonlinear master equation (2.173) for the evolution between $\overline{\bullet}$ clicks.

Probability $w_0(t)$ is the no-click probability of Sect. 2.5.2,

$$w_0 = \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t}\rho^{(\mathrm{ss})}\right\} \qquad \text{with} \qquad \mathcal{C} = \eta_{\overline{\bullet}}\mathcal{A}\,, \tag{2.220}$$

where $\rho_{t=0} = \rho^{(\mathrm{ss})}$ in (2.214) is appropriate now. The one-click probability $w_1(t)$ is given by

$$w_1(t) = r\int_0^t \mathrm{d}t'\ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta (t-t')}\rho_{t=t'+0}\right\} \operatorname{tr}\left\{\mathcal{C}\rho_{t=t'-0}\right\} \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}\right\}\,, \tag{2.221}$$

where the last factor is the probability $w_0(t')$ for no click before t', the first factor is the probability for no click after t', and

$$r\mathrm{d}t'\ \operatorname{tr}\left\{\mathcal{C}\rho_{t=t'-0}\right\} \tag{2.222}$$

is the probability for a click at t'. In accordance with (2.175) and (2.181), the statistical operators just before and after the click at t' are

$$\rho_{t=t'-0} = \frac{\mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}}{\operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}\right\}} \tag{2.223}$$

and

$$\rho_{t=t'+0} = \frac{\mathcal{A}\rho_{t=t'-0}}{\operatorname{tr}\left\{\mathcal{A}\rho_{t=t'-0}\right\}} = \frac{\mathcal{C}\rho_{t=t'-0}}{\operatorname{tr}\left\{\mathcal{C}\rho_{t=t'-0}\right\}}\,, \tag{2.224}$$

respectively, so that

$$\rho_{t=t'+0}\ \operatorname{tr}\left\{\mathcal{C}\rho_{t=t'-0}\right\} \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}\right\} = \mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}\,, \tag{2.225}$$

and a remarkable simplification happens, inasmuch as

$$w_1(t) = \int_0^t \mathrm{d}t'\ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta (t-t')} r\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t'}\rho^{(\mathrm{ss})}\right\} \tag{2.226}$$

involves but a single trace as the equivalent replacement of the product of three traces with which we started in (2.221).

Upon writing (2.226) as a double integral

$$w_1(t) = \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_0\,\delta(t_0+t_1-t)\ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t_1} r\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t_0}\rho^{(\mathrm{ss})}\right\} \tag{2.227}$$

it is reasonably obvious (and can be demonstrated by a simple induction) that

$$\begin{aligned} w_n(t) = \int_0^\infty \mathrm{d}t_n \int_0^\infty \mathrm{d}t_{n-1} \cdots \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_0\,\delta(t_0+t_1+\cdots+t_{n-1}+t_n-t) \\ \times\ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_\eta t_n} r\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t_{n-1}} r\mathcal{C}\cdots\mathrm{e}^{\mathcal{L}_\eta t_1} r\mathcal{C}\,\mathrm{e}^{\mathcal{L}_\eta t_0}\rho^{(\mathrm{ss})}\right\} \end{aligned} \tag{2.228}$$

or

$$w_n(t) = \mathrm{tr}\left\{\mathcal{W}_n(t)\rho^{(\mathrm{ss})}\right\} . \tag{2.229}$$

The operator $\mathcal{W}_n(t)$ thus introduced,

$$\mathcal{W}_n(t) = \int_0^\infty \mathrm{d}t_n \cdots \int_0^\infty \mathrm{d}t_0\, \delta(t_0 + \cdots + t_n - t)\, \mathrm{e}^{\mathcal{L}_\eta t_n} r\mathcal{C}\, \mathrm{e}^{\mathcal{L}_\eta t_{n-1}} \cdots r\mathcal{C}\, \mathrm{e}^{\mathcal{L}_\eta t_0} , \tag{2.230}$$

obeys the recurrence relation

$$\mathcal{W}_n(t) = \int_0^t \mathrm{d}t'\, \mathcal{W}_0(t-t')\, r\mathcal{C}\, \mathcal{W}_{n-1}(t') \qquad \text{for } n = 1, 2, 3, \ldots \tag{2.231}$$

that commences with

$$\mathcal{W}_0(t) = \mathrm{e}^{\mathcal{L}_\eta t} . \tag{2.232}$$

As always, we will find it convenient to use a generating function,

$$\mathcal{W}(x,t) = \sum_{n=0}^\infty x^n \mathcal{W}_n(t) . \tag{2.233}$$

The recurrence relation for the $\mathcal{W}_n(t)$'s then turns into an integral equation for $\mathcal{W}(x,t)$,

$$\mathcal{W}(x,t) = \mathcal{W}_0(t) + \int_0^t \mathrm{d}t'\, \mathcal{W}_0(t-t')\, xr\mathcal{C}\, \mathcal{W}(x,t') . \tag{2.234}$$

The standard technique for handling such finite-range convolutions utilizes Laplace transforms because the familiar identity

$$\int_0^\infty \mathrm{d}t\ \mathrm{e}^{-\gamma t} \int_0^t \mathrm{d}t'\, f(t-t')g(t') = \int_0^\infty \mathrm{d}t\ \mathrm{e}^{-\gamma t} f(t) \int_0^\infty \mathrm{d}t'\ \mathrm{e}^{-\gamma t'} g(t') \tag{2.235}$$

leads to a factorization. With

$$\int_0^\infty \mathrm{d}t\ \mathrm{e}^{-\gamma t} \mathcal{W}_0(t) = \big(\gamma - \mathcal{L}_\eta\big)^{-1} \tag{2.236}$$

this gets us to

$$\begin{aligned} \int_0^\infty \mathrm{d}t\ \mathrm{e}^{-\gamma t} \mathcal{W}(x,t) &= \big[1 - \big(\gamma - \mathcal{L}_\eta\big)^{-1} xr\mathcal{C}\big]^{-1} \big(\gamma - \mathcal{L}_\eta\big)^{-1} \\ &= \Big(\big(\gamma - \mathcal{L}_\eta\big)\big[1 - \big(\gamma - \mathcal{L}_\eta\big)^{-1} xr\mathcal{C}\big]\Big)^{-1} \\ &= \big(\gamma - \mathcal{L}_\eta - xr\mathcal{C}\big)^{-1} , \end{aligned} \tag{2.237}$$

and the inverse Laplace transform is elementary,

$$\mathcal{W}(x,t) = \mathrm{e}^{(\mathcal{L}_\eta + xr\mathcal{C})t} . \tag{2.238}$$

As we noted at (2.216), all eigenvalues of $\mathcal{L}_\eta$ have negative real parts, and so the Laplace transform (2.236) of $\mathcal{W}_0(t)$ surely exists for $\gamma \geq 0$. The same remark applies to $\mathcal{W}(x,t)$ for $-(1-\eta_{\overline{\bullet}})/\eta_{\overline{\bullet}} < x < 1$ because

$$\mathcal{L}_\eta + xr\mathcal{C} = \mathcal{L}_0 - (1-x)r\mathcal{C} = \mathcal{L}_0 - r(1-x)\eta_{\overline{\bullet}}\mathcal{A} \tag{2.239}$$

is $\mathcal{L}_\eta$ with $\eta_{\overline{\bullet}}$ replaced by $(1-x)\eta_{\overline{\bullet}}$ so that $\mathcal{L}_\eta + xr\mathcal{C}$ is just another operator of the same kind as $\mathcal{L}_\eta$ if the "effective detection efficiency" $(1-x)\eta_{\overline{\bullet}}$ is in the range $0\cdots 1$, which restricts x to the range stated above.

After putting things together we obtain

$$\sum_{n=0}^{\infty} x^n w_n(t) = \operatorname{tr}\left\{ \mathrm{e}^{(\mathcal{L}_\eta + xr\mathcal{C})t}\rho^{(\mathrm{ss})}\right\} \tag{2.240}$$

as the generating function for the counting probabilities $w_n(t)$. In view of what we observed above about $\mathcal{L}_\eta + xr\mathcal{C}$, the right-hand side of (2.240) is equal to the no-click probability $w_0(t)$ for detection efficiency $(1-x)\eta_{\overline{\bullet}}$. Accordingly, $w_0(t)$ determines all $w_n(t)$ through its dependence on $\eta_{\overline{\bullet}}$. As an immediate consequence of this observation, we get a statement about the moments of the counting statistics,

$$\sum_{n=0}^{\infty}\binom{n}{k} w_n(t) = \frac{1}{k!}\left(\frac{\partial}{\partial x}\right)^k \sum_{n=0}^{\infty} x^n w_n(t)\Big|_{x=1} = \eta_{\overline{\bullet}}^k \left[\eta_{\overline{\bullet}}^{-k} w_k(t)\Big|_{\eta_{\overline{\bullet}}\to 0}\right] . \tag{2.241}$$

For $k=0$, we check the normalization,

$$\sum_{n=0}^{\infty} w_n(t) = \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_0 t}\rho^{(\mathrm{ss})}\right\} = 1 ; \tag{2.242}$$

for $k=1$, we get the average number of $\overline{\bullet}$ clicks,

$$\begin{aligned}\sum_{n=0}^{\infty} n w_n(t) &= \int_0^t \mathrm{d}t' \ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_0(t-t')} r\eta_{\overline{\bullet}}\mathcal{A}\,\mathrm{e}^{\mathcal{L}_0 t'}\rho^{(\mathrm{ss})}\right\} \\ &= \eta_{\overline{\bullet}} rt \ \operatorname{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\} ,\end{aligned} \tag{2.243}$$

which can be understood as a statement about the ergodicity of the process [39]; and for $k=2$, we learn something about the variance of the counting statistics,

$$\begin{aligned}&\sum_{n=0}^{\infty} n(n-1) w_n(t) \\ &\quad = 2\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \ \operatorname{tr}\left\{ \mathrm{e}^{\mathcal{L}_0(t-t')} r\eta_{\overline{\bullet}}\mathcal{A}\,\mathrm{e}^{\mathcal{L}_0(t'-t'')} r\eta_{\overline{\bullet}}\mathcal{A}\,\mathrm{e}^{\mathcal{L}_0 t''}\rho^{(\mathrm{ss})}\right\} \\ &\quad = (\eta_{\overline{\bullet}} rt)^2 \ \operatorname{tr}\left\{\mathcal{A}E(\mathcal{L}_0 t)\mathcal{A}\rho^{(\mathrm{ss})}\right\} ,\end{aligned} \tag{2.244}$$

with

$$E(y) = \frac{2}{y^2}\big(\,\mathrm{e}^{y} - 1 - y\big)\ . \tag{2.245}$$

In traces of integrals such as (2.243) and (2.244), the exponential on the left and on the right can be ignored because $\mathcal{L}_0$ is trace conserving and $\rho^{(\mathrm{ss})}$ is its right eigenvector to eigenvalue zero.

Here, too, we get Poissonian statistics for (2.218), namely

$$\sum_{n=0}^{\infty} x^n w_n(t) = \mathrm{e}^{-(1-x)r\eta_{\overset{\bullet}{-}}qt}\ , \qquad w_n(t) = \frac{(r\eta_{\overset{\bullet}{-}}qt)^n}{n!}\,\mathrm{e}^{-r\eta_{\overset{\bullet}{-}}qt}\ , \tag{2.246}$$

for which

$$\sum_{n=0}^{\infty}\binom{n}{k} w_n(t) = \frac{(r\eta_{\overset{\bullet}{-}}qt)^k}{k!}\ . \tag{2.247}$$

In particular, we note that

$$\sum_{n=0}^{\infty} n(n-1)w_n(t) = \left[\sum_{n=0}^{\infty} n w_n(t)\right]^2 \tag{2.248}$$

holds for the Poissonian counting statistics (2.246).

A convenient, yet rough, measure for the deviation from Poissonian statistics is the so-called Fano–Mandel factor $Q(t)$,

$$Q = \frac{\displaystyle\sum_{n=0}^{\infty} n(n-1)w_n}{\displaystyle\sum_{n=0}^{\infty} n w_n} - \sum_{n=0}^{\infty} n w_n\ , \tag{2.249}$$

a normalized variance. The normalization is such that $Q = 0$ for Poissonian counting statistics, as one verifies easily with (2.248). For $-1 \le Q < 0$ one speaks of sub-Poissonian statistics, and of super-Poissonian statistics for $Q > 0$.

For the count of $\overset{\bullet}{-}$ clicks, (2.243) and (2.244) give

$$Q_{\overset{\bullet}{-}}(t) = \eta_{\overset{\bullet}{-}}rt\left[\frac{\mathrm{tr}\left\{\mathcal{A}E(\mathcal{L}_0 t)\mathcal{A}\rho^{(\mathrm{ss})}\right\}}{\mathrm{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\}} - \mathrm{tr}\left\{\mathcal{A}\rho^{(\mathrm{ss})}\right\}\right]\ . \tag{2.250}$$

We use the damping bases to get a tractable numerical expression,

$$Q_{\overset{\bullet}{-}}(t) = \eta_{\overset{\bullet}{-}}rt\sum_{n=0}^{\infty}\sum_{k=-\infty}^{\infty}\big(1-\delta_{n,0}\delta_{k,0}\big)\frac{\mathrm{tr}\left\{\mathcal{A}\rho_n^{(k)}\right\}E(\lambda_n^{(k)}t)\ \mathrm{tr}\left\{\check{\rho}_n^{(k)}\mathcal{A}\rho_0^{(0)}\right\}}{\mathrm{tr}\left\{\mathcal{A}\rho_0^{(0)}\right\}}\ , \tag{2.251}$$

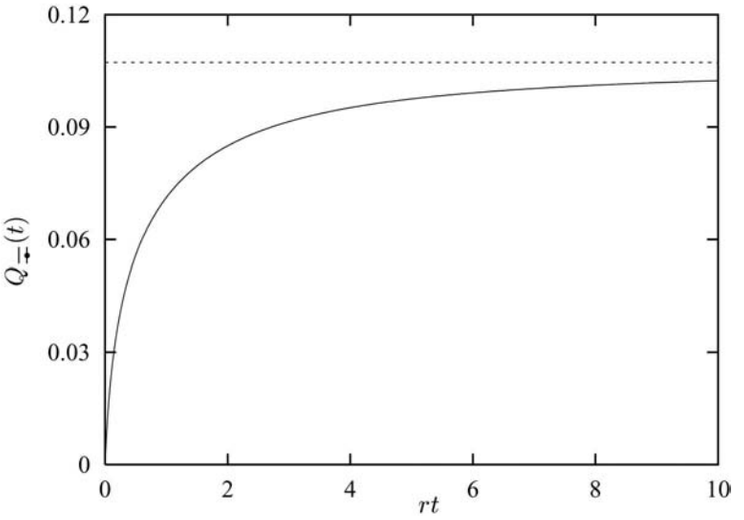


Fig. 2.10. Fano–Mandel factor for the counting statistics in parity measurements on an unpumped resonator. The solid —— line shows $Q_{\bullet\!-}(t)$ of (2.252), the horizontal dashed − − − line is the asymptotic value of (2.253)

where the $n = 0, k = 0$ term is removed since $\rho_0^{(0)} = \rho^{(\mathrm{ss})}$, $\breve{\rho}_0^{(0)} = 1$, $\lambda_0^{(0)} = 0$, and $E(y = 0) = 1$. For the parity measurements of Figs. 2.6–2.9, one can evaluate the sum and gets

$$Q_{\bullet\!-}(t) = \frac{\eta_{\bullet\!-} r}{(\nu+1)(2\nu+1)} \int_0^t \frac{\mathrm{d}t'}{2At} \ln\Bigl(1 + 4\nu(\nu+1)\bigl(1 - \mathrm{e}^{-At'}\bigr)\Bigr) \,, \tag{2.252}$$

which is plotted in Fig. 2.10. For $At \ll 1$ and $At \gg 1$ the limiting forms

$$Q_{\bullet\!-}(t) \simeq \begin{cases} \dfrac{\nu}{2\nu+1}\eta_{\bullet\!-} rt & \text{for } At \ll 1 \,, \\ \dfrac{\ln(2\nu+1)}{(\nu+1)(2\nu+1)}\eta_{\bullet\!-} r/A & \text{for } At \gg 1 \,, \end{cases} \tag{2.253}$$

obtain, as is confirmed by Fig. 2.10. Other examples of Fano–Mandel factors for counting statistics are presented in some figures of [39].

The pioneering measurement in 1990 of atom statistics in a real-life micromaser experiment is reported in [45] and linked to the photon counting statistics in [46]. Measured Fano–Mandel factors from some later experiments can be found in [44].

Homework Assignments

26 What is the correlation function for clicks of either kind, that is: without caring if it's a $\bullet\!-$ click or a $-\!\bullet$ click?

27 The two cross-correlation functions are identical in the example of (2.200)–(2.206), see (2.204). Is this always the case?

28 Since the next click is bound to come sooner or later, consistency requires that $P_{\rm next}(t)$ is normalized to unit integral. Show that this is indeed so for $P_{\rm next}(t)$ of (2.215).

29 A $\overline{-\bullet}$ click happens at $t = 0$. What is the probability that the next click is a $\underline{-\bullet}$ click?

30 Use the methods of Sect. 2.5.2 to find an expression for the average waiting time between successive $\overline{-\bullet}$ clicks, between successive $\underline{-\bullet}$ clicks.

31 Determine the short-time behavior of the various $P_{\rm next}$'s of Fig. 2.9 and compare with the plots.

32 Show that the exponential function e^F of an operator F responds to variations δF in accordance with

$$\delta\,\mathrm{e}^F = \int_0^1 \mathrm{d}\tau\ \mathrm{e}^{\tau F}\delta F\,\mathrm{e}^{(1-\tau)F}\ , \tag{2.254}$$

which epitomizes all of perturbation theory.

33 Use (2.254) to extract $w_1(t)$ and $w_2(t)$ from (2.240). Compare with (2.228).

34 Consider the probability $w_{nm}(t)$ of detecting n atoms in state $\overline{-\bullet}$ and m atoms in state $\underline{-\bullet}$ during a period of duration t. Show that

$$\sum_{n,m=0}^{\infty} x^n y^m w_{nm}(t) = \mathrm{tr}\left\{\mathrm{e}^{(\mathcal{L}_\eta + xr\eta_{\overline{-\bullet}}\mathcal{A} + yr\eta_{\underline{-\bullet}}\mathcal{B})t}\rho^{(\mathrm{ss})}\right\} \tag{2.255}$$

is the appropriate generalization of the generating function (2.240).

35 For $\mathcal{A}$ of (2.177), evaluate the $k = 0$ traces in (2.251) and so confirm (2.252).

36 Find the leading correction to the approximation given in (2.253) for $At \gg 1$.

Acknowledgments

We congratulate the organizers of the CohEvol seminar and workshop in Dresden. They succeeded in putting together a stimulating meeting in a most enjoyable atmosphere.

Much of the work reported in these lectures was done at the Max-Planck-Institut für Quantenoptik in Garching. BGE is greatly indebted to Herbert Walther for his generous hospitality and support over all those years.

The evolution of the damping-basis method from a crude idea to a powerful tool would not have occurred without the ingenuity and stamina of Hans Briegel. BGE will always remember this collaboration with great affection.

At the time of the CohEvol seminar and workshop, BGE enjoyed the hospitality of the Atominstitut in Vienna. He wishes to thank Helmut Rauch and Gerald Badurek for the splendid environment they provided, and the Technical University of Vienna for financial support. BGE is equally grateful to Goong Chen and Marlan Scully for the visiting professorship they arranged at Texas A&M University, where these notes were finalized.

GM wishes to express her sincere gratitude for the strong encouragement and support by Herbert Walther and the members of his Garching group. And she thanks all participants of the workshop for the many memorable hours spent together.

Appendix

Here are some facts about special functions that are useful for homework assignments 6 and 9. The expansion of (2.73) in powers of $\mathrm{e}^{-\mathrm{i}\omega t} \propto \big[\alpha(t)/\alpha^*(t)\big]^{\frac{1}{2}}$ is done with the aid of

$$\mathrm{e}^{\frac{1}{2}x(y-1/y)} = \sum_{k=-\infty}^{\infty} y^k \mathrm{J}_k(x)\;, \tag{2.256}$$

the most important generating function for Bessel functions of integer order,

$$\mathrm{J}_k(x) = (-1)^k \mathrm{J}_{-k}(x) = \mathrm{i}^{|k|-k}\mathrm{J}_{|k|}(x) = \mathrm{i}^{|k|-k}\sum_{m=0}^{\infty}\frac{(-1)^m}{m!\,(m+|k|)!}\,\big(\tfrac{1}{2}x\big)^{|k|+2m}\;. \tag{2.257}$$

They in turn act as a generating function for Laguerre polynomials,

$$\mathrm{J}_{|k|}\big(2\sqrt{xy}\,\big) = (xy)^{\frac{1}{2}|k|}\,\mathrm{e}^{-y}\sum_{n=0}^{\infty}\frac{y^n}{(n+|k|)!}\mathrm{L}_n^{(|k|)}(x)\;. \tag{2.258}$$

After a suitable Laplace transform this becomes

$$(1+y)^{-|k|-1}\,\mathrm{e}^{\frac{xy}{1+y}} = \sum_{n=0}^{\infty}(-y)^n\mathrm{L}_n^{(|k|)}(x)\;, \tag{2.259}$$

which is another useful generating function for the Laguerre polynomials

$$\mathrm{L}_n^{(|k|)}(x) = \sum_{m=0}^{n}\binom{n+|k|}{m+|k|}\frac{(-x)^m}{m!}\;. \tag{2.260}$$

The integral relations

$$\mathrm{L}_n^{(|k|)}(x) = \frac{1}{n!}\,\mathrm{e}^x x^{-\frac{1}{2}|k|}\int_0^{\infty}\mathrm{d}y\;\mathrm{e}^{-y}y^{n\,+\,\frac{1}{2}|k|}\mathrm{J}_{|k|}\big(2\sqrt{xy}\,\big) \tag{2.261}$$

and

$$\int_0^{\infty}\mathrm{d}y\;\mathrm{e}^{-y}\mathrm{J}_{|k|}\big(2\sqrt{uy}\,\big)\mathrm{J}_{|k|}\big(2\sqrt{vy}\,\big) = \mathrm{e}^{-(u\,+\,v)}\mathrm{I}_k\big(2\sqrt{uv}\,\big) \tag{2.262}$$

are worth knowing, where

$$\mathrm{I}_k(x) = \mathrm{I}_{-k}(x) = \mathrm{i}^{-k}\mathrm{J}_k(\mathrm{i}x) \tag{2.263}$$

are modified Bessel functions of integer order. As a preparation for homework assignment 6 you might want to derive first

$$\sum_{n=0}^{\infty} \frac{n!}{(n+|k|)!} x^n \mathrm{L}_n^{(|k|)}(y) \mathrm{L}_n^{(|k|)}(z) = \frac{(xyz)^{-\frac{1}{2}|k|}}{1-x} \mathrm{e}^{-\frac{x}{1-x}(y+z)} \mathrm{I}_k\left(2\frac{\sqrt{xyz}}{1-x}\right) \tag{2.264}$$

by combining some of these relations fittingly.

References

1. J. R. Klauder, E. C. G. Sudarshan: *Fundamentals of Quantum Optics* (W. A. Benjamin, New York 1970)
2. R. Loudon: *The Quantum Theory of Light* (Oxford University Press, New York 1973)
3. W. H. Louisell: *Quantum Statistical Properties of Radiation* (John Wiley, New York 1973)
4. H. M. Nussenzveig: *Introduction to Quantum Optics* (Gordon and Breach, New York 1974)
5. M. Sargent III, M. O. Scully, W. E. Lamb Jr.: *Laser Physics* (Addison-Wesley, Reading 1974)
6. L. Allen, J. H. Eberly: *Optical Resonance and Two-Level Atoms* (John Wiley, New York 1975)
7. H. Haken: *Light*, Vols. I and II (North-Holland, Amsterdam 1981)
8. P. L. Knight, L. Allen: *Concepts of Quantum Optics* (Pergamon Press, Oxford 1983)
9. P. Meystre, M. Sargent III: *Elements of Quantum Optics* (Springer, Berlin 1990)
10. H. Carmichael: *An Open System Approach to Quantum Optics* (Springer, Heidelberg 1991)
11. W. Vogel, D.-G. Welsch: *Lectures on Quantum Optics* (Akademie Verlag, Berlin 1994)
12. D. F. Walls, G. J. Milburn: *Quantum Optics* (Springer, Heidelberg 1994)
13. M. O. Scully, M. S. Zubairy: *Quantum Optics* (Cambridge University Press, Cambridge 1997)
14. C.W. Gardiner, P. Zoller: *Quantum Noise* (Springer, Heidelberg 2000)
15. W. P. Schleich: *Quantum Optics in Phase Space* (Wiley-VCH, Berlin 2001)
16. A. S. Holevo: Rep. Math. Phys. **33**, 95 (1993)
17. A. S. Holevo: J. Funct. Anal. **131**, 255 (1995)
18. B. Vacchini: *Quantum optical versus quantum Brownian motion master-equation in terms of covariance and equilibrium properties*, Preprint arXiv:quant-ph/0204071 (2002)
19. B.-G. Englert: *Elements of Micromaser Physics*, Preprint arXiv:quant-ph/0203052 (2002)
20. H.-J. Briegel, B.-G. Englert: Phys. Rev. A **47**, 3311 (1993); errata in [37]
21. B.-G. Englert, M. Naraschewski, A. Schenzle: Phys. Rev. A **50**, 2667 (1994); the argument of the Laguerre polynomial has the wrong sign in (2.17)
22. B.-G. Englert: J. Phys. A **22**, 625 (1989)
23. A. Royer: Phys. Rev. A **15**, 449 (1977)
24. A. Grossmann: Commun. Math. Phys. **48**, 191 (1976)

25. S. M. Barnett, S. Stenholm: J. Mod. Opt. **47**, 2869 (2000)
26. H. Risken: *The Fokker–Planck Equation*, 2nd edn. (Springer, Heidelberg 1989)
27. G. Lindblad: Commun. Math. Phys. **48**, 119 (1976)
28. R. Alicki, K. Lendi: *Quantum Dynamical Semigroups and Applications* (Springer, Heidelberg 1987)
29. A. S. Holevo: *Statistical Structure of Quantum Theory* (Springer, Berlin 2001)
30. G. Raithel, C. Wagner, H. Walther, L. M. Narducci, M. O. Scully: 'The Micromaser: A Proving Ground for Quantum Physics'. In: *Cavity Quantum Electrodynamics*, ed. by P. R. Berman (Academic Press, New York 1994) pp. 57–121
31. S. Haroche, J.-M. Raimond: 'Manipulation of Non-Classical Field States in a Cavity by Atomic Interferometry'. In: *Cavity Quantum Electrodynamics*, ed. by P. R. Berman (Academic Press, New York 1994) pp. 123–170
32. H. Walther: Opt. Spectrosc. **91**, 327 (2001)
33. J.-M. Raimond, M. Brune, S. Haroche: Rev. Mod. Phys. **73**, 565 (2001)
34. H. Walther, B. T. H. Varcoe, B.-G. Englert: *Cavity Quantum Electrodynamics* (in preparation for Rep. Prog. Phys., to be submitted shortly)
35. B. T. H. Varcoe, S. Brattke, M. Weidinger, H. Walther: Nature (London) **403**, 743 (2000)
36. S. Brattke, B.-G. Englert, B. T. H. Varcoe, H. Walther: J. Mod. Opt. **47**, 2857 (2000)
37. H.-J. Briegel, B.-G. Englert: Phys. Rev. A **52**, 2361 (1995)
38. B.-G. Englert, N. Sterpi, H. Walther: Opt. Commun. **100**, 526 (1993)
39. H.-J. Briegel, B.-G. Englert, N. Sterpi, H. Walther: Phys. Rev. A **49**, 2962 (1994)
40. B. Kümmerer, H. Maassen: *An Ergodic Theorem for Quantum Counting Processes*, LANL eprint quant-ph/0102134 (2001)
41. N. D. Mermin: 'Whose Knowledge?'. In: *Quantum (Un)speakables. From Bell to Quantum Information*, ed. by R. A. Bertlmann, A. Zeilinger (Springer, Berlin 2002) pp. 171–180
42. T. A. Brun, J. Finkelstein, N. D. Mermin: Phys. Rev. A **65**, 032315 (2002)
43. C. M. Caves, C. A. Fuchs, R. Schack: *Conditions for compatibility of quantum state assignments*, LANL eprint quant-ph/0206110 (2002)
44. B.-G. Englert, M. Löffler, O. Benson, B. Varcoe, M. Weidinger, H. Walther: Fortschr. Phys. **46**, 897 (1998)
45. G. Rempe, F. Schmidt–Kaler, H. Walther: Phys. Rev. Lett. **64**, 2783 (1990)
46. G. Rempe, H. Walther: Phys. Rev. A **42**, 1650 (1990)

3 Stochastic Resonance

Kurt Wiesenfeld[1], Thomas Wellens[2], and Andreas Buchleitner[2]

[1] Georgia Institute of Technology, Atlanta, GA 30332, USA
[2] Max-Planck-Institut für Physik komplexer Systeme, D-01187 Dresden, Germany

3.1 Introduction

In this chapter, we discuss a surprising discovery made in the 1980*s* known as stochastic resonance. It concerns a cooperative effect seen in certain nonlinear systems in the presence of random noise. The signature of stochastic resonance is that the coherence of the system output improves with an increase of random noise, at least over some range of noise levels. The subject has been vigorously studied over the past decade. Stochastic resonance is now known to occur in numerous examples spanning a wide range of physical systems. The lion's share of the research has focussed on systems that can be adequately described by classical physics, though some attention has been devoted to quantum mechanical realizations.

The purpose of this chapter is to develop the main features of the theory in a way which is accessible to the non-expert. There already exist general overviews of stochastic resonance [1,2]. These tell the interesting story of how the idea of stochastic resonance was originally proposed to explain why the Earth suffers Ice Ages with apparently near-periodic regularity, how that idea found its true application in phenomena as diverse as laser dynamics and predator-detection of the simple crayfish, and current efforts to see whether it could explain outstanding mysteries in human perception in the visual and auditory systems. There is also a comprehensive technical review article on stochastic resonance [3].

In contrast to these references, this chapter aims to present the theory of stochastic resonance at a quantitative level suitable for graduate students in the physical sciences. The goal is to provide a solid basis from which to explore the (by now rather large) technical literature on stochastic resonance. We also hope to convey the current frontiers of the subject and where open questions remain. Despite great progress, stochastic resonance remains an active and exciting area of research.

3.2 Some Mathematical Tools

The quantitative theory of stochastic resonance involves the study of fluctuating dynamical systems. The basic tools used to develop the theory are those of stochastic processes.

The first question is this: how do we incorporate randomness into a model? Here, we have in mind randomness which is separate from that which arises in

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 107–138, 2002.

quantum mechanics. The idea is that some aspect of the system is not under perfect control, and suffers fluctuations about its nominal value. For example, suppose we have a simple mass-spring system, which is subject to very complicated and highly erratic forces due to collisions with the surrounding air molecules. The equation of motion for the displacement x is

$$m\ddot{x} = -k\,(x - \ell) + \xi(t)\ , \tag{3.1}$$

where m is the mass, k is the spring constant, ℓ is the spring's unstretched length, and the overdot denotes differentiation with respect to time. The forcing function ξ is for all practical purposes random. How do we deal with this equation mathematically? We imagine that the function ξ is deterministic, but changes each time we run the experiment. We also suppose that we have statistical information about the ensemble of functions $\xi_1, \xi_2, \xi_3, ...$, and nothing more. Then we can imagine solving the sequence of problems

$$\begin{aligned} m\ddot{x}_1 &= -k\,(x_1 - \ell) + \xi_1(t)\ , \\ m\ddot{x}_2 &= -k\,(x_2 - \ell) + \xi_2(t)\ , \\ m\ddot{x}_3 &= -k\,(x_3 - \ell) + \xi_3(t)\ , \end{aligned} \tag{3.2}$$

and so on, generating an ensemble of outputs x_j. Since we have only statistical information about the ξ_j, we can only hope to recover statistical information about the x_j.

What kind of statistical information do we need to supply about ξ? Well, we typically specify information about how ξ is correlated with itself at different times,

$$\langle \xi(t) \rangle = \int \xi\ p(\xi, t)\ \mathrm{d}\xi, \tag{3.3}$$

$$\langle \xi(t_1)\xi(t_2) \rangle = \int\int \xi_1\xi_2\ p(\xi_1, t_1; \xi_2, t_2)\ \mathrm{d}\xi_1\mathrm{d}\xi_2, \tag{3.4}$$

$$\langle \xi(t_1)\xi(t_2)\xi(t_3) \rangle = \int\int\int \ldots \tag{3.5}$$

and so on. In the above, $p(\xi, t)$ stands for the probability density of ξ as a function of time, $p(\xi_1, t_1; \xi_2, t_2)$ is the joint probability density of ξ at two times, and so forth. The second of these plays a central role, and is called the *autocorrelation function*, or sometimes just the correlation function, for short.

We call the process ξ stationary if the moments depend only on the time-differences, so that

$$\begin{aligned} &\langle \xi(t) \rangle = \langle \xi(0) \rangle\ , \\ &\langle \xi(t_1)\xi(t_2) \rangle = \langle \xi(0)\xi(t_2 - t_1) \rangle\ , \\ &\langle \xi(t_1)\xi(t_2)\xi(t_3) \rangle = \langle \xi(0)\xi(t_2 - t_1)\xi(t_3 - t_1) \rangle\ , \\ &\ldots \end{aligned} \tag{3.6}$$

For physical reasons we often expect the random process ξ to be stationary.

An important quantity is the power spectrum, which gives information about a dynamical process in terms of its frequency content. If we consider a function $\xi(t)$ over the time interval $t \in (0, t_{\max})$, then we can form the Fourier transform

$$\tilde{\xi}(\Omega) = \int_0^{t_{\max}} \xi(t) \mathrm{e}^{-\mathrm{i}\Omega t} \mathrm{d}t \; . \tag{3.7}$$

The power spectrum is

$$S(\Omega) = \lim_{t_{\max}\to\infty} \frac{1}{2\pi t_{\max}} \left|\tilde{\xi}(\Omega)\right|^2 \; . \tag{3.8}$$

For stationary processes one can compute the power spectrum via the Wiener–Khintchine theorem,

$$S(\Omega) = 4 \int_0^{\infty} C(\tau) \cos(\Omega\tau) \mathrm{d}\tau, \tag{3.9}$$

where C is the autocorrelation function [4].

3.3 Example: Driven Linear System

Let's consider a simple example. Later, we'll use the results of this example to compare against a system which displays stochastic resonance. But for now we just want to illustrate the various tools we just introduced in a concrete example. Consider a linear system driven by noise and a periodic input, with governing equation

$$\dot{x} = -\gamma x + \xi + \varepsilon \cos(\omega t), \;\; x(0) = 0, \tag{3.10}$$

where ξ is a random function and γ, ε, and ω are constants. We need to specify the statistical properties of ξ. Let's take

$$\langle \xi(t) \rangle = 0 \; , \tag{3.11}$$

$$\langle \xi(t) \xi(t') \rangle = \kappa \delta(t - t') \; . \tag{3.12}$$

The parameter κ represents the strength of the fluctuations. By themselves, these two conditions don't uniquely specify the random process ξ, but for the quantities we're going to be interested in they are all we need. It is common to call a random process with these properties 'white noise'. The name refers to the power spectrum of ξ. The autocorrelation function of ξ depends on the time difference only, so we calculate the power spectrum using the Wiener–Khintchine theorem (3.9), with result

$$S(\Omega) = 2\kappa \; . \tag{3.13}$$

The noise is 'white' because it contains equal power in every frequency bin.

The solution to (3.10) is

$$x(t) = \mathrm{e}^{-\gamma t} \int_0^t \mathrm{e}^{\gamma t'} \left(\xi(t') + \varepsilon \cos \omega t'\right) \mathrm{d}t' \; . \tag{3.14}$$

The ensemble average of this is

$$\langle x(t)\rangle = \langle \mathrm{e}^{-\gamma t}\int_0^t \mathrm{e}^{\gamma t'}\xi(t')\,\mathrm{d}t'\rangle + \langle \mathrm{e}^{-\gamma t}\int_0^t \mathrm{e}^{\gamma t'}\varepsilon\cos(\omega t')\,\mathrm{d}t'\rangle\ . \tag{3.15}$$

In the first term, the angular brackets can be moved all the way in until they hit ξ since this is the only random quantity, and using (3.11) we conclude that this term vanishes. There are no random factors in the second term: the ensemble average has no effect and so the angular brackets are just dropped. Thus,

$$\langle x(t)\rangle = \frac{\varepsilon}{\sqrt{\gamma^2+\varepsilon^2}}\sin(\omega t+\phi) - \mathrm{e}^{-\gamma t}\frac{\varepsilon\sin\phi}{\sqrt{\gamma^2+\varepsilon^2}}\ , \tag{3.16}$$

where $\sin\phi = \gamma/\sqrt{\gamma^2+\varepsilon^2}$. After a transient time, the mean response is periodic with an amplitude which is independent of the noise strength.

Meanwhile, the correlation function $\langle x(t)x(t+\tau)\rangle$ consists of four terms. Of these, the two cross-terms drop out since each involves a single factor of ξ which therefore vanishes upon ensemble averaging. This leaves two terms:

$$\begin{aligned}\langle x(t)x(t+\tau)\rangle &= \mathrm{e}^{-\gamma(2t+\tau)}\int_0^t\int_0^{t+\tau}\mathrm{e}^{\gamma t'}\mathrm{e}^{\gamma t''}\langle\xi(t')\xi(t'')\rangle\mathrm{d}t''\mathrm{d}t' \\ &\quad+\varepsilon^2\mathrm{e}^{-\gamma(2t+\tau)}\int_0^t\int_0^{t+\tau}\mathrm{e}^{\gamma t'}\mathrm{e}^{\gamma t''}\cos(\omega t')\cos(\omega t'')\,\mathrm{d}t''\mathrm{d}t'\ .\end{aligned} \tag{3.17}$$

Evaluating the integrals yields

$$\begin{aligned}\langle x(t)x(t+\tau)\rangle &= \frac{\kappa}{2\gamma}\mathrm{e}^{-\gamma\tau}+\varepsilon^2\frac{\sin[\omega t+\phi]\sin[\omega(t+\tau)+\phi]}{\gamma^2+\omega^2} \\ &= \frac{\kappa}{2\gamma}\mathrm{e}^{-\gamma\tau}+\frac{1}{2}\frac{\varepsilon^2}{\gamma^2+\omega^2}\left[\cos\omega\tau+\cos(2\omega t+\omega\tau+2\phi)\right]\ ,\end{aligned} \tag{3.18}$$

where in the first equation we neglected the transient response by taking the limit $t \gg 1/\gamma$. Notice, however, that even after the transient has died away, *the correlation function is not stationary*. That is, the correlation function depends on both the initial time t and the time difference τ. On the other hand, the dependence on t is merely periodic, a consequence of the periodic forcing function.

It is certainly possible to measure and otherwise investigate the systematic variations in output that occur over one period of the drive [5] – the variance might be largest at a particular point in the drive cycle, for instance. However, this is relatively subtle information, and not of central interest in the study of stochastic resonance. Instead, it is convenient to make an additional time average (in t) over one period, and take

$$C(\tau) = \frac{\omega}{2\pi}\int_0^{2\pi/\omega}\langle x(t)x(t+\tau)\rangle\mathrm{d}t\ , \tag{3.19}$$

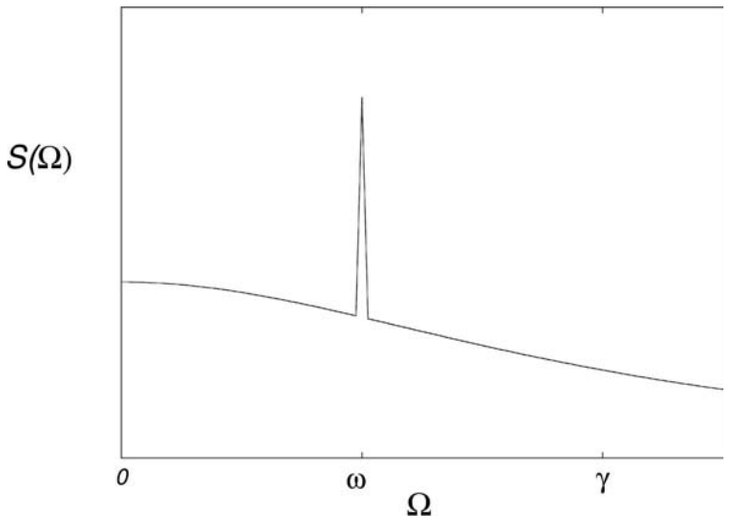


Fig. 3.1. Power spectrum for the linear system driven by both white noise and a periodic input, consisting of a sharp peak at the signal frequency ω, and a Lorentzian noise background, see (3.21)

where $2\pi/\omega$ is the driving period. As the notation indicates, the resulting correlation function depends only on the time difference τ,

$$C(\tau) = \frac{\kappa}{2\gamma}\mathrm{e}^{-\gamma\tau} + \frac{1}{2}\frac{\varepsilon^2}{\gamma^2+\omega^2}\cos\omega\tau \ . \tag{3.20}$$

Note that the periodic influence of the drive has not been destroyed by the extra averaging. The correlation function is now stationary, so we get the power spectrum by taking the Fourier transform:

$$S(\Omega) = \frac{2\kappa}{\gamma^2+\Omega^2} + \frac{\varepsilon^2}{\gamma^2+\omega^2}\delta\left(\Omega-\omega\right) \ . \tag{3.21}$$

The result is plotted in Fig. 3.1. We see that the power spectrum consists of a broadband part plus a narrow spike at the signal frequency. For low enough signal frequencies, the broadband spectrum is essentially flat. A useful measure of the output coherence is the signal-to-noise ratio (SNR). We divide the strength of the delta function spike at the signal frequency by the level of broadband noise at this same frequency. In our example, this gives

$$\mathrm{SNR} \ = \ \frac{\varepsilon^2}{\gamma^2+\omega^2}\frac{\gamma^2+\omega^2}{2\kappa} \ = \ \frac{\varepsilon^2}{2\kappa} \ . \tag{3.22}$$

It is worth making an additional technical point here. In experiments (or simulations), one typically integrates the power spectrum over some finite but small bandwidth $\Delta\Omega$ to determine both the output signal and noise powers. The bandwidth should be large enough to pick up all the power under the spike at the

signal frequency, but still small enough so that the broadband spectrum is flat over the interval. If we carry out these integrations for the present example, the signal part is unchanged but the noise part picks up a bandwidth factor, with result

$$\mathrm{SNR} \;=\; \frac{\varepsilon^2}{2\kappa\Delta\Omega}\;. \tag{3.23}$$

From a theoretical perspective, the bandwidth factor is essentially arbitrary, and we won't bother to include it.

Our primary interest is in the behavior of the SNR as a function of the input noise strength κ, and in this example we see that the SNR monotonically decreases with κ. Raising the noise level always degrades the quality of the output. You may say that this last statement is intuitively obvious, and that only an overeducated ninny would bother to demonstrate this fact by doing a calculation. But in fact this intuition can be wrong! This is what makes stochastic resonance such an intriguing phenomenon.

3.4 Mean First Passage Time and Kramers Escape Formula

The following example serves to illustrate certain aspects of the Fokker–Planck approach to analyzing stochastic processes and also presents an important result we will use later. We begin with a nonlinear stochastic problem

$$\dot{x} = -\frac{dV}{dx} + \xi\;, \tag{3.24}$$

where the potential function V has the shape illustrated in Fig. 3.2.

We ask: if a particle is initially placed at the lefthand minimum x_1, how long will it take for the particle to reach the righthand minimum x_3 for the first time? The answer will vary from realization to realization, and we call the expectation value the 'mean first passage time'. In the limit of weak noise, the result is

$$\langle \mathcal{T} \rangle = \frac{2\pi}{\sqrt{-V''(x_1)V''(x_2)}}\; \mathrm{e}^{2\Delta V/\kappa}\;, \tag{3.25}$$

where the primes denote differentiation with respect to x and ΔV is the energy difference between the potential maximum at x_2 and minimum at x_1. This is the justly famous Kramers formula. It is widely (not to say universally!) applicable because, as we can see, the details of the potential are irrelevant. Only the energy barrier and the curvatures at the minimum and maximum enter the formula. For a derivation of this, see Gardiner's book [6]. Here is the basic idea. Consider the evolution of the probability density. Initially, it is a delta function, $p(x,0) = \delta(x - x_1)$, and its subsequent time evolution $p(x,t)$ follows a partial differential equation, called the Fokker–Planck equation [6]. Since we are interested in the very first instant a particle reaches x_3, we solve the Fokker–Planck equation with an absorbing boundary at x_3. The solution represents the

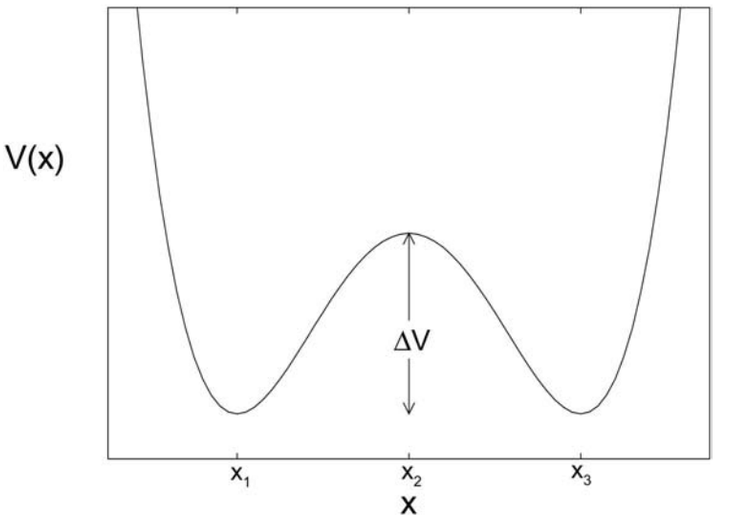


Fig. 3.2. Effective particle potential for Kramers' escape under the governing equation (3.24). The potential exhibits stable equilibria at x_1 and x_3, and an unstable equilibrium at x_2, on top of the potential barrier of height ΔV

conditional probability that the particle is at position x at time t given that it started at position x_1. Therefore, the probability that the particle has escaped after time t is

$$p_{\rm esc}(t;x_1) = 1 - \int_{-\infty}^{x_3} {\rm d}x \; p_{\rm cond}(x,t|x_1,0) \; . \tag{3.26}$$

As time passes, $p_{\rm esc}$ increases monotonically from zero toward one. The first passage time distribution is just the derivative

$$\mathcal{T}(t;x_1 \to x_3) = \frac{\partial}{\partial t} p_{\rm esc} \; , \tag{3.27}$$

and the mean time is

$$\langle \mathcal{T} \rangle = \int_0^\infty t \; \mathcal{T}(t;x_1 \to x_3) {\rm d}t \; . \tag{3.28}$$

This last result can be written directly in terms of the potential function V:

$$\langle \mathcal{T} \rangle = \alpha \int_{x_1}^{x_3} {\rm d}y \; {\rm e}^{\alpha V(y)} \int_{-\infty}^{y} {\rm d}z \; {\rm e}^{-\alpha V(z)} \; , \tag{3.29}$$

where $\alpha = 2/\kappa$. For weak noise, κ is small, so α is large, and the integrals can be evaluated in terms of the extremal values of V, which leads to the relatively detail-free form of the Kramers formula.

3.5 Rate Equation Description

The Fokker–Planck equation governs the time evolution of the probability distribution, but we will be able to use a simplified description for one of the theories of stochastic resonance developed in the next section. The simpler version is called the rate equation description, and reduces the continuous phase space to a discrete set of attractors. For a damped particle in a bistable potential there are two stable states, and we make the approximation that the system is either in one state or the other. This is reasonable if the relaxation times are short compared with the typical transition times between the states. The probability density is then written

$$p(x,t) \approx n_+(t)\delta(x-x_+) + n_-(t)\delta(x-x_-) \ , \tag{3.30}$$

where, for example

$$n_+ = \int_{x_2}^{\infty} p(x,t)\mathrm{d}x \ , \tag{3.31}$$

and similarly for $n_-(t)$. If the separation of time scales is obeyed, then the precise location of x that divides the phase space is unimportant, though the logical place for it is at the potential maximum x_2. We also imagine that the transition rates between states are specified. These rates may be given by the Kramers formula, or they may be determined in some other way (including possibly experimental measurements). The dynamics is then governed by the linear ordinary differential equations

$$\begin{aligned} \dot{n}_+ &= W_{\mathrm{up}} n_- - W_{\mathrm{down}} n_+ \ , \\ \dot{n}_- &= W_{\mathrm{down}} n_+ - W_{\mathrm{up}} n_- \ . \end{aligned} \tag{3.32}$$

These can be reduced to a single equation since $n_+ + n_- = 1$.

3.6 Two State Theory of Stochastic Resonance

We now have the tools assembled to develop the theory of two state stochastic resonance. More sophisticated treatments can be found [3,7], but the simple theory here has the virtue that it is quite general, and is the one most often used in the literature. Its chief disadvantage is that the transition rates are assumed to be given.

The basic situation is illustrated in Fig. 3.3. There are two stable states, with the system making transitions between them. To fix ideas, you can imagine a heavily damped particle moving in a double well potential subject to noise. We also include a time periodic influence which asymmetrically alters the rates, favoring first transitions 'up' and then a half-period later favoring transitions 'down'. This amounts to rocking the potential back and forth as shown in Fig. 3.3.

The goal of the theory is to calculate the output signal to noise ratio (SNR) as a function of the input noise strength. The theory splits into two parts. First,

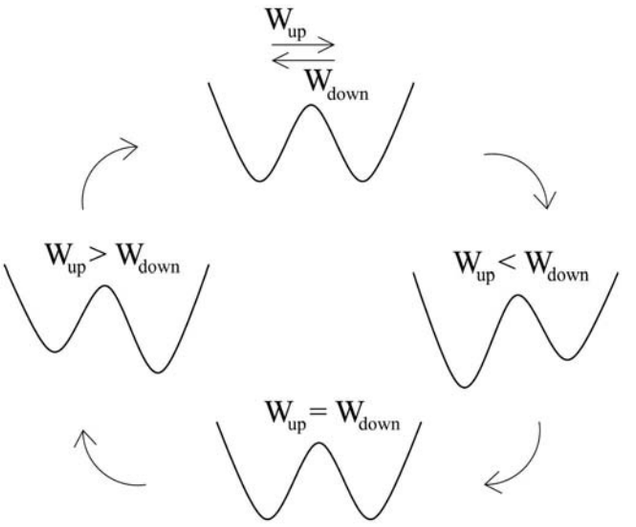


Fig. 3.3. Modulated potential function for two state SR. The periodic driving force which is rocking the bistable potential back and forth passes its extrema at one fourth (middle right) and three fourths (middle left) of the driving field period. This induces a periodic modulation of the transition rates (3.34)

we suppose that the transition rates are known, and we calculate the SNR in terms of these rates. This part is very general. Second, we relate the rates to the various system parameters. This part depends on the details of the system under consideration.

From (3.32) and the condition $n_+ + n_- = 1$, we have

$$\dot{n}_+ = -\left(W_{\rm up} + W_{\rm down}\right) n_+ + W_{\rm up} \ . \tag{3.33}$$

We assume time dependent transition rates

$$\begin{aligned} W_{\rm up} \quad &= W_0 + \varepsilon \cos \omega t \ , \\ W_{\rm down} &= W_0 - \varepsilon \cos \omega t \ . \end{aligned} \tag{3.34}$$

As time passes the output x switches between two values, say $x = \pm c$ (to be identified with x_1 and x_3 in Fig. 3.1).

3.7 The Unmodulated Case

It is instructive to consider first the unmodulated case ($\varepsilon = 0$). This allows us to see the overall structure of the calculation without getting bogged down in complicated expressions.

We can immediately integrate (3.33)

$$n_+(t) = {\rm e}^{-2W_0(t-t_0)} \left[n_+(t_0) - \frac{1}{2} \right] + \frac{1}{2} \ , \tag{3.35}$$

where $n_+(t_0)$ is the initial condition. Note that this determines:

1 the *conditional* probabilities

$$\begin{aligned} p_c(x=c,t|x=c,t_0) &= \frac{1}{2}\mathrm{e}^{-2W_0(t-t_0)}+\frac{1}{2}\,, \\ p_c(x=c,t|x=-c,t_0) &= -\frac{1}{2}\mathrm{e}^{-2W_0(t-t_0)}+\frac{1}{2}\,, \end{aligned} \tag{3.36}$$

where we set $n_+(t_0)=1$ in the first of these, and $n_+(t_0)=0$ in the second. In this notation, the first of these is read 'the conditional probability that $x=c$ at time t given that $x=c$ at the initial time t_0'.

Equation (3.35) also determines

2 the *equilibrium* probabilities

$$\begin{aligned} p_{\mathrm{eq}}(x=c) &= n_+(t_0\to-\infty)=\frac{1}{2}\,, \\ p_{\mathrm{eq}}(x=-c) &= 1-n_+(t_0\to-\infty)=\frac{1}{2}\,. \end{aligned} \tag{3.37}$$

From these, we can easily compute the correlation function.

In general, when the state variable can take on a continuous range of values, the correlation function is given by

$$\begin{aligned} \langle x(t)x(t+\tau)\rangle &= \int\int xx'\; p_{\mathrm{joint}}(x,t;x',t+\tau)\;\mathrm{d}x\,\mathrm{d}x' \\ &= \int\int xx'\; p_c(x',t+\tau|x,t)p_{\mathrm{eq}}(x,t)\;\mathrm{d}x\,\mathrm{d}x'\,, \end{aligned} \tag{3.38}$$

where p_{joint} denotes the joint probability to find the particle in state x at time t, and in state x' at time $t+\tau$. For our two state system, the double integral collapses to a discrete sum of four terms:

$$\begin{aligned} \langle x(t)x(t+\tau)\rangle = {} & c^2p_c(c,t+\tau|c,t)p_{\mathrm{eq}}(c,t)+(c)(-c)p_c(c,t+\tau|-c,t)p_{\mathrm{eq}}(-c,t) \\ & +(-c)(c)p_c(-c,t+\tau|c,t)p_{\mathrm{eq}}(c,t) \\ & +(-c)^2p_c(-c,t+\tau|-c,t)p_{\mathrm{eq}}(-c,t). \end{aligned} \tag{3.39}$$

Collecting the various terms, and substituting into this last equation, we find the very simple result

$$\langle x(t)x(t+\tau)\rangle = c^2\mathrm{e}^{-2W_0\tau}\,, \tag{3.40}$$

and so (with (3.9))

$$\begin{aligned} S(\Omega) &= 4\int_0^\infty \langle x(t)x(t+\tau)\rangle\cos\Omega\tau\mathrm{d}\tau \\ &= 4c^2\frac{2W_0}{\Omega^2+4W_0^2}\,. \end{aligned} \tag{3.41}$$

The power spectrum is a Lorentzian with half-width $2W_0$.

3.8 Time Dependent Rates

Now put the modulation back in ($\varepsilon \neq 0$). The rate equation (3.33) reads

$$\dot{n}_+ = -2W_0 n_+ + W_0 + \varepsilon \cos \omega t \ , \tag{3.42}$$

with solution

$$n_+(t) = \mathrm{e}^{-2W_0(t-t_0)} \left[n_+(t_0) - \frac{1}{2} - \frac{\varepsilon \cos(\omega t_0 - \phi)}{\sqrt{4W_0^2 + \omega^2}} \right] + \frac{1}{2} + \frac{\varepsilon \cos(\omega t_0 - \phi)}{\sqrt{4W_0^2 + \omega^2}} \ . \tag{3.43}$$

This is nearly the same as (3.35), except for two periodic terms due to the modulation. The equilibrium probabilities are

$$\begin{aligned} p_{\mathrm{eq}}(c,t) \quad &= n_+(t_0 \to -\infty) = \frac{1}{2} \frac{\varepsilon \cos(\omega t_0 - \phi)}{\sqrt{4W_0^2 + \omega^2}} \ , \\ p_{\mathrm{eq}}(-c,t) &= 1 - p_{\mathrm{eq}}(c,t) \ . \end{aligned} \tag{3.44}$$

The conditional probabilities are also determined by (3.43). For example, consider the probability that $x(t+\tau) = c$ given that $x(t) = c$. One simply sets $n_+(t_0) = 1$ and $t - t_0 = \tau$ in (3.43). If we denote this process of substitution by double brackets, we have

$$p_c(c, t+\tau | c, t) = [[n_+(t_0) = 1; t - t_0 = \tau]] \ . \tag{3.45}$$

Similarly,

$$p_c(c, t+\tau | -c, t) = [[n_+(t_0) = 0; t - t_0 = \tau]] \ . \tag{3.46}$$

The two other conditional probabilities we need are simply constructed from the first two:

$$\begin{aligned} p_c(-c, t+\tau | c, t) \quad &= 1 - p_c(c, t+\tau | c, t) \ ; \\ p_c(-c, t+\tau | -c, t) &= 1 - p_c(c, t+\tau | -c, t) \ . \end{aligned} \tag{3.47}$$

Armed with the equilibrium and conditional probabilities, it is a straightforward (if tedious) exercise to construct the correlation function using (3.39), with result

$$\begin{aligned} \frac{1}{c^2} \langle x(t) x(t+\tau) \rangle = \mathrm{e}^{-2W_0 \tau} &- 4\mathrm{e}^{-2W_0 \tau} \frac{\varepsilon^2 \cos^2(\omega t - \phi)}{4W_0^2 + \omega^2} \\ &+ \frac{4\varepsilon^2}{4W_0^2 + \omega^2} \cos(\omega t - \phi) \cos(\omega t + \omega \tau - \phi) \ . \end{aligned} \tag{3.48}$$

As expected, the periodic modulation leads to a correlation function that depends on both t and τ. We do an additional phase averaging (like (3.19))

$$C(\tau) = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} \langle x(t) x(t+\tau) \rangle \mathrm{d}t \ , \tag{3.49}$$

and since the time average of $\cos^2(\omega t - \phi) = 1/2$, and the time average of $\cos(\omega t - \phi)\cos(\omega t + \omega\tau - \phi) = (1/2)\cos\omega\tau$, we find

$$C(\tau) = \left\{ c^2 \left(1 - \frac{2\varepsilon^2}{4W_0^2 + \omega^2} \right) \right\} \mathrm{e}^{-2W_0\tau} + \left\{ \frac{2c^2\varepsilon^2}{4W_0^2 + \omega^2} \right\} \cos\omega\tau \ . \tag{3.50}$$

Use the Wiener–Khintchine theorem (3.9) to determine the power spectrum:

$$\begin{aligned} S(\Omega) &= 4 \int_0^\infty C(\tau)\cos(\Omega\tau)\mathrm{d}\tau \\ &= 4c^2 \left(1 - \frac{2\varepsilon^2}{4W_0^2 + \omega^2} \right) \frac{2W_0}{4W_0^2 + \Omega^2} + \frac{4\pi c^2\varepsilon^2}{4W_0^2 + \omega^2} \delta(\Omega - \omega) \ . \end{aligned} \tag{3.51}$$

The first term gives the broadband part of the power output. Note that it *diminishes* with increasing ε in such a way that the total power is conserved: $\int_0^\infty S(\Omega)\mathrm{d}\Omega = \text{const}$.

From $S(\Omega)$, we can write down the signal-to-noise ratio SNR. Just divide the coefficient of the delta function by the value of the broadband term at $\Omega = \omega$. The resulting expression simplifies to

$$\mathrm{SNR} = \frac{\pi\varepsilon^2}{2W_0} \left(1 - \frac{2\varepsilon^2}{4W_0^2 + \omega^2} \right)^{-1} \ . \tag{3.52}$$

For small modulations, $\varepsilon \ll 1$, and so

$$\mathrm{SNR} \approx \frac{\pi\varepsilon^2}{2W_0} \ . \tag{3.53}$$

This is the main result of the first part of the theory. What we need now is to express ε and W_0 in terms of the specific system parameters, which means we need a theory for the transition rates W_{up} and W_{down}. This part of the theory depends on the details of the particular system. We consider here a popular example, the so-called overdamped particle in a double well potential. The dynamics is governed by the Langevin equation

$$\dot{x} = +\gamma x - x^3 + a\cos\omega t + \xi \ , \tag{3.54}$$

where ξ is white noise. This is an example of (3.24) described earlier during the discussion of the Kramers escape formula, and in fact we will use that formula to determine the transition rates. The corresponding potential $V(x,t)$ is

$$V(x,t) = -\frac{1}{2}\gamma x^2 + \frac{1}{4}x^4 - ax\cos\omega t \tag{3.55}$$

which when $a = 0$ describes a symmetric bistable system, and when $a \neq 0$ is periodically tilted in an antisymmetric fashion (see Fig. 3.3). We assume that the modulation strength a is sufficiently weak that the system remains in the high-barrier/low-noise limit throughout, in which case the Kramers formula holds.

Actually, the problem is a little more subtle since we are considering a time-dependent potential, while the Kramers formula is derived for the time independent case. It remains valid, however, if the modulation frequency is sufficiently low, which is sometimes referred to as the *adiabatic limit* of the time dependent problem.

We equate the transition rates to the reciprocal of the mean first passage time (3.25), so that

$$W = \frac{1}{2\pi}\sqrt{V''_{\min}|V''_{\max}|}\exp\left(-\frac{2}{\kappa}\Delta V\right) . \tag{3.56}$$

Consider first the unmodulated case $a = 0$. Then $V'' = 3x^2 - \gamma$, the potential minima lie at $x_{1,3} = \pm\sqrt{\gamma}$, the potential maximum is at $x_2 = 0$, and so the two barrier heights are equal and given by $\Delta V = \gamma^2/4$. Putting this into the rate formula yields

$$W = \frac{\gamma}{\sqrt{2}\pi}\mathrm{e}^{-\gamma^2/(2\kappa)} . \tag{3.57}$$

The two rates are equal because the unmodulated potential is symmetric. Now consider the modulated case, with $a \neq 0$ but small. Then

$$x_{\min} = \pm\sqrt{\gamma} + \text{correction} , \tag{3.58}$$

$$x_{\max} = 0 + \text{correction} ; \tag{3.59}$$

$$\Delta V \approx V(0) - V(x^2 = \gamma) = \frac{1}{4}\gamma^2 \pm a\sqrt{\gamma}\cos\omega t , \tag{3.60}$$

so that the rate formula becomes

$$W(t) = \frac{\gamma}{\sqrt{2}\pi}\exp\left\{-\frac{2}{\kappa}\left(\frac{1}{4}\gamma^2 \pm a\sqrt{\gamma}\cos\omega t\right)\right\} \tag{3.61}$$

$$= \frac{\gamma}{\sqrt{2}\pi}\mathrm{e}^{-\gamma^2/(2\kappa)}\exp\left\{\mp a\frac{2\sqrt{\gamma}}{\kappa}\cos\omega t\right\} . \tag{3.62}$$

If we expand the last exponential factor for small a, and equate $W(t)$ to the form assumed in the rate equation part (3.34) of the theory we get the required expressions

$$W_0 = \frac{\gamma}{\sqrt{2}\pi}\mathrm{e}^{-\gamma^2/2\kappa} , \tag{3.63}$$

$$\varepsilon = \frac{a\sqrt{2}\gamma^{3/2}}{\pi\kappa}\mathrm{e}^{-\gamma^2/2\kappa} . \tag{3.64}$$

This is the main result of the second part of the theory. We can put this together with the general expression for the signal-to-noise ratio. Using the simpler form (3.53), the result reduces to

$$\mathrm{SNR} = \frac{\sqrt{2}a^2\gamma^2}{\kappa^2}\mathrm{e}^{-\gamma^2/(2\kappa)} . \tag{3.65}$$

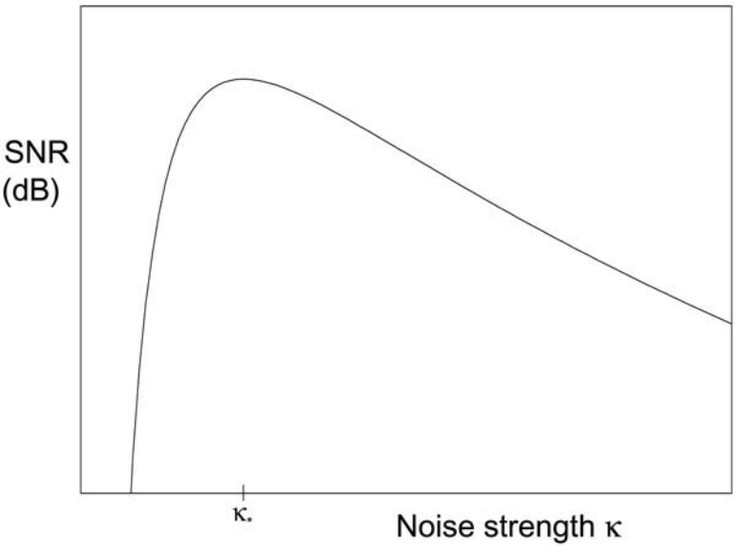


Fig. 3.4. Signal to noise ratio (SNR) vs. input noise strength for the two state theory. By virtue of (3.65), the SNR exhibits a maximum at an optimal, non-vanishing value κ_* of the noise strength κ. This is the characteristic feature of stochastic resonance. SNR is measured in 'decibels'(in short dB), defined as SNR (dB) $= 10 \log_{10}$ SNR

This is the final result of the two state theory. Figure 3.4 shows a semilog plot of the SNR, as a function of noise strength κ. The function in (3.65) is very flat and goes to zero in the noise-free limit (so the logarithm goes to $-\infty$ as $\kappa \to 0$), and the SNR decays as $1/\kappa^2$ for large noise. Of course, the Kramers formula breaks down for large noise, so the derived expression in this limit can't be taken seriously, although we expect it to be qualitatively right since the SNR must fall off (one would think!) for sufficiently high κ. In any event, the exciting result is that the SNR passes through a maximum at an intermediate noise level κ_*. From the formula, we readily find that $\kappa_* = \gamma^2/4 = \Delta V$, where ΔV is the barrier height in the unmodulated limit ($a = 0$).

Stochastic resonance has turned out to be very easy to see experimentally. However, there is one competing effect which is often encountered which may diminish or even mask entirely the maximum in Fig. 3.4. This is the so-called intrawell motion, a name borrowed from the double-well potential picture and which refers to the fact that the potential minima will in general vary with the time-dependent signal. This means that the output of the system is not constant between switching events; rather, it oscillates with small amplitude at the signal frequency, which of course contributes a small amount of output power at the signal frequency. In the limit of small noise, the power contained in the switching events is very small, and the residual power due to the intrawell oscillations is no longer negligible. We can get a quantitative measure of this effect by studying the dynamics restricted to one of the potential minima (since the oscillations

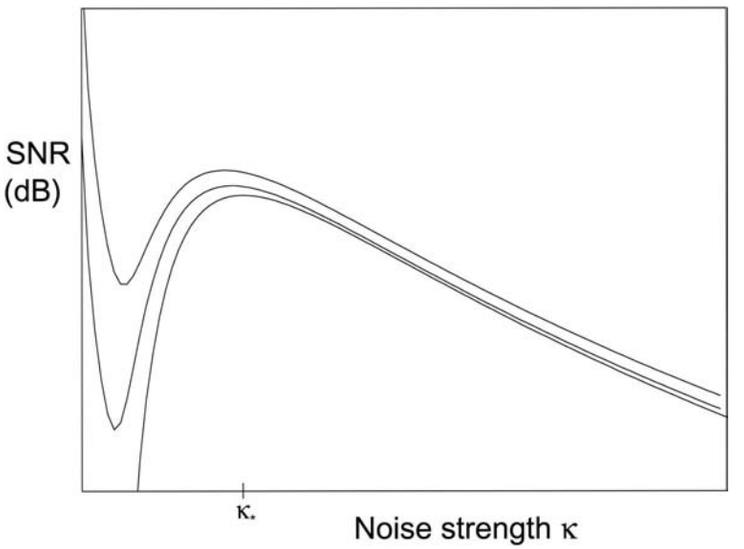


Fig. 3.5. Effect of intrawell motion on the output SNR. As the relative amount of intrawell motion increases, its contribution increasingly masks the underlying stochastic resonance behavior by increasing the SNR value at small values of the noise intensity κ (i.e., below κ_*)

are small in the important limit). We already did this calculation earlier in the chapter, when we studied as an example the dynamics described by (3.10). We found that

$$\mathrm{SNR}_{\mathrm{intra}} = \frac{\varepsilon^2}{2\kappa} \,, \tag{3.66}$$

which diverges at zero noise strength and is otherwise monotonically decreasing with κ. Whether or not this contribution washes out the maximum in expression (3.65) depends on how 'soft' the potential is about the two minima. Said another way, the more nearly the output follows a strict two-state waveform, the less overlap there is between the intrawell and interwell contributions to the SNR (see Fig. 3.5). If you read through the literature on stochastic resonance carefully, you will find that in some experiments the data has been filtered [8] to make it strictly two-state before the power spectrum is computed. In other cases, the stochastic resonance effect is so strong that such filtering is unnecessary.

Finally, we note that although in most cases the SNR is used for a quantitative analysis of the stochastic resonance, some authors prefer to look at the output signal power S alone, i.e. the coefficient in front of the δ-function in the power spectrum, (3.51). By this measure, a system exhibits stochastic resonance if S has a local maximum at some non-zero input noise intensity. (Compare this against the behavior of a linear system: since superposition applies, S is independent of noise intensity.) In the case of stochastic resonance, S exhibits a maximum at usually roughly the same temperature as the SNR [3,9].

3.9 Two-State Examples: Classical and Quantum Stochastic Resonance

You can find numerous physical examples of stochastic resonance in the technical literature. Let's take a brief look here at three of these. Together, they give an idea of how diverse physical problems can be described by the same underlying theory.

The first example is taken from the original experimental paper reporting stochastic resonance [10]. This experiment involved a simple electronic switch known as a Schmitt trigger. This is a situation where the output is truly two-state: the output voltage is either plus or minus fifteen volts, so we don't have to worry about the complication of intrawell motion. The input to the trigger is the sum of two independent sources, one noisy and one periodic. The amplitude of the periodic input is kept very small – too small to induce switching of the trigger on its own – while the amplitude of the noise can be varied as desired. The experimental results are reflected in Fig. 3.6. The first two, Figs. 3.6a and b, show the output power spectrum for two different levels of input noise. The middle panel (b) corresponds to the larger value of input noise. Two effects are obvious: the sharp line at the signal frequency ($\omega/2\pi = 23$ Hz in the experiment) is larger, and the broadband noise level is smaller. (The latter effect, though striking, turns out to be a feature rarely seen in experiments. It is predicted by the full expression (3.51) though one can see there that it is a relatively minor effect, quantitatively speaking.) It is obvious from the data that the output signal to noise ratio has increased with the addition of input noise, providing a clear demonstration of stochastic resonance. From each power spectrum, one can pull off the signal to noise ratio. The results from a large set of such runs, for different noise strengths and different signal amplitudes, are shown in Fig. 3.6c.

The second example is taken from experiments on a ring laser [11]. In this kind of laser, the cavity is in the form of a closed ring, and laser light can propagate in either the clockwise or counterclockwise sense, but not both. Which of these modes is observed depends on fluctuations in the initial conditions. If the laser is turned off and then turned on again, there is a 50/50 chance it will be in the same mode. It is possible to modify these odds to prefer one of the modes over the other. In the experiment, this was done using a crystal inserted in the laser cavity, and acoustically exciting the crystal with the sum of a small periodic signal and a relatively large noise component. The output was the intensity of the clockwise laser mode. A time series of this output looks like random switching between two states, but there is a regularity to the switching which becomes evident at an optimal level of input noise. The signal to noise ratio can be measured in the same way as in the Schmitt trigger experiment, and the results are shown in Fig. 3.7. Since this system is very nearly two-state, with very little 'intrawell motion', the predictions of the two state theory fit the data extremely well.

The third example is drawn from a theoretical paper investigating the possibility of quantum stochastic resonance [12] . The specific problem studied concerned a two state system coupled to a heat bath of harmonic oscillators. The

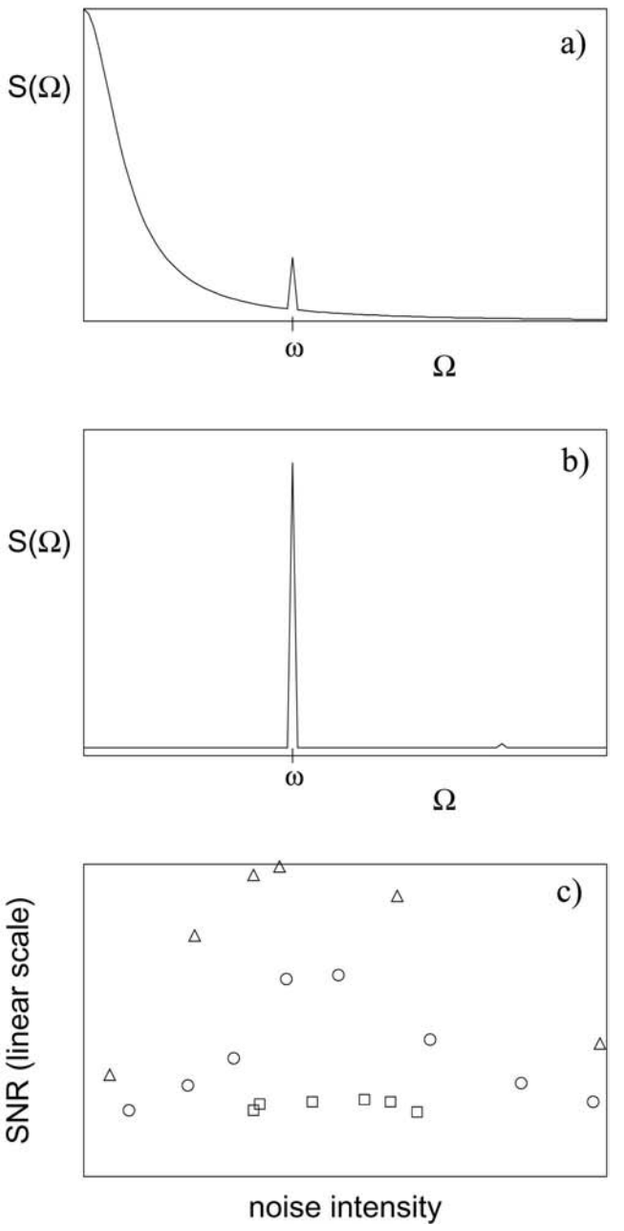


Fig. 3.6. Experimental results from the Schmitt trigger (after Ref. [10]). The top two panels show the power spectrum before (**a**) and after (**b**) adding noise. The bottom panel (**c**) shows a plot gathered from several spectra of the SNR *vs.* input noise level , for three different levels of input signal amplitude (increasing from □ through ∘ to △)

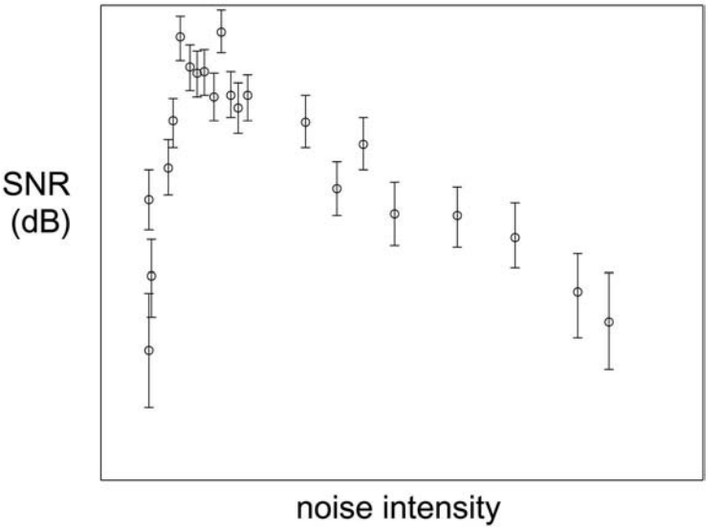


Fig. 3.7. Results from the Georgia Tech ring laser experiment [11], where the intensity of the clockwise laser mode is optimally synchronized with a weak signal injected through an acousto-optical modulator placed in the cavity, at nonvanishing noise level

problem is fundamentally quantum mechanical because the transitions are due to tunneling through the energy barrier separating the two metastable states. As a result, the Kramers formula is not the appropriate theory for the transition rates; instead, one uses the dissipative quantum tunneling formula of Chakravarty and Leggett [13]. The periodic 'signal' in this problem can arise either through a time dependent temperature T or a time dependent barrier height. The results are intriguing: while indeed stochastic resonance is predicted, the conditions when it is expected are different than in the classical case. This is due to the different temperature dependence of the quantum tunneling rates as compared to the classical Kramers rates. In particular, if the underlying potential is symmetric, i.e. for any purely even potential like (3.55), the tunneling rates are proportional to $T^{2\alpha-1}$ (where the dissipation coefficient α quantifies the strength of coupling to the heat bath), instead of depending exponentially on T, like the classical Kramers rate (3.25) (replace there κ by T). As a consequence, quantum stochastic resonance in a symmetric potential is observed only in the regime $\alpha > 1$ of strong dissipation [14]. For weak dissipation, on the other hand, quantum stochastic resonance is predicted when including an asymmetric part to the potential [12] (which would *diminish* the effect in the classical case).

A further, fundamental difference of quantum and classical stochastic resonance is that a rate equation description of the dynamics is not always correct for quantum systems. Indeed, it is well known that, in the absence of any dissipation, a quantum mechanical two state system exhibits purely coherent oscillations, and remnants of this coherent dynamics are also present in the case of very weak dissipation at low temperatures. This regime has been examined in [7],

where stochastic resonance was found only in parameter regimes where incoherent tunneling dominates over coherent transitions. Hence, also in the quantum case, the basic mechanism of stochastic resonance can be understood in terms of a two state model (with quantum noise activated transition rates), although the exact quantitative behavior may be amended by quantum coherence [15].

3.10 Quantum Stochastic Resonance in the Micromaser

In the above discussed dissipative two level system, transitions between the two states are induced by quantum tunneling. Another genuinely quantum mechanical effect which may influence the transition rates of a bistable system is the noise arising from the measurement process : according to the (a priori unpredictable) measurement result, the quantum system is projected onto the corresponding eigenstate of the measured observable. In the following, we will present an example of quantum stochastic resonance, where the measurement noise plays an important role.

Our bistable system will be a single mode of the quantized radiation field (i.e., a harmonic oscillator) sustained by a microwave cavity of very high quality. (By this a very small decay rate γ of the radiation field within the cavity is understood. The storage time γ^{-1} can reach several hundred milliseconds in most advanced experiments [16].) In order to control and to monitor the state of the photon field, we let the field interact with a two-level atom, whose state can be measured after exit from the cavity. If the atom is initially in its upper energy eigenstate $|u\rangle$, and the field in the Fock state $|n\rangle$, then the final quantum state after the atom-field interaction U reads, according to the Jaynes–Cummings model [17] (see Sects. 2.1.1 and 4.11.3 in this volume):

$$U\ |u,n\rangle\ =\ \cos(\phi\sqrt{n+1})|u,n\rangle\ -\ \mathrm{i}\sin(\phi\sqrt{n+1})|d,n+1\rangle. \tag{3.67}$$

Here, the vacuum Rabi angle $\phi = gt_{\mathrm{int}}$ is given by the atom-field coupling strength g and the interaction time t_{int}. As is evident from (3.67), the probability of detecting the atom finally in the lower state $|d\rangle$, and thereby emitting a photon into the field, $|n\rangle \to |n+1\rangle$, is given by

$$\beta_n\ =\ \sin^2(\phi\sqrt{n+1}), \tag{3.68}$$

whereas with probability $1-\beta_n$ the field remains in the state $|n\rangle$.

While this clearly demonstrates the random influence of the atomic detection on the photon field, we still do not see any bistability of the photon field. As we will show in the following, the latter can be induced if we let the photon field interact with a *sequence* of two-level atoms, and also take into account the cavity dissipation. We consider a steady flux r of atoms, initially in the upper state $|u\rangle$, which arrive in the cavity at random, uncorrelated times. We assume that the average time interval r^{-1} between two consecutive atoms is much larger than the interaction time t_{int} of a single atom with the cavity. Then the probability of finding two (or more) atoms in the cavity at the same time can be neglected, and we are dealing with a one-atom or micromaser [18].

In addition to the interaction with the atoms, the photon field is also coupled to the cavity walls, which are cooled to a low temperature $T < 1$ K, such that the mean number $n_b = (\exp(\hbar\omega/kT) - 1)^{-1}$ [19] of photons in thermal equilibrium is smaller than one. According to the standard master equation of the damped harmonic oscillator (employing weak-coupling and Markov approximation) [19], the influence of the heat bath onto the photon field is fully characterized by the temperature T (or, equivalently, n_b) and the cavity decay rate γ, which quantifies the coupling strength to the heat bath. Specifically, given n photons inside the cavity, the probability of emission of a photon from the heat bath into the cavity is given by $\gamma n_b(n+1)$, whereas $\gamma(n_b+1)n$ is the probability of absorption. Thus, in total, the dynamics of the maser field can be described as a jump process between neighboring photon numbers $n \to n \pm 1$ with the following transition rates :

$$t_n^+ = r\ \beta_n\ +\ \gamma\ n_b(n+1)\ , \tag{3.69}$$

$$t_n^- = \gamma\ (n_b+1)\ n\ . \tag{3.70}$$

The average of one single realization of this jump process over a sufficiently long time approaches the following stationary photon number distribution [6]:

$$p_n^{(\mathrm{ss})}\ =\ p_0^{(\mathrm{ss})} \prod_{k=1}^{n} \frac{t_{k-1}^+}{t_k^-}, \tag{3.71}$$

where $p_0^{(\mathrm{ss})}$ is determined by normalization. We assume that the atomic flux r is much larger than the cavity decay rate γ, such that the stationary distribution is far away from thermal equilibrium. Under these circumstances a double-peaked stationary distribution, indicating a bistable photon field, can establish if t_n^+ and t_n^- as a function of n intersect (at least) three times: at n_1 and n_2 (defining a stable equilibrium), and at n_3 (corresponding to an unstable equilibrium, with $n_1 < n_3 < n_2$). An example is shown in Fig. 3.8.

In this case, the photon number will be found almost always near one of the two maxima at n_1 or n_2, and transitions between these metastable states occur. The transition rates $W_{1,2}$ (from n_1 to n_2, and vice versa) of these 'macroscopic' jumps of the photon field can be expressed in terms of the rates for the microscopic jumps [6]:[1]

$$W_1^{-1} = \sum_{n=n_1}^{n_2-1} [p_n^{(\mathrm{ss})} t_n^+]^{-1} \sum_{m=0}^{n} p_m^{(\mathrm{ss})}\ ,$$

$$W_2^{-1} = \sum_{n=n_1+1}^{n_2} [p_n^{(\mathrm{ss})} t_n^+]^{-1} \sum_{m=n}^{\infty} p_m^{(\mathrm{ss})}\ . \tag{3.72}$$

[1] For large n's, if the discrete photon number can be approximated by a continuous variable n, the transition rates may also be obtained by a Kramers analysis, see Sect. 3.4. However, the effective micromaser potential derived in [21] explicitly depends on T, leading to a temperature dependence of the transition rates different from the classical Kramers law (3.29).

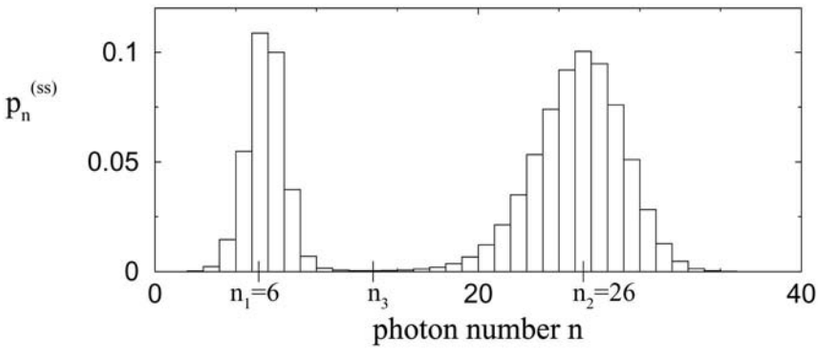


Fig. 3.8. Double-peaked stationary state of the photon number distribution of the micromaser field in the bistable operation mode. The experimental parameters (comparable to the laboratory scenario described in [20]) are: temperature $T = 0.6$ K (corresponding to a mean thermal photon number $n_b = 0.2$), atomic flux $r = 40\gamma$, and vacuum Rabi angle $\phi = 1.03$

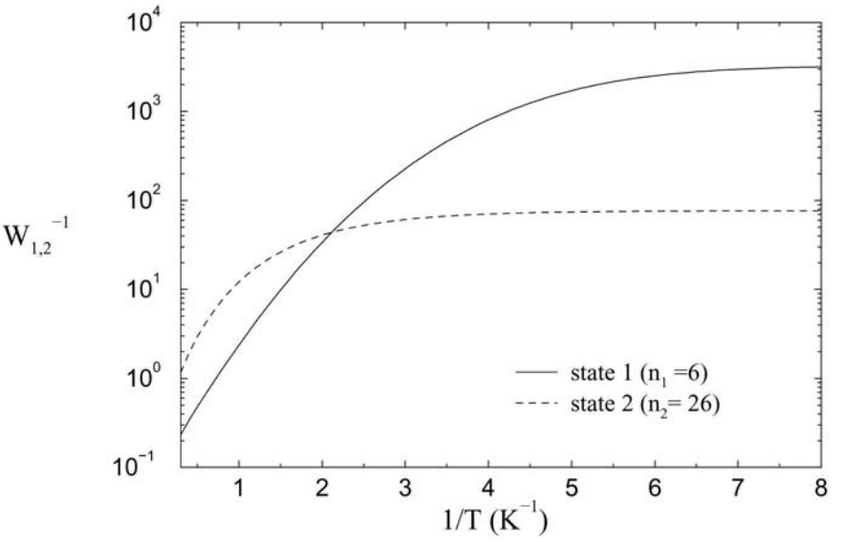


Fig. 3.9. Average residence times $W_{1,2}^{-1}$ in the two metastable states as a function of the inverse temperature $1/T$. Atomic flux $r = 40\gamma = 667$ s^{-1}, vacuum Rabi angle $\phi = 1.03$. A deviation from Kramers' law (a straight line in the semilogarithmic plot) is detected, due to quantum-noise induced transitions at $T = 0$

In Fig. 3.9, we plotted the transition rates as a function of the temperature T, for the same experimental parameters as in Fig. 3.8.

Evidently, the rate W_2 from n_2 to n_1 is more sensitive to the temperature than W_1. The reason is that the microscopic rates t_n^-, given by (3.69), which are responsible for transitions from n_2 to n_1, are proportional to $n_b + 1$, and

therefore (since $n_b < 1$) do not depend on n_b as strongly as the t_n^+'s, which are proportional to n_b. As as result, the two rates intersect at about $T = 0.5\,\mathrm{K}$. A further difference from the classical Kramers rates (3.25) is that they do not vanish even at $T = 0$. This is due to the presence of the noise associated with the atomic detections (the measurement noise) on the one hand, leading to $t_n^+ > 0$, and the fact that the heat bath can absorb photons at random times even at $T = 0$, on the other hand, leading to $t_n^- > 0$. These noise sources are of genuinely quantum mechanical origin, what explains the deviation from the classical Kramers law.

Stochastic resonance appears if we add a periodic signal to our system, e.g., by modulating the atomic flux

$$r(t) = \langle r \rangle + \Delta r \sin(\omega t). \tag{3.73}$$

The modulation amplitude Δr should be small enough to prevent deterministic transitions between the two metastable states, and the modulation period $2\pi/\omega$ should be of the same order of magnitude as the transition rates $W_{1,2}^{-1}$ (close to the intersection $W_1 = W_2$ in Fig. 3.9), in order to enable a matching of time scales (driving frequency $\simeq$ transition frequency). A numerical simulation of the maser dynamics in the presence of such a periodic signal is shown in Fig. 3.10, for three different temperatures of the environment.

Here, we plotted the time evolution of the probability β to detect an exiting atom in the lower state $|d\rangle$, determined by averaging the atomic detection results ('1' or '0' for detection in $|d\rangle$ or $|u\rangle$, respectively) over small time windows of length $\Delta t = 1$ s (ca. 700 detection events). Since β depends on the photon number inside the cavity, see (3.68), the observed quantum jumps between $\beta_{n_1} \simeq 0.15$ and $\beta_{n_2} \simeq 0.65$ are a signature of the jumps of the photon field between the two metastable states around $n_1 = 6$ and $n_2 = 26$. As also predicted by Fig. 3.9, for the lowest temperature $T = 0.3$ K, Fig. 3.10a, the average residence time in state 1 is much longer than the driving period $2\pi/\omega = 42$ s. Consequently, the individual quantum jumps occur at unpredictable times. However, if we add the right amount of noise to the system, i.e., increase the temperature to $T = 0.6$ K, we observe almost periodic transitions: in most cases, the photon field jumps from state 1 to 2 and back again once per modulation period, see Fig. 3.10b. If we further increase the temperature, Fig. 3.10c, the cooperativity between signal and noise is lost again. This illustrates nicely the stochastic resonance effect: the most regular behavior occurs at a finite, nonvanishing noise level.

For a quantitative analysis of the synchronization effect, we can calculate the power spectra either by Fourier transformation of numerically simulated atomic detection sequences (such as in Fig. 3.10), or by employing the two-state model described by (3.33), using the (adiabatically modulated) transition rates given in (3.72). The results are shown in Fig. 3.11, where we plot the strength S of the signal peak as a function of the temperature. While, in both cases, a stochastic resonance maximum occurs at $T \simeq 0.6$ K – the same temperature where the optimal synchronization is observed in Fig. 3.10 – the quantitative behavior is quite different, due to the intrawell modulation, which enhances

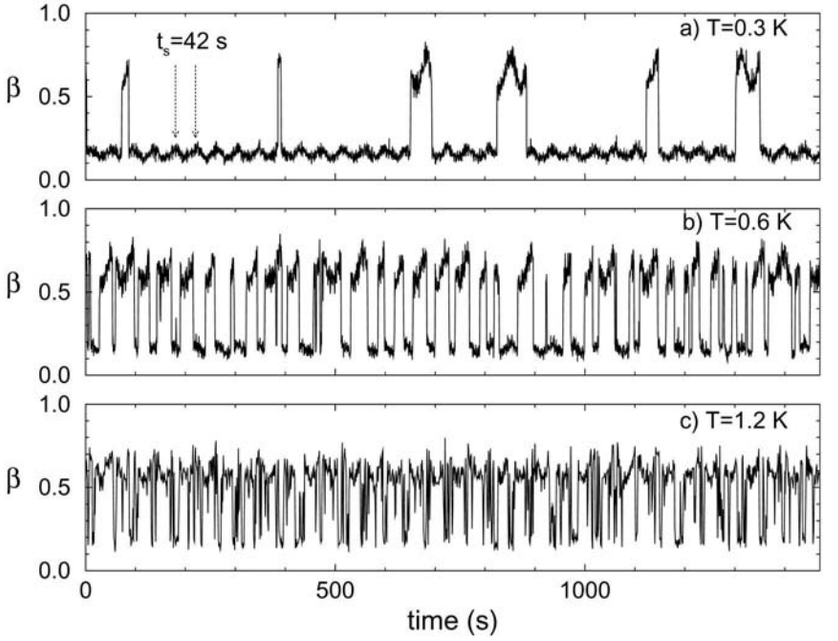


Fig. 3.10. Time evolution of the probability β to detect an atom in $|d\rangle$, with periodically modulated atomic flux $r(t)$, according to (3.73), with mean value $\langle r\rangle = 40\gamma$, modulation amplitude $\Delta r = 6.9\gamma$, and period $2\pi/\omega = 42$ s. Vacuum Rabi angle $\phi = 1.03$. The noise-induced synchronization of quantum jumps is poor for the lowest temperature (too rare quantum jumps), optimal for the intermediate temperature (almost regular quantum jumps), and again poor for the highest temperature (too frequent quantum jumps). Also note the clearly observable intrawell motion for the lowest temperature

(at low T) or suppresses (at high T) the signal as compared to the two-state model [22]. Let us note that the two-state model does *not* exhibit a maximum of the SNR, but rather monotonically increases as a function of T [22], until the two-state approximation breaks down for high temperatures. (This can be traced back to an untypical behavior of the modulated transition rates, whose modulation amplitudes of which increase with increasing temperature, whereas the modulation amplitude (3.25) of the classical Kramers rates is approximately constant in the relevant temperature region.) Although an increase of the SNR with increasing temperature may also be considered as a fingerprint of stochastic resonance, this shows that the signal strength S may, in some cases, give a better quantitative picture of stochastic resonance than the SNR. On the other hand, the exact model (i.e., without two-state approximation) *does* exhibit a maximum of the SNR, since the intrawell dynamics reduce the signal output S at high temperatures, see Fig. 3.11.

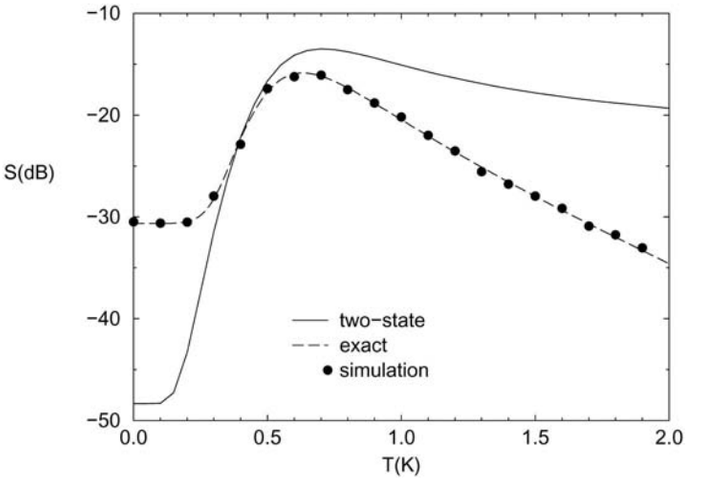


Fig. 3.11. Signal output S as a function of the temperature T, for the same parameters as in Fig. 3.10. A stochastic resonance maximum is observed at $T \simeq 0.6$ K. The circles show the results of the simulation of the maser dynamics (run for 75 000 s for each value of T), which agree perfectly well with the results of an exact calculation of the power spectrum (dashed line) [22]. The deviation from two state model (solid line) reveals the influence of the intrawell dynamics

Finally, we want to discuss briefly the case of 'coherent pumping', where the atoms are injected into the cavity in a coherent superposition

$$|\psi\rangle \;=\; a|u\rangle \;+\; b\,|d\rangle \tag{3.74}$$

of their energy eigenstates. Above, we considered the case $a = 1$, $b = 0$, of 'incoherent pumping'(no initial coherence between the upper and the lower atomic eigenstate). However, by applying a suitable classical microwave pulse on the atoms just before entering the cavity, we may, in principle, choose arbitrary values of a and b. In general, an initial atomic coherence between $|u\rangle$ and $|d\rangle$ will induce nonvanishing coherences of the cavity field between different photon numbers, what prevents the simple description (3.69), (3.70) of the maser dynamics in terms of a jump process between neighboring photon numbers. Nevertheless, for suitably chosen experimental parameters, the photon field exhibits - just as in the case of incoherent pumping - a bistable behavior, with transition rates $W_{1,2}$ between two metastable states (although the calculation of the transition rates is more complicated [15]). In contrast to the case of incoherent pumping, however, the quantum jumps of the photon field can now be monitored also by measuring other components of the atomic Bloch vector on exit from the cavity, e.g., when detecting the atoms in $(|u\rangle \pm |d\rangle)/\sqrt{2}$ instead of $|u\rangle$, $|d\rangle$. Furthermore, we can inscribe the weak periodic signal directly into the initial atomic coherence, by modulating $a(t), b(t)$. A detailed discussion of the thereby achievable stochastic resonance in the atomic coherence can be found in [8,15].

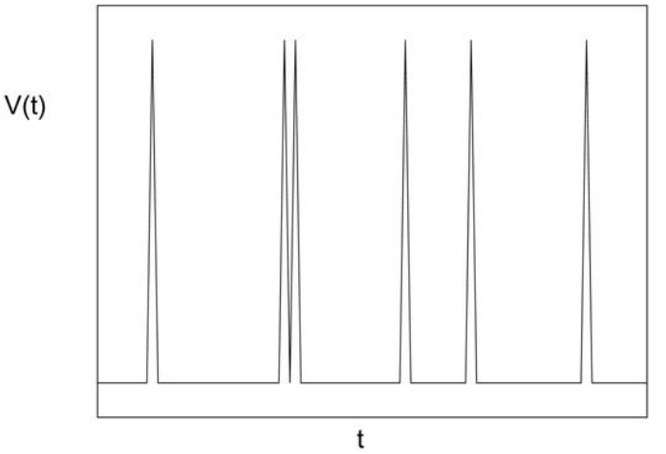


Fig. 3.12. Time series of an excitable system such as a neuron. Most of the time the system resides in its rest state, with output signal $V = 0$. Excitation to its bursting state causes a narrow spike in $V(t)$

3.11 Stochastic Resonance in Excitable Systems

The two state theory describes many systems known to exhibit stochastic resonance. Many, but not all. The original idea and its early development explicitly considers hopping between two stable states. But it has turned out that stochastic resonance occurs more generally. Today, we think that there are a handful of different mechanisms, all of which show the same general properties which we identify with stochastic resonance.

One place where the two state theory won't do is in so-called excitable systems. These are systems which spend most of their time in a resting state, but can enter into a transient bursting state if perturbed strongly enough. Neurons are a common example. A typical time series is shown in Fig. 3.12. Each excitation event is represented as a narrow spike. Most of the time the system is in its rest state, where the output V is zero.

We now develop the theory of stochastic resonance for such a system. The basic idea of the calculation is as follows. First, we imagine that the spike train is characterized by an event rate α which depends on a weak periodic influence as well as a (possibly strong) random influence. In the absence of any input signal, we assume that the events occur randomly and independently at an average rate α_0 which depends on the input noise level. The effect of a weak periodic signal is assumed to periodically modulate the event rate, so that

$$\alpha(t) = \alpha_0 + \varepsilon \cos \omega t \; . \tag{3.75}$$

Here, ω is the frequency of the input signal and ε is a constant which depends on the size of the signal.

We also assume that each event contributes a narrow spike to the output, according to

$$V = \sum_i F(t - t_i) \,, \tag{3.76}$$

where the event times are $t_1, t_2, \ldots$. The pulses are narrow, and it will turn out that only the pulse area (and not its detailed shape) is important. Note that overlapping pulses simply add in the output. Under the assumption that the probability of an event depends only on the instantaneous value of α and not on the arrival times of any previous events, we can calculate the correlation function $\langle V(t)V(t+\tau)\rangle$, and then the power spectrum and the signal to noise ratio.

With this plan in mind, let's first consider the case where α is constant, corresponding to the case of no input signal. Then the problem reduces to the classical 'shot effect' [4]. Suppose that there occur exactly K events in a long time interval $t \in (0, t_{\max})$. We denote the corresponding output by V_K. Then

$$\begin{aligned}\langle V_K(t)V_K(t+\tau)\rangle &= \int_0^{t_{\max}} \frac{\mathrm{d}t_1}{t_{\max}} \ldots \int_0^{t_{\max}} \frac{\mathrm{d}t_K}{t_{\max}} \sum_{i=1}^{K}\sum_{j=1}^{K} F(t-t_i)F(t+\tau-t_j) \\ &= \sum_{i=1}^{K}\sum_{j=1}^{K} \left\{ \int_0^{t_{\max}} \frac{\mathrm{d}t_1}{t_{\max}} \ldots \int_0^{t_{\max}} \frac{\mathrm{d}t_K}{t_{\max}} F(t-t_i)F(t+\tau-t_j) \right\} . \end{aligned} \tag{3.77}$$

The double sum has K^2 terms. Of these, there are K terms with $i = j$ and $K(K-1)$ terms with $i \neq j$:

$$\begin{aligned}\langle V_K(t)V_K(t+\tau)\rangle &= K \int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} F(t-t_i)F(t+\tau-t_i) \\ &\quad + K(K-1) \int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} \\ &\quad \times \int_0^{t_{\max}} \frac{\mathrm{d}t_j}{t_{\max}} F(t-t_i)F(t+\tau-t_j) \,. \end{aligned} \tag{3.78}$$

To explicitly evaluate the integrals we assume that the pulse shape F is a rectangle of width Δt and height H. Then

$$\int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} F(t-t_i)F(t+\tau-t_i) = \frac{H^2}{t_{\max}} \left(\Delta t - |\tau|\right) \,, \tag{3.79}$$

if $|\tau| < \Delta t$ and is otherwise zero. Plotted as a function of τ, this is a triangle of width $2\Delta t$ and area $H^2\Delta t^2/t_{\max}$. In the limit as $\Delta t \to 0$ and $H \to \infty$ such that the product $H\Delta t$ remains constant, this becomes a Dirac delta function

$$\int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} F(t-t_i)F(t+\tau-t_i) = \frac{(H\Delta t)^2}{t_{\max}} \delta(\tau) \,. \tag{3.80}$$

Similarly, we can evaluate the double integral in (3.78),

$$\int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} \int_0^{t_{\max}} \frac{\mathrm{d}t_j}{t_{\max}} F(t-t_i)F(t+\tau-t_j) \approx \left[\int_0^{t_{\max}} \frac{\mathrm{d}t_i}{t_{\max}} F(t-t_i)\right]^2 = \left(\frac{H\Delta t}{t_{\max}}\right)^2 , \tag{3.81}$$

where the first line neglects the small correction for pulses which overlap each other or the endpoints of the interval $(0, t_{\max})$. Thus,

$$\langle V_K(t)V_K(t+\tau)\rangle = \frac{K}{t_{\max}} \left(H\Delta t\right)^2 \delta(\tau) + \frac{K(K-1)}{t_{\max}^2} \left(H\Delta t\right)^2 . \tag{3.82}$$

This expression assumes that there are exactly K events in the interval $(0, t_{\max})$. Now, the probability that there are exactly K events in a time $t_{\max}$ depends on the event rate α according to the Poisson distribution:

$$P_K(t_{\max}) = \frac{(\alpha t_{\max})^K}{K!} \mathrm{e}^{-\alpha t_{\max}} . \tag{3.83}$$

Performing the weighted average of (3.82) over all possible K gives us the correlation function

$$\begin{aligned} C(\tau) &= \sum_{K=0}^{\infty} \langle V_K(t)V_K(t+\tau)\rangle P_K(t_{\max}) \\ &= \left(H\Delta t\right)^2 \sum_{K=0}^{\infty} \left\{ \frac{K}{t_{\max}} \delta(\tau) + \frac{K\,(K-1)}{t_{\max}^2} \right\} \frac{(\alpha t_{\max})^K}{K} \mathrm{e}^{-\alpha t} \\ &= \left(H\Delta t\right)^2 \left\{ \alpha\delta(\tau) + \alpha^2 \right\} , \end{aligned} \tag{3.84}$$

and the corresponding power spectrum is, using (3.9),

$$S(\Omega) = \left(H\Delta t\right)^2 \left\{ 2\alpha + 4\pi\alpha^2\delta(\Omega) \right\} . \tag{3.85}$$

We now repeat the calculation, this time using a time dependent rate $\alpha(t)$. We modify the Poisson distribution (3.83) as follows

$$P_K(t_{\max}) = \frac{1}{K!} Z^K \mathrm{e}^{-Z} , \tag{3.86}$$

where

$$Z(t_{\max}) = \int_0^{t_{\max}} \alpha(t)\mathrm{d}t . \tag{3.87}$$

As a check, this properly reduces to (3.83) when α is constant.

If there are exactly K events in $(0, t_{\max})$,

$$\begin{aligned} \langle V_K(t)V_K(t+\tau)\rangle = \sum_{i=1}^{K} \sum_{j=1}^{K} \int_0^{t_{\max}} \frac{\alpha(t_1)\mathrm{d}t_1}{Z} \cdots \\ \times \int_0^{t_{\max}} \frac{\alpha(t_K)\mathrm{d}t_K}{Z} F(t-t_i)F(t+\tau-t_j) . \end{aligned} \tag{3.88}$$

Group terms according to whether or not $i = j$,

$$\begin{aligned}\langle V_K(t)V_K(t+\tau)\rangle &= \frac{K}{Z}\int_0^{t_{\max}} \alpha(t_i)F(t-t_i)F(t+\tau-t_i)\mathrm{d}t_i \\ &+ \frac{K(K-1)}{Z^2}\int_0^{t_{\max}} \alpha(t_i)F(t-t_i)\mathrm{d}t_i \\ &\times \int_0^{t_{\max}} \alpha(t_j)F(t+\tau-t_j)\mathrm{d}t_j \ . \end{aligned} \tag{3.89}$$

If F is sharply peaked,

$$\int_0^{t_{\max}} \alpha(t_i)F(t-t_i)\mathrm{d}t_i \approx \alpha(t)\int_0^{t_{\max}} \mathrm{d}t_i F(t-t_i) = \alpha(t)H\Delta t \ , \tag{3.90}$$

and

$$\begin{aligned}\int_0^{t_{\max}} \alpha(t_j)F(t+\tau-t_j)\mathrm{d}t_j &\approx \alpha(t+\tau)\int_0^{t_{\max}} \mathrm{d}t_j F(t+\tau-t_j) \\ &= \alpha(t+\tau)H\Delta t \ , \end{aligned} \tag{3.91}$$

as well as

$$\begin{aligned}\int_0^{t_{\max}} \alpha(t_i)F(t-t_i)F(t+\tau-t_i)\mathrm{d}t_i &\approx \alpha(t)\int_0^{t_{\max}} F(t-t_i)F(t+\tau-t_i)\mathrm{d}t_i \\ &= \alpha(t)\left(H\Delta t\right)^2\delta(\tau) \ , \end{aligned} \tag{3.92}$$

so that

$$\begin{aligned}\langle V_K(t)V_K(t+\tau)\rangle &= \frac{K}{Z}\alpha(t)\left(H\Delta t\right)^2\delta(\tau) \\ &+ \frac{K(K-1)}{Z^2}\alpha(t)\alpha(t+\tau)\left(H\Delta t\right)^2 \ , \end{aligned} \tag{3.93}$$

if there are exactly K events in $(0, t_{\max})$. Taking the weighted sum over all K yields the full correlation function :

$$\begin{aligned}C(\tau;t) &= \sum_{K=0}^{\infty}\langle V_K(t)V_K(t+\tau)\rangle P_K(t_{\max}) \\ &= \left(H\Delta t\right)^2\left\{\alpha(t)\delta(\tau)+\alpha(t)\alpha(t+\tau)\right\} \ . \end{aligned} \tag{3.94}$$

As we have come to expect, the presence of the periodic signal results in an expression which depends on both τ and t, so we perform a phase average to eliminate the t-dependence. For example, suppose

$$\alpha(t) = \alpha_0 + \alpha_1\cos\left(\omega t+\psi\right) \ . \tag{3.95}$$

Then we have

$$\begin{aligned}\langle C(\tau;t)\rangle_\psi &= \frac{1}{2\pi}\int_0^{2\pi} C(\tau;t)\mathrm{d}\psi \\ &= (H\Delta t)^2\left\{\delta(\tau)\alpha_0+\alpha_0^2\right. \\ &\quad \left.+\alpha_1^2\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega t+\psi)\cos(\omega t+\omega\tau+\psi)\mathrm{d}\psi\right\} \\ &= (H\Delta t)^2\left\{\delta(\tau)\alpha_0+\alpha_0^2+\frac{1}{2}\alpha_1^2\cos\omega\tau\right\}\ , \end{aligned} \tag{3.96}$$

and the power spectrum becomes

$$S(\Omega) = (H\Delta t)^2\left\{2\alpha_0+\frac{4}{\pi}\alpha_0^2\delta(\Omega)+\frac{2}{\pi}\alpha_1^2\delta(\Omega-\omega)\right\}\ . \tag{3.97}$$

More generally, if the rate α is periodic, we can write

$$\alpha(t) = \alpha_0+\sum_{j=1}^{\infty}\alpha_j\cos\left(j\omega t+\psi_j\right)\ . \tag{3.98}$$

Going through the phase average as before, we arrive at the power spectrum

$$S(\Omega) = (H\Delta t)^2\left\{2\alpha_0+\frac{4}{\pi}\alpha_0^2\delta(\Omega)+\frac{2}{\pi}\sum_{j=1}^{\infty}\alpha_j^2\delta(\Omega-j\omega)\right\}\ . \tag{3.99}$$

This describes a constant broadband background plus a series of spikes at the signal frequency ω and its harmonics. Reading off the signal to noise ratio yields

$$\mathrm{SNR} = \frac{\alpha_1^2}{\pi\alpha_0}\ . \tag{3.100}$$

To complete the theory, we need to know how the event rate depends on the system parameters, in particular the input noise intensity. Since this depends on the specific details of the system, we consider an example. Suppose α obeys a Kramers-type formula

$$\alpha(t) = \exp\left[-\frac{U}{\kappa}\left(1+\eta\cos\omega t\right)\right]\ , \tag{3.101}$$

where κ is the noise strength and U,η, and ω are constants. The justification for this form in terms of a particular Langevin model can be found elsewhere [23]; U plays the role of the potential barrier and η is proportional to the signal amplitude. The rate can be Fourier expanded as in (3.98). For the SNR we need only the lowest two coefficients α_0 and α_1, with result

$$\mathrm{SNR} = \frac{8I_1^2(z)}{\pi I_0(z)}\mathrm{e}^{-U/\kappa}\ , \tag{3.102}$$

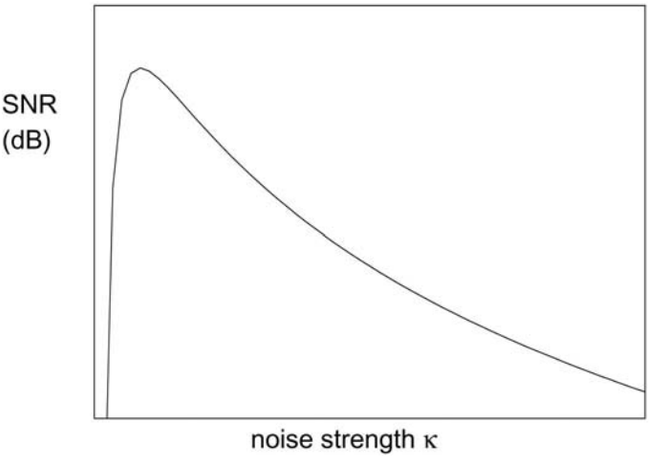


Fig. 3.13. Signal to noise ratio vs. noise input strength from the theory for excitable systems. Much as for bistable systems treated in the preceding sections, the maximum SNR is reached at an optimal, non-vanishing noise level

where I_n is the modified Bessel function of order n and $z = \eta U/\kappa$. A plot of the SNR *vs.* κ is shown in Fig. 3.13. The theory predicts an increase in SNR over some range of noise input. The theory does a fair job explaining data from, e.g., experiments on crayfish mechanoreceptors and simulations of the Fitzhugh-Nagumo equation [23]. The biggest discrepancy is that the theory predicts too rapid a fall off at high noise. This could be due to the breakdown of some fundamental assumption such as statistical independence of consecutive events, or may simply reflect that the Kramers-type formula (3.101) doesn't properly capture the true rate dependence. In fact, there is nothing published (to our knowledge) on what the correct rate formula should be for either the crayfish neuron or the Fitzhugh-Nagumo equation, so there is no reason to expect close agreement. The important point to take away is that the qualitative behavior is common to the various systems, and is the same whether the underlying dynamics represents bistable switching or excitable bursting.

3.12 The Frontier of Stochastic Resonance

Noise is usually considered a nuisance, but in systems which display stochastic resonance an increase in input noise improves their sensitivity to discriminate weak signals. Stochastic resonance is now firmly established as a common phenomenon which appears in a wide variety of physical situations. And the list of examples continues to grow.

The study of stochastic resonance is by now a mature field. There are several similar but distinct fundamental mechanisms that can give rise to the effect, two of which we covered in this chapter. The quantum mechanical version of

stochastic resonance is far less studied than its classical counterpart, and the future may hold great advances in this direction. Indeed, quantum opticians are recently getting aware of this robust phenomenon [8,15,22,24,25,26,27,28]. In the classical context, the great unresolved issue is whether there are any practical applications of stochastic resonance.

Ideas for applications fall into two categories. The first involves engineering technology. There have been proposals for direct use in electronics, resulting in at least two U.S. government patents [29,30]. Another technological use is to retrofit threshold detectors whose performance have degraded with age. The most interesting example of this type may be the biomedical application of specially designed stockings for people with diminished balancing ability to help them stand up [31].

The other category of applications is biological systems: does Mother Nature already use stochastic resonance in some of her detectors? Sensory neurons are notoriously noisy, and stochastic resonance might account for the exquisite sensitivity of some animals to detect weak coherent signals. The first experiments on biological stochastic resonance were reported in 1993 using mechanoreceptors of the crayfish Procambarus clarkii [32]. Two years later, experiments demonstrated stochastic resonance at the sub-cellular level in lipid bilayer membranes [33]. Other examples include the mechanosensory systems used by crickets to detect air currents [34] and by rats to detect pin pricks [35]. Experiments on hair cells, important auditory detectors in many vertebrates (including humans), are especially suggestive [36]: the optimal noise level appears to coincide with the naturally occurring level set by equilibrium thermal fluctuations! Perhaps these cells evolved to take maximum advantage of these inevitable environmental fluctuations.

The change in context from physical to life sciences has led researchers to reconsider and refine even some very basic issues. For example, the extension of stochastic resonance to excitable systems was primarily motivated by experiments on neurons. In a similar way, a theoretical mechanism which employs a randomly fluctuating rate [37] raises interesting fundamental questions concerning its connection with microscopic stochastic descriptions. Another question of great importance is: what is the most appropriate measure of 'output performance'? In the biological context information transmission is more significant than signal-to-noise ratio, but it may be that the most relevant measure – whatever it is – depends on the particular application. Related to this is the question: what kind of signals are most relevant? Truly periodic signals are uncommon in the natural world, and the nonlinear nature of stochastic resonance suggests that the study of complicated signals cannot be reduced to a superposition of elemental periodic ones. And what about other important biological properties of sensory neurons such as adaptation and refraction? These effects are absent from existing theories of stochastic resonance. Finally, the biological context has given new impetus to the study of stochastic resonance in arrays of elements [38,39], where it appears both that overall performance can be improved and that tuning of the noise strength may be unnecessary.

References

1. K. Wiesenfeld, F. Jaramillo: Chaos **8**, 539 (1998)
2. F. Moss, K. Wiesenfeld: Scientific American **273**, 66 (1995)
3. L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni: Rev. Mod. Phys. **70**, 223 (1998)
4. see, *e.g.*, S.O. Rice, in: *Selected Papers on Noise, Stochastic Processes*, ed. by N. Wax (Dover, New York 1954)
5. P. Jung: Phys. Rep. **234**, 175 (1993)
6. C. Gardiner: *Handbook of Stochastic Methods*, 2nd ed. (Springer, Berlin 1983)
7. M. Grifoni, P. Hänggi: Phys. Rev. Lett. **76**, 1611 (1996), Phys. Rev. E **54**, 1390 (1996)
8. T. Wellens, A. Buchleitner: Phys. Rev. Lett. **84**, 5118 (2000)
9. A. Bulsara, L. Gammaitoni: Physics Today **49**, 39 (March 1996)
10. S. Fauve, F. Heslot: Phys. Lett. A **97**, 5 (1983)
11. B. McNamara, K. Wiesenfeld, R. Roy: Phys. Rev. Lett. **60**, 2626 (1988)
12. R. Löfstedt, S. N. Coppersmith: Phys. Rev. Lett. **72**, 1947 (1994)
13. S. Chakravarty, A. J. Leggett: Phys. Rev. Lett. **52**, 5 (1984)
14. I. Goychuk, P. Hänggi: Phys. Rev. E **59**, 5137 (1999)
15. T. Wellens, A. Buchleitner: Chem. Phys. **268**, 131 (2001)
16. B. T. H. Varcoe *et al.*: Nature **403**, 743 (2000)
17. E. T. Jaynes, F.W. Cummings: Proc. IEEE **51**, 89 (1963)
18. J. Krause, M. O. Scully, H. Walther: Phys. Rev. A **34**, 2032 (1986)
19. P. Meystre, M. Sargent III: *Elements of Quantum Optics* (Springer, Berlin 1990)
20. O. Benson, G. Raithel, H. Walther: Phys. Rev. Lett. **72**, 3506 (1994)
21. P. Filipowicz, J. Javanainen, P. Meystre: Phys. Rev. A **34**, 3077 (1986)
22. T. Wellens, A. Buchleitner: J. Phys. A **32**, 2895 (1999)
23. K. Wiesenfeld *et al.*: Phys. Rev. Lett. **72**, 2125 (1994)
24. A. Buchleitner, R. N. Mantegna: Phys. Rev. Lett. **80**, 3932 (1998)
25. L. Viola *et al.*: Phys. Rev. Lett. **84**, 5466 (2000)
26. S. F. Huelga, M. Plenio: Phys. Rev. A **62**, 52111 (2000)
27. L. Sanchez-Palencia *et al.*: Phys. Rev. Lett. **88**, 133903 (2002)
28. P. K. Rekdal, B.-S. K. Skagerstam: Physica A **305**, 404 (2002)
29. A. D. Hibbs: 'Detection and communications device employing stochastic resonance', U.S. Patent No. 5574369 (1996)
30. A. R. Bulsara *et al.*: 'Controlled stochastic resonance circuit', U.S. Patent No. 6285249 (2001)
31. J. Niemi, A. Priplata, M. Salen, J. Harry, J. J. Collins: *Noise-enhanced balance control*, preprint (2001)
32. J. K. Douglass, L. Wilkens, E. Pantazelou, F. Moss: Nature (London)**365**, 337 (1993)
33. S. M. Bezrukov, I. Vodyanoy: Nature (London) **378**, 362 (1995)
34. J. E. Levin, J. P. Miller: Nature (London) **380**, 165 (1996)
35. J. J. Collins, T. T. Imhoff, P. Grigg: J. Neurophysiology **76**, 642 (1996)
36. F. Jaramillo, K. Wiesenfeld: Nature Neuro. **1**, 384 (1998)
37. S. M. Bezrukov, I. Vodyanoy: Nature (London) **385**, 319 (1997)
38. J. J. Collins, C. C. Chow, T. T. Imhoff: Nature (London), **376**, 236-238 (1995)
39. J. F. Lindner, B. K. Meadows, W. L. Ditto, M. E. Inchiosa, A. R. Bulsara: Phys. Rev. Lett. **75**, 3-6 (1995)

4 Quantum Markov Processes

Burkhard Kümmerer

Technische Universität Darmstadt, Fachbereich Mathematik,
Schloßgartenstraße 7, D-64289 Darmstadt

Introduction

These notes give an introduction to some aspects of quantum Markov processes. Quantum Markov processes come into play whenever a mathematical description of irreversible time behaviour of quantum systems is aimed at. Indeed, there is hardly a book on quantum optics without having at least a chapter on quantum Markov processes. However, it is not always easy to recognize the basic concepts of probability theory in families of creation and annihilation operators on Fock space. Therefore, in these lecture notes much emphasis is put on explaining the intuition behind the mathematical machinery of classical and quantum probability. The lectures start with describing how probabilistic intuition is cast into the mathematical language of classical probability (Sects. 4.1–4.3). Later on, we show how this formulation can be extended such as to incorporate the Hilbert space formulation of quantum mechanics (Sects. 4.4,4.5). Quantum Markov processes are constructed and discussed in Sects. 4.6,4.7, and we add some further discussions and examples in Sects. 4.8–4.11.

It follows a detailed description of the contents of these lectures. The first three sections provide the necessary background from the theory of classical Markov processes. Even in discussions on quantum optics the classical theory still has its place, not only as a motivation, but also when discussing stochastic behaviour through phase space methods (e.g. P–, Q–, and Wigner representations).

The mathematical treatment of classical stochastic behaviour naturally breaks into two parts: First there is what could be called the phenomenological description of stochasticity: It is confined to develop a mathematical description of what one really sees. After the discussion of a motivating example (Sect. 4.1) this is discussed in Sect. 4.2. In a second step one goes, however, much further: One assumes that the world can be described by certain mathematical models, for example by random variables on an abstract probability space. In most cases this space is not believed to be a realistic model of the world, but nevertheless it turns out to be extremely useful. This approach is disscussed in Sect. 4.3.

In the following Sect. 4.4 we give a formulation of traditional quantum mechanics. It then needs to be extended in order to include also the description of classical systems. Within this framework we are finally in a position to define the basic notions of probability, including the notion of a Markov process, such that it is applicable to classical as well as to quantum systems. This is done in Sect. 4.5. After this preparations we construct and discuss quantum Markov processes in

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 139–198, 2002.

Sect. 4.6. It turns out that in a quantum context the construction of Markov processes is different from what one is used to from the classical theory. As an illustration a class of simple processes describing a spin-1/2-particle in a stochastic magnetic field is discussed in Sect. 4.7. Associated to any Markov process is a semigroup of completely positive transition operators. Complete positivity is introduced in Sect. 4.8 while continuous semigroups of completely positive operators and their Lindblad generators are discussed in Sect. 4.9. In Sect. 4.10 a different type of quantum Markov processes is described which is used in the description of repeated quantum measurement. We discuss the ergodic theory of these processes and show that a single long sequence of repeated measurements contains the same information as a sample of many such sequences. Finally we show in Sect. 4.11 that the micromaser is a realization of a quantum Markov process and we discuss an application of the general theory to the preparation of quantum states of this system.

4.1 An Example

This first section motivates some fundamental concepts with a simple example.

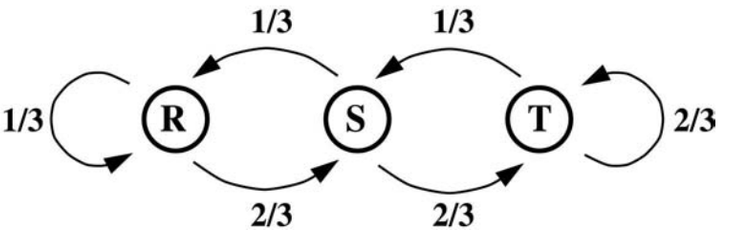
4.1.1 The System

Suppose, we are observing over some time a travelling salesman. He is commuting between three different cities, called (for historical reasons) R, S, and T. On a single day we find that he is doing at most one travelling (or he stays overnight in the city where he is). After some time of observation we find out that he travels between R, S, and T with certain probabilities. More precisely, suppose that we find out that

i) the probability for taking a route does not change in time,
ii) the probability for taking a route (say S $\to$ R) depends only on the starting point (S), but not on the history (how he did come to S).

The first property is called *homogeneity* (in time), the second is the *Markov property* : It appears as if the salesman had no memory.

To have a concrete example in mind suppose we observe a behaviour as indicated by the following diagram:



For example, when in R, then the salesman travels with a probability of 2/3 to S, with a probability of 1/3 he stays at R overnight. There seems to be no direct way to come from R to T within one day (so we omitted this route which is taken with probability 0).

4.1.2 One Time Step

It turns out to be convenient and useful to write this information into the entries of a matrix: We number the cities R, S, and T from 1 to 3 and write the probability for taking the route from city i to city j into the (i, j)–entry of a 3×3–matrix T, i.e.,

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix} .$$

This matrix is called *transition matrix* (this is why we call it T) and it contains all information on the process. By its probabilistic interpretation, all its entries have to be positive real numbers between 0 and 1. Moreover, when being in a city, the salesman has to do something: Therefore the entries in each individual row have to sum up to 1. This is equivalent to the requirement

$$T \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

In general, such matrices with non-negative entries and rows summing up to 1 are called *stochastic matrices*. They play a fundamental role in the phenomenological description of Markovian behaviour and we will meet various generalizations of them.

4.1.3 Many Time Steps

Suppose now that we observed the salesman starting in S but we have not been able to follow which route he has taken afterwards. So the best we can say on his whereabouts the next day is his probability distribution, in our case $(1/3, 0, 2/3)$. And what can we say about the day after? By summing over all possible paths we will now find the probability distribution $(1/9, 4/9, 4/9)$.

For example, the probability to come from S to S within two days is computed as

$$\begin{aligned} \mathbb{P}(\mathrm{S} \xrightarrow{2} \mathrm{S}) &= \mathbb{P}(\mathrm{S} \to \mathrm{R} \to \mathrm{S}) + \mathbb{P}(\mathrm{S} \to \mathrm{S} \to \mathrm{S}) + \mathbb{P}(\mathrm{S} \to \mathrm{T} \to \mathrm{S}) \\ &= \quad \tfrac{1}{3} \cdot \tfrac{2}{3} \quad + \quad 0 \cdot 0 \quad + \quad \tfrac{2}{3} \cdot \tfrac{1}{3} \\ &= \quad \tfrac{4}{9} . \end{aligned}$$

It is one of the basic tools of probability theory that summing over paths is 'the same' as matrix multiplication. Indeed, computing the probability of coming from city i to city j within two days gives

$$\begin{aligned} T_{\text{two days}}(i,j) &= \textstyle\sum_k T(i,k)T(k,j) \\ &= T^2(i,j)\,. \end{aligned}$$

Similarly, for n days we obtain

$$\begin{aligned} T_{n\text{ days}}(i,j) &= \textstyle\sum_{k_1,k_2,\ldots,k_{n-1}} T(i,k_1)T(k_1,k_2)\ldots T(k_{n-1},j) \\ &= T^n(i,j)\,. \end{aligned}$$

These equations are referred to as *Chapman–Kolmogorov equations*. Therefore, the salesmans probability distribution after two days can be computed as

$$\left(\frac{1}{9},\frac{4}{9},\frac{4}{9}\right) = \left(\frac{1}{3},0,\frac{2}{3}\right)\cdot T = (0,1,0)\cdot T^2$$

(row by matrix multiplication).

In general, if we know at a certain day the salesman's probability distribution to be (p_1,p_2,p_3) (where $p_i \geq 0\,,\ \sum_i p_i = 1$) then our knowledge n days later, when we haven't been able to do any observation, is adequately described by

$$(p_1,p_2,p_3)\cdot T^n\,.$$

We end this discussion with the following observation: $\pi := \left(\frac{1}{7},\frac{2}{7},\frac{4}{7}\right)$ is a stationary probability distribution, i.e., $\pi\cdot T = \pi$. Moreover, when starting with an arbitrary initial probability distribution π_0, then $\lim_{n\to\infty}\pi_0\cdot T^n = \pi$. Thus, when we are not able to perform further observations, we find a high probability for the salesman to be in T – perhaps he is living there.

4.1.4 Outlook

In the discussion of this simple example we already touched many basic ideas of a phenomenological description of Markovian behaviour.

All of what comes now is merely a reformulation of these ideas into more elaborate mathematical terms. Unfortunately, this is necessary when one wants to describe systems whose behaviour in time still follows the same intuitive ideas, but their state spaces are more complicated than consisting just of three elements like $\{\mathrm{R},\mathrm{S},\mathrm{T}\}$. A second complication comes when we consider continuous time and we can no longer follow step by step of what the system is doing.

There is an immense literature on Markov chains with *finite* state space since they model appropriately a large number of situations of practical interest. Almost every book on probability contains chapters on this subject, like [5,7,27]. Large parts of these discussions are concerned with the asymptotic behaviour

of those processes like existence of periodic states, convergence to stationary probability distributions, etc., but this is outside the scope of the present lectures. It should be noted that already the extension of these results to the case of a countable state space is not trivial. The books [29] and [11] contain a lot of information on this particular subject.

4.2 Markovian Behaviour on General State Spaces

When we consider a system whose possible states cannot be collected into a finite set and we want to describe this even in continuous time, then there are at least three problems to face which force us to blow up the mathematical apparatus enormously:

1) If there is a continuum Ω_0 of possible states, then the probability for moving from a state $x \in \Omega_0$ to a state $y \in \Omega_0$ will be zero in most cases. Therefore, instead of considering transition probabilities between states, we should consider transitions from a state $x \in \Omega_0$ into a subset $A \subseteq \Omega_0$ such that their probabilities do no longer vanish.
2) When trying to associate probabilities for transitions from x to arbitrary subsets $A \subseteq \Omega_0$ in a consistent way one is led into deep mathematical problems. The way out is to allow A to be a member only of a σ-algebra of subsets of Ω_0. Thus measure theory comes into play.
3) In continuous time there is no smallest time step. Therefore, one either has to consider all time steps simultaneously – which leads to considering continuous semigroups of transitions – or one tries to consider infinitesimal time steps – which leads to a description by partial differential equations. But in both cases one first has to create 'good' spaces where such formulations can fruitfully be carried through.

4.2.1 Transition Probabilities with Densities

If the state space Ω_0 is not too general, $\mathbb{R}^n$ or $\mathbb{C}^n$ would be typical cases, then one can still hope that transition probabilities can be described by a density function $p(x, y)$ on $\Omega_0 \times \Omega_0$ such that the probability of coming from $x \in \Omega_0$ to a subset $A \subseteq \Omega_0$ is given by

$$\mathbb{P}(x, A) = \int_A p(x, y) \mathrm{d}y \ .$$

Perhaps the most prominent example is the Gaussian density given by

$$p(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \mathrm{e}^{-(x-y)^2/2\sigma^2}$$

for $\Omega_0 = \mathbb{R}$.

In such a formulation a Markovian behaviour in discrete time is determined by a *transition probability density function* p on $\Omega_0 \times \Omega_0$ with values in the non-negative real numbers such that

$$\int_{\Omega_0} p(x,y)\mathrm{d}y = 1$$

for all initial values $x \in \Omega_0$. The function p describes the transition probabilities for one time step.

If we assume again, as in our first example, that the transition probabilities do not change in time and don't depend on the history, then we can compute from $p =: p_1$ the transition probability density function for two time steps by

$$p_2(x,y) := \int_{\Omega_0} p(x,z)p(z,y)\mathrm{d}z$$

and for n steps by

$$\begin{aligned} p_n(x,y) &:= \textstyle\int_{\Omega_0} p(x,z)\cdot p_{n-1}(z,y)\mathrm{d}z \\ &= \textstyle\int_{\Omega_0^{n-1}} p(x,z_1)\cdot p(z_1,z_2)\cdot\ldots\cdot p(z_{n-1},y)\mathrm{d}z_1\ldots\mathrm{d}z_{n-1}\,. \end{aligned}$$

Again, this is called the Chapman–Kolmogorov equation.

As in our introductory example in Sect. 4.1, if we start with an initial probability density p_0 on Ω_0, it will change to

$$p_n(x) = \int_{\Omega_0} p_0(z)p_n(z,x)\mathrm{d}z$$

after n time steps (in this notation one has to distinguish between p_i with one and p_i with two arguments).

So far we did not further specify the type of functions used for the transition probability densities. Indeed, here comes a problem: Such a probability density doesn't need to be continuous, it even doesn't need to be a function: Even in the case when nothing is happening and the system does not change its state, the appropriate choice for p would be the δ-function $\delta(x-y)$, which is not a function but a measure. Similarly, if the system jumps from one state into another state with some positive probability, we cannot expect p to be a classical function. Nevertheless, with some precautions and good intuition such functions can be used.

4.2.2 Transition Kernels

What follows is a brief outline of the most general description of Markovian behaviour. This will then be specialized in Sect. 4.2.3 to a situation which is more tractable. For convenience we first review some basic notions of measure theory.

A *σ-algebra* Σ of Ω is a set of subsets of Ω which contains the empty set, it contains with each set also its complement and it contains with any countable set of subsets also its union. On spaces like $\mathbb{R}$ and $\mathbb{C}$ we consider always the *Borel-σ-algebra* which is generated by the open subsets.

A *measure* μ on (Ω, Σ) is a map $\mu : \Sigma \to [0, \infty] := \mathbb{R}_+ \cup \infty$ which is *σ-additive*, i.e., additive on countable unions of disjoint subsets in Σ. If $\mu(\Omega) = 1$ then μ is called a *probability measure* and (Ω, Σ, μ) a *probability space*. Further information on the basic notions of measure theory may be found, e.g. in [28].

Fixing some further notation, if Ω is any set and Σ a σ-algebra of subsets of Ω, we call (Ω, Σ) a *measurable space*. If (Ω, Σ) and (Ω', Σ') are measurable spaces then a map $f : \Omega \to \Omega'$ is called *measurable* if $f^{-1}(A) \in \Sigma$ for all $A \in \Sigma'$. If μ is a (probability) measure on (Ω, Σ) such a map f induces a (probability) measure μ' on (Ω', Σ') via $\mu'(A) := \mu(f^{-1}(A))$, which is called its *distribution* .

The basic idea behind the following is to work persistently with the probabilities for jumping from a state $\omega \in \Omega$ into a subset $A \subseteq \Omega$. Fixing the state ω this should be a probability measure on the subsets:

A *kernel* on a measurable space (Ω, Σ) is a mapping $K : \Omega \times \Sigma \to [0, \infty]$ such that

- for all $A \in \Sigma : \Omega \ni \omega \mapsto K(\omega, A) \in [0, \infty]$ is measurable,
- for all $\omega \in \Omega : \Sigma \ni A \mapsto K(\omega, A)$ is a measure on (Ω, Σ).

If $K(\omega, \Omega) = 1$ for all $\omega \in \Omega$ then K is called a *Markov kernel* and the number $K(\omega, A) \in [0, 1]$ should be interpreted as the probability of coming from ω into the set A.

A Markov kernel K transforms a probability measure μ on (Ω, Σ) into the probability measure

$$\mu \circ K : \Sigma \ni A \mapsto \int_\Omega K(\omega, A) \mathrm{d}\mu(\omega) \ . \qquad (*)$$

Given two kernels K and L they can be composed to a new kernel $K \circ L$ defined as

$$K \circ L(\omega, A) := \int_\Omega L(y, A) K(\omega, \mathrm{d}y) \ , \qquad (**)$$

where $K(\omega, \mathrm{d}y)$ denotes integration with respect to the probability measure $\Sigma \ni A \mapsto K(\omega, A)$.

Markov kernels generalize stochastic matrices: If $T = (\tau_{ij})_{i,j}$ is any stochastic $n \times n$-matrix then on $\Omega := \{1, \ldots, n\}$ a kernel $K = K_T$ is obtained as

$$K_T(i, A) = \sum_{j \in A} \tau_{ij}$$

for $A \subseteq \Omega$. Conversely, every kernel on Ω is of this type. Equation (*) generalizes the fact that a stochastic matrix T transforms a probability distribution π into the probability distribution $\pi \cdot T$, and (**) generalizes matrix multiplication.

Similarly, Markov kernels generalize a description by transition probability densities: If $p : \Omega \times \Omega \to \mathbb{R}$ is such a density then $K(\omega, A) := \int_A p(\omega, y)\mathrm{d}y$ will be a Markov kernel.

Thus, in a risky formulation, Markov kernels are something like stochastic matrices with probability measures as columns.

Alike transition matrices and transition probability densities, Markov kernels determine a Markovian system: If a kernel K on Ω_0 models the behaviour within one time step then, under the same assumptions as before, $K \circ K$ models the behaviour within two time steps and the n-fold composition $K \circ \ldots \circ K$ for n time steps.

Despite their abstractness Markov kernels are used and even generalized in physics, for example in the theory of repeated measurement ([3]). A version of such kernels is discussed in Sect. 4.10.

4.2.3 Transition Operators

While the formulation with transition probability densities is too special the notion of a Markov kernel is so general that one cannot expect to have a good description of, e.g., infinitesimal behaviour in continuous time. Fortunately, there is a formulation of intermediate generality which is very useful both for classical probability theory as well as for possible generalizations to quantum mechanics. Basically it is the transition from the Schrödinger picture to the Heisenberg picture, but now performed within the framework of the classical theory:

Consider, again, the example of a stochastic matrix $T = (\tau_{ij})_{i,j}$. If $\pi = (\pi_1, \ldots, \pi_n)$ is the vector of a probability distribution then its change after one time step is given by $\pi \cdot T$, i.e., by row by matrix multiplication. But there is also matrix by column multiplication, and column vectors, in this interpretation, are random variables – the classical counterpart of observables to which they are generalized in quantum mechanics. The time behaviour of random variables is dual to the time behaviour of probability distributions where the duality is given by the expectation value, i.e., the scalar product: Indeed, if x is a column vector, then $\langle \pi T, x\rangle = \langle \pi, Tx\rangle$ (note that T has positive, hence real, entries).

In the general case, if K is a Markov kernel on (Ω, Σ) and $f : \Omega \to \mathbb{R}$ is a bounded measurable function then

$$(T_K f)(\omega) := \int_\Omega f(\omega')K(\omega, \mathrm{d}\omega')$$

is again a bounded measurable function and K can be considered as a linear operator T_K on the space of such functions. This operator has obvious properties like $(T_K f) \geq 0$ if $f \geq 0$ and $T_K(\mathbb{1}) = \mathbb{1}$, where $\mathbb{1}$ is the constant function with value 1. In addition, T_K has some (order) continuity property. Under reasonable conditions on (Ω, Σ) (e.g. Ω *Polish*, i.e., homeomorphic to a separable complete metric space), which are satisfied for most cases of interest like for $\Omega = \mathbb{R}^n$, $\mathbb{C}^n$, there is a biunique correspondence between Markov kernels and such operators. The basic advantage of turning from K to T_K is the fact that in many cases

the operator T_K can be restricted to some good Banach space of functions on Ω where techniques from functional analysis can be used. For example, if K is a kernel on $\mathbb{R}^n$ then in most cases of practical interest T_K can be considered as an operator on the space $\tilde{C}_0(\mathbb{R}^n)$ of continuous functions f on $\mathbb{R}^n$ whose limit $\lim_{|x|\to\infty} f(x)$ exists. The operator T_K on these functions is dual to the transformation induced by K on the probability measures: For every such measure μ and for all $f \in \tilde{C}_0(\mathbb{R}^n)$ we obtain

$$\int_{\mathbb{R}^n} T_k f \mathrm{d}\mu = \int_{\mathbb{R}^n} f \mathrm{d}(\mu \circ K) \ .$$

It is worth mentioning that the passage from K to T_K is multiplicative: If L is another such kernel, then $T_{K\circ L} = T_K \cdot T_L$, in particular $T_{K^n} = T_K^n$, for $n \in \mathbb{N}$.

Suppose, finally, that for the Markov kernel K there does exist a transition probability density p. Then T_K can be written as

$$(T_K f)(\omega) = \int_\Omega f(\omega') p(\omega, \omega') \mathrm{d}\omega \ ,$$

i.e., T_K has an integral kernel.

But there are many more cases in which T_K is a well defined operator on $\tilde{C}_0(\mathbb{R}^n)$ although no transition probability density does exist. The cases where jump processes are involved are important examples for such a situation. In the following an operator of the type T_K will be called *transition operator*.

4.2.4 Continuous Time

Although the case of continuous time is not the main focus of these notes we briefly describe the passage from discrete to continuous time. Common to all descriptions of Markovian behaviour in discrete time was the fact that there is a description of a first time step from which the description of the n-th time step was obtained by iteration. The description of n-th time step behaviour in terms of the one time step behaviour is obtained from the Chapman–Kolmogorov equations. In continuous time there is no first time step so we have to consider all times simultaneously and the Chapman–Kolmogorov equations turn into a semigroup law.

Semigroups of Transitions. In the case of a finite state space the transition probabilities between time zero and time $t \geq 0$ are described by a stochastic matrix T_t, and since the composition of time intervals corresponds to matrix multiplication (cf. Sect. 4.1.3) the family $(T_t)_{t\geq 0}$ has the properties

i) $T_0 = \mathbb{1}$, the identity matrix,
ii) $T_{s+t} = T_s \cdot T_t$ for $s, t \geq 0$ (semigroup law).

The semigroup law can be written more impressively: Consider a sequence of time steps

$$0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t \ ,$$

and denote the entries of the matrix T_s by τ_{ij}^s. Then we may write

$$T_t = T_{t-t_{n-1}} \cdot T_{t_{n-1}-t_{n-2}} \cdot \ldots \cdot T_{t_2-t_1} \cdot T_{t_1}$$

and

$$\tau_{i,j}^t = \sum_{k_1,\ldots,k_{n-1}} \tau_{i,k_{n-1}}^{t-t_{n-1}} \cdot \tau_{k_{n-1},k_{n-2}}^{t_{n-1}-t_{n-2}} \cdot \ldots \cdot \tau_{k_2,k_1}^{t_2-t_1} \cdot \tau_{k_1,j}^{t_1} \ .$$

This, again, deserves the name Chapman–Kolmogorov equation.

Turning to the description by transition probability densities we must have a function p_t on $\Omega_0 \times \Omega_0$ for every $t \geq 0$ such that

i) $p_0(x,y) = \delta_{x,y}$,
ii) $p_{s+t}(x,y) = \int_{\Omega_0} p_s(x,z)p_t(z,y)\mathrm{d}z$ for $s,t \geq 0$.

Again, the semigroup property ii) can be split up as above by inserting intermediate time steps.

Similarly, a family K_t, $t \geq 0$, of Markov kernels on (Ω, Σ) describes a Markovian behaviour if

i) $K_0(x,A) = \chi_A(x) = \begin{cases} 1 \text{ , if } x \in A \ , \\ 0 \text{ otherwise} \ , \end{cases}$

ii) $K_{s+t} = K_s \circ K_t$ for $s,t \geq 0$.

Finally, any family $(T_t)_{t\geq 0}$ of positive identity preserving transition operators should satisfy

i) $T_0 = \mathbb{1}$,
ii) $T_{s+t} = T_s \cdot T_t$ for $s,t \geq 0$.

Infinitesimal Generators. Instead of considering transitions for all positive times simultaneously, one would like to turn to a differential description of the infinitesimal time step. From this the time evolution can be characterized as the solution of a differential equation.

To begin with, consider $(T_t)_{t\geq 0}$, a semigroup of stochastic matrices as above. It is a non-trivial fact that in the finite dimensional case continuity in t of this semigroup implies its differentiability. Hence there is an $n \times n$-matrix L, called its *generator*, such that

$$T_t = \mathrm{e}^{Lt} \ ,$$

hence

$$\dot{T}_t := \frac{\mathrm{d}}{\mathrm{d}t} T_t = L \cdot T_t \quad , \quad T_0 = \mathbb{1} \ ,$$

which is a system of ordinary linear differential equations of first order.

Written more explicitly, if π is any initial probability distribution vector and if we put $\pi_t := \pi \cdot T_t$, then the time evolution satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_t = \pi_t L \ , \ \pi_0 = \pi \ .$$

Dually to this a column vector x behaves in time according to

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = L(x_t)\ ,\ x_0 = x\ ,$$

where $x_t := T_t(x)$.

An enormous amount of efforts has been made in the second half of the 20^{th} century to generalize such a differential description from the finite dimensional to more general cases. It is one of the main advantages of the formulation in terms of transition operators that here such a differential description can be obtained in many cases. The starting point of all these considerations is the following fundamental result:

Theorem 1. *Let E be any Banach space and $(T_t)_{t\geq 0}$ be a family of bounded linear operators on E such that*

i) $T_0 = \mathrm{Id}$*, the identity on* E*,*
ii) $T_{s+t} = T_s \cdot T_t$ *for* $s, t \geq 0$*,*
iii) $\lim_{t\downarrow 0} \|T_t x - x\| = 0$ *for all* $x \in E$ *(i.e.,* $(T_t)_{t\geq 0}$ *is strongly continuous).*

Then there exists a – usually unbounded but closed – operator L defined on a dense domain $\mathcal{D} \subseteq E$ such that $x_t := T_t x$ satisfies for all $x \in \mathcal{D}$ the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = Lx_t\ ,\ x_0 = x \qquad (t \geq 0)\ .$$

In this case L is called the *generator* of T_t, and one writes $T_t = \mathrm{e}^{Lt}$. In many cases, however, e^{Lt} has no obvious independent meaning, e.g. it cannot always be defined through a power series. More information on the general theory of semigroups of operators and their generators can be found in many books, for example in [4].

From Generators to Partial Differential Equations. If we specialize now from general semigroups of operators to semigroups of transition operators, a lot more information is available. The following result is a paradigmatic example for this type of information. Let E be the Banach space $\tilde{C}_0(\mathbb{R}^n)$ equipped with the supremum norm as considered in Sect. 4.2.3, and let $(T_t)_{t\geq 0}$ be a semigroup of positive, identity preserving operators satisfying the assumptions of the previous theorem.

Theorem 2. *If $f \in E$ is twice continuously differentiable and we put $f_t := T_t f$ then $\partial_t f_t(x)$ exists on $\mathbb{R}^n$ and satisfies a partial differential equation of the form*

$$\begin{aligned}\tfrac{\partial}{\partial t} f_t(x) &= \textstyle\sum_i a_i(x,t)\, \tfrac{\partial}{\partial x_i} f_t(x) \\ &+ \textstyle\sum_{i,j}\ \tfrac{1}{2}\, b_{ij}(x,t)\ \tfrac{\partial^2}{\partial x_i \partial x_j} f_t(x) \\ &+ \textstyle\int_{\mathbb{R}^n} f_t(y)\mathrm{d}w(y,t)\end{aligned}$$

for suitable functions a_i, b_{ij} and measures $w(\cdot, t)$.

Further inspection of this equation shows that the first summand on the right hand side describes a deterministic drift, the second leads to a diffusion and the third summand describes a superposition of jump processes.

In the literature a similar equation can sometimes be found for transition probability densities under the name *differential Chapman–Kolmogorov equation*. However, as long as jump processes are present, its mathematical meaning is difficult to establish, as it involves differential calculus on δ-functions. On the other hand, without jumps there are resonable situations when such an equation for the transition probability densities $(p_t)_{t\geq 0}$ holds. It then takes a form like

$$\begin{aligned}\frac{\partial}{\partial t}p_t(x,y) = &-\sum_i \frac{\partial}{\partial x_i}(a_i(x,t)\,p_t(x,y)) \\ &+\sum_{i,j}\ \tfrac{1}{2}\ \frac{\partial^2}{\partial x_i\,\partial x_j}\ (b_{ij}(x,t)\cdot p_t(x,y))\end{aligned}$$

which is referred to as the *Fokker–Planck equation*.

From a mathematical point of view it has many advantages to study transition operators on spaces like $\tilde{C}_0(\mathbb{R}^n)$. However, from the point of view of intuition one would like to see probabilities to evolve in time. Indeed, this can be done: Probability measures are elements in the Banach space dual of $\tilde{C}_0(\mathbb{R}^n)$. (The dual of a Banach space E is the space of all continuous linear functionals on E and a probability measure μ on $\mathbb{R}^n$ induces such a functional $\varphi_\mu : \tilde{C}_0(\mathbb{R}^n) \to \mathbb{C} : f \mapsto \int_{\mathbb{R}^n} f\mathrm{d}\mu$.) And the Banach space adjoint of a transition operator T_t on $\tilde{C}_0(\mathbb{R}^n)$ maps probability measures into probability measures. Even better, in many cases probability measures with densities in $L^1(\mathbb{R}^n)$ are mapped into itself and one can consider the adjoints of T_t on $L^1(\mathbb{R}^n)$. Then one can again write down a partial differential equation for this evolution which is of the same type as the equation for f_t above.

4.3 Random Variables and Markov Processes

So far we merely developed a mathematical language in order to describe what we see: A system changing its state randomly. Our vocabulary developed so far is built around the notions of state space, probability, and transition probability. Although many problems in probability can be successfully discussed on this level – the treatment in [7] is an impressive example for this – there are questions which require the introduction of concepts of a more theoretical nature like random variables and path spaces.

4.3.1 Example and Motivation

The Problem. Considering, once again, our introductory example of Sect. 4.1 of a travelling salesman. We noted that every initial probability distribution finally converges to $(1/7, 2/7, 4/7)$. What is the probabilistic meaning of this? One possible answer is that we observe many salesmen who are, at the beginning of

our observation, distributed over the cities according to a given initial distribution. If they are travelling independently of each other according to our assumed transition probabilities then after many days we will find them distributed approximately like $(1/7, 2/7, 4/7)$.

However, one could guess even more: If we follow the route of only one traveller, we expect that he will touch every city rather soon. But according to the Markov property, each day can be considered as a starting point of a new travel. From this point of view one long travel contains many travelling routes, one starting at the first day, one starting with the second, and so on. Thus we can expect that even one salesman will meet, in the course of many days, the cities R, S, and T with a relative frequency of $1/7, 2/7$, and $4/7$, respectively. How can we make such a statement mathematically precise? The first idea might be to say that every route meets these cities with the above relative frequencies. But this is not true, because a route changing only between T and S is not excluded by our laws – it is just very very unlikely. In order to formulate this rigorously we need a probability measure on the set of all possible routes. Then we can hope for a statement of the type that on 'most' routes the various cities are met with the relative frequencies as indicated above.

Construction of a Stochastic Process. The mathematical model which allows to prove such statements is formulated as follows: For $\Omega_0 = \{\mathrm{R}, \mathrm{S}, \mathrm{T}\}$ or for any other finite set Ω_0 we form the space

$$\Omega := \Omega_0^{\mathbb{N}_0}$$

of all one-sided sequences of elements of Ω_0; in our example they represent all possible travelling routes – likely or not. In general, elements of Ω are called *paths* and Ω is the *path space* of the process.

The space Ω is already very large and we cannot assign in general a probability to each of its subsets. Therefore one has to single out a σ-algebra Σ of subsets which is generated by the *cylinder sets*

$$\Lambda_{i_0,\ldots,i_k} := \{\omega \in \Omega : \omega_0 = i_0, \omega_1 = i_1, \ldots, \omega_k = i_k\} ,$$

for any choice of elements $i_0, \ldots, i_k \in \Omega_0$ and any $k \in \mathbb{N}$. In our example it consists of all routes starting with a fixed sequence $(i_0, \ldots, i_k)$ of cities. As in our example we assume that the transition probabilities are given by the entries of a stochastik matrix $T = (\tau_{ij})$.

Theorem 3. *(Kolmogorov) If $\pi_0 = (p_1, \ldots p_n)$ is any initial probability distribution on Ω_0 then there exists a unique probability measure μ on (Ω, Σ) such that*

$$\mu(\Lambda_{i_1\,,\,\ldots\,,\,i_k}) = \pi_0(i_0) \cdot t_{i_0,i_1} \cdot \ldots \cdot t_{i_{k-1},i_k}$$

for all cylinder sets.

This is very intuitive, as the probability of the set of all paths starting with $(i_0, \ldots, i_k)$ is the probability of having a route $(i_0, \ldots, i_k)$. The σ-algebra Σ,

however, contains also other sets which are very different from cylinder sets. Nevertheless their probability is already uniquely determined.

As indicated in the introduction the probability space (Ω, Σ, μ) is a theoretical construct: It is not the probability space of a part of the real world, but an artificial model of it. But suppose that a demon (he is a relative of Maxwell's demon who knows the state of the world completely) tells us that the world is realizing the state $\omega \in \Omega$ of our artificial world then the state of the system Ω_0 at time n can be read out as the value of

$$X_n : \Omega \to \Omega_0 : \omega = (\omega_0, \omega_1, \ldots) \mapsto \omega_n$$

at $\omega \in \Omega$. The map X_n is an example of a *random variable* and the family $(X_n)_{n\in\mathbb{N}_0}$ is an example of a *stochastic process*. If we are not lucky enough to meet such a demon then the random variables X_n are still useful. For example, the distribution induced by X_n on Ω_0 is the correct probability distribution $\pi_0 \cdot T^n$ on Ω_0 after n steps.

As time goes by today's tomorrow is tomorrow's today; that is, if $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$ is any path then ω_1 is the starting point of another path $(\omega_1, \omega_2, \ldots)$ which starts one time step later. Therefore, it makes sense to describe time translation by the left shift on Ω (omitting the first component).

Solving the Problem. The above considerations should motivate the definitions to come. But first we show that (Ω, Σ, μ) is an appropriate mathematical model in order to answer questions like the above on relative frequencies of visits. Indeed, in our example the following holds:

Theorem 4.

$$\lim_{N\to\infty} \frac{1}{N} \left|\{k : 0 \le k \le N-1 : \omega(k) = \mathrm{R}\}\right| = \frac{1}{7}$$

μ-almost surely on $\omega \in \Omega$.

That is, the set of paths $\omega \in \Omega$, where this statement is not correct, has measure zero; therefore, the probability of a 'correct' behaviour is 1, other paths are very, very unlikely.

The theorem says that the probability 1/7 of being in R, averaged over all paths – *space average* – is the same as the probability of being in R when following one path in time – *time average*. It is a very special case of the famous *individual ergodic theorem* of Birkhoff (cf. Sect. 4.3.4). A generalization of this result to the outcomes of a repeated quantum measurement is discussed in Sect. 4.10.

4.3.2 One Random Variable

We proceed to introduce the stochastic language. Throughout the following (Ω, Σ, μ) is a probability space and (Ω_0, Σ_0) is a measurable space.

A measurable mapping $X : (\Omega, \Sigma) \to (\Omega_0, \Sigma_0)$ (i.e., X is a mapping from Ω to Ω_0 which is Σ – Σ_0 measurable) is also called a *random variable*. This is only a

new name for a known mathematical object. The name should appeal to a certain intuition as indicated in the above examples: The random variable X describes the influence of a world (Ω, Σ, μ) onto a state space (Ω_0, Σ_0). In particular, it induces a distribution ν on $(\Omega_0, \Sigma_0) : \nu(A) = \mu(X^{-1}(A)) =: \mathbb{P}(X \in A)$. The last expression should appeal to the intuitive meaning that $\nu(A)$ is the probability that X takes a value in A.

The case of real-valued random variables, i.e., $\Omega_0 = \mathbb{R}$, is of special importance. Characteristic properties of such a random variable are described by its expectation

$$\begin{aligned}\mathbb{E}(X) :&= \int_\Omega X(\omega)\mathrm{d}\mu(\omega)\\ &= \int_{\mathbb{R}} x\mathrm{d}\nu(x)\end{aligned}$$

and its variance

$$\begin{aligned}V(X) :&= \int_\Omega (X - \mathbb{E}(X))^2(\omega)\mathrm{d}\mu(\omega)\\ &= \mathbb{E}\left((X - \mathbb{E}(X))^2\right) .\end{aligned}$$

4.3.3 Two Random Variables

If $X_1, X_2 : (\Omega, \Sigma, \mu) \to (\Omega_0, \Sigma_0)$ are two random variables then the probability $\mathbb{P}(X_1 \in A \text{ and } X_2 \in B)$ that X_1 has a value in $A \in \Sigma_0$ and X_2 has a value in $B \in \Sigma_0$ is given by

$$\mu\left(X_1^{-1}(A) \cap X_2^{-1}(B)\right) .$$

It defines a probability measure $\nu_{1,2}$ on $\Omega_0 \times \Omega_0$, called the *joint distribution* of X_1 and X_2, since it is the distribution of the random variable

$$(X_1, X_2) : \Omega \ni \omega \mapsto (X_1(\omega), X_2(\omega)) \in \Omega_0 \times \Omega_0 .$$

The distributions ν_1 and ν_2 induced by X_1 and X_2 on (Ω_0, Σ_0) are recovered as *marginal distributions* of $\nu_{1,2} : \nu_1(A) = \nu_{1,2}(A \times \Omega_0),\ \nu_2(B) = \nu_{1,2}(\Omega_0 \times B)$. Similarly, joint distributions of n random variables are defined.

A joint distribution $\nu_{1,2}$ on $\Omega_0 \times \Omega_0$ is just a probability measure on this product space. Fixing a set $B \in \Sigma_0$ then $\Sigma_0 \ni A \mapsto \nu_{1,2}(A \times B)$ is a measure on Σ_0 dominated by the marginal distribution ν_1. Therefore, by the Radon–Nikodym theorem there is a density function on Ω_0, call it $\omega \mapsto K(\omega, B)$, such that

$$\nu_{1,2}(A \times B) = \int_A K(\omega, B)\mathrm{d}\nu_1(\omega) ,$$

and it is not difficult to guess that K is a kernel on Ω_0. Conversely, any kernel K on $(\Omega_0, \Sigma_0, \nu_1)$ determines a probability measure $\nu_{1,2}$ on $\Omega_0 \times \Omega_0$ via the above identity.

Therefore, given an initial probability measure ν_1 on (Ω_0, Σ_0) there is a canonical biunique correspondence between

- Probability measures $\nu_{1,2}$ on $\Omega_0 \times \Omega_0$ with first marginal distribution ν_1, i.e., $\nu_1(A) = \nu_{1,2}(A \times \Omega_0)$,

- transition kernels on (Ω_0, Σ_0) with an 'initial distribution' ν_1.

and, in many cases of practical interest,

- transition operators T_K on a space of functions on Ω_0 with an 'initial distribution' ν_1.

4.3.4 Many Random Variables

In the following $\mathbb{T}$ stands for any of the time parameter sets $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{R}_+$, $\mathbb{R}$.

A family of random variables $(X_t)_{t\in\mathbb{T}} : (\Omega, \Sigma, \mu) \to (\Omega_0, \Sigma_0)$ is also called a *stochastic process*. A random variable X_t is interpreted as the description of the influence of a world (Ω, Σ, μ) onto the system (Ω_0, Σ_0) under observation. Therefore, in this description random variables change in time and we are working in a sort of Heisenberg picture. If $\omega \in \Omega$ then $\mathbb{T} \ni t \mapsto X_t(\omega) \in \Omega_0$ is again called a *path* and $\Omega_0^{\mathbb{T}}$ is the *path space* of this process. Under very general circumstances we still obtain a canonical probability measure on the path space $\Omega_0^{\mathbb{T}}$:

Theorem 5. *If Ω_0 is a Polish space and Σ_0 its Borel-σ-algebra then there is a unique probability measure ν on the σ-algebra generated by the cylinder sets of $\Omega_0^{\mathbb{T}}$ which has the following properties: If for $t \in \mathbb{T}$ we define the random variable*

$$Y_t : \Omega_0^{\mathbb{T}} \ni \omega = (\omega_s)_{s\in\mathbb{T}} \mapsto \omega_t \in \Omega_0 \ ,$$

then for any finite set $t_1, \ldots, t_n$ of times the joint distribution of $Y_{t_1}, \ldots, Y_{t_n}$ on $\Omega_0 \times \ldots \times \Omega_0$ is the same as the joint distribution of $X_{t_1}, \ldots, X_{t_n}$.

The stochastic process $(Y_t)_{t\in\mathbb{T}}$ is called the *canonical realization* of $(X_t)_{t\in\mathbb{T}}$, and for most purposes one can work with this canonical realization. Again, we can realize time translation by t by the left shift σ_t on $\Omega_0^{\mathbb{T}}$, and we note that

$$X_0 \circ \sigma_t = X_t \ , \text{ for all } t \in \mathbb{T} \ .$$

A stochastic process $(X_t)_{t\in\mathbb{T}}$ is called *stationary* if joint probabilities do not change under time translation:

$$\mathbb{P}(X_{t_1} \in A_1; \ldots; X_{t_n} \in A_n) = \mathbb{P}(X_{t_1+s} \in A_1; \ldots; X_{t_n+s} \in A_n)$$

for all times $t_1, \ldots t_n$, for all 'time translations' s, and for all measurable subsets $A_1, \ldots, A_n \in \Omega_0$. Equivalently, if $f_1, \ldots f_n$ are bounded measurable functions on Ω_0, then

$$\mathbb{E}(f_1 \circ X_{t_1} \cdot \ldots \cdot f_n \circ X_{t_n}) = \mathbb{E}(f_1 \circ X_{t_1+s} \cdot \ldots \cdot f_n \circ X_{t_n+s}) \ ,$$

for all times $t_1, \ldots, t_n$ and s. In this stationary case the measure ν on the path space defined in the above theorem is invariant under time translations σ_t on the path space. In this situation we can formulate a version of the individual ergodic theorem of G.D. Birkhoff. For convenience we do this in discrete time.

The stationary stochastic process $(Y_n)_n$ is called *ergodic* if whenever $A \subseteq \Omega_0^{\mathbb{N}}$ is measurable and $\sigma^{-1}(A) = A$ then $\nu(A) = 0$ or $\nu(A) = 1$.

Theorem 6. *(Birkhoff) For any* $f \in L^1(\Omega_0^{\mathbb{N}}, \nu)$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} (f \circ \sigma^n)(\omega)$$

exists for ν*-almost all* $\omega \in \Omega_0^{\mathbb{N}}$*, and converges to a translation invariant function.*

If the process is ergodic then the only translation invariant functions are multiples of the identity.

The ergodic theorem in Sect. 4.3.1 is obtained from this general theorem if we put $f(\omega) = 1$ if $\omega(0) = \mathrm{R}$, and $f(\omega) = 0$ in all other cases ($\omega \in \Omega_0^{\mathbb{N}}$). In Sect. 4.10 an extension of this ergodic theorem to trajectories of repeated quantum measurement is discussed.

For more information on the ergodic theory of stationary stochastic processes we refer to [5] and [26].

4.3.5 Conditional Expectations

In order to formulate the Markov property of a stochastic process in such a way that it can be generalized to the quantum context, we need the notion of a conditional expectation.

We consider a probability space (Ω, Σ, μ). Alike some earlier occasions we use the symbol $\mathbb{P}$ when we want to talk about probabilities. For example, $\mathbb{P}(\omega \in B)$ is the probability of finding a state ω in the subset B, hence $\mathbb{P}(\omega \in B) = \mu(B)$. If we have already the information $\omega \in A$ then our expectation for finding $\omega \in B$ changes to the *conditional probability*

$$\mathbb{P}(\omega \in B | \omega \in A) := \mathbb{P}(B|A) := \frac{\mu(A \cap B)}{\mu(A)} \quad \text{, provided that} \quad \mu(A) \neq 0 \; .$$

For fixed A we thus obtain a new probability measure $B \mapsto \mathbb{P}(B|A)$ on (Ω, Σ) which is supported by A.

To be slightly more general, suppose that Σ_0 is a σ-subalgebra of Σ, which is generated by the sets of a partition $\Omega = A_1 \dot\cup \ldots \dot\cup A_n$ of Ω into disjoint subsets $A_1, \ldots, A_n$. Suppose further that we do not know, as before, that ω is in, say A_3, but we know that $\omega \in A_1$ with probability p_1, $\omega \in A_2$ with probability $p_2, \ldots, \omega \in A_n$ with probability p_n, i.e., we know a probability measure ν on Σ_0.

Then we should change the probability measure μ to a measure $\tilde\nu$ given by

$$\tilde\nu(B) = p_1 \cdot \mathbb{P}(B|A_1) + p_2 \cdot \mathbb{P}(B|A_2) + \ldots + p_n \cdot \mathbb{P}(B|A_n) \; .$$

The Radon–Nikodym theorem allows us to do the same in the general situation.

Theorem 7. *Let* (Ω, Σ, μ) *be a probability space,* $\Sigma_0 \subseteq \Sigma$ *a* σ*-subalgebra, and* ν *a probability measure on* (Ω, Σ_0) *which is absolutely continuous with respect to the restriction of* μ *to* Σ_0 *(i.e., if* $\mu(A) = 0$ *then* $\nu(A) = 0$ *for a subset* $A \in \Sigma_0$*).*

Then there is a unique function $g \in L^1(\Omega, \Sigma_0, \mu)$ (in particular, g is measurable with respect to Σ_0) such that

$$\int f \mathrm{d}\nu = \int fg \mathrm{d}\mu \quad \textit{for all} \quad f \in L^\infty(\Omega, \Sigma_0, \mu) \ .$$

Thus, if we define a measure $\tilde{\nu}$ on (Ω, Σ) by

$$\int f \mathrm{d}\tilde{\nu} := \int fg \mathrm{d}\mu \quad \textit{for all} \quad f \in L^\infty(\Omega, \Sigma, \mu)$$

then we obtain an affine injection $\nu \mapsto \tilde{\nu}$ from probability measures on (Ω, Σ_0) as above into the probability measures on (Ω, Σ).

This result can be 'dualized' to the following version.

Theorem 8. *In the situation above, for $f \in L^1(\Omega, \Sigma, \mu)$, there is a unique function $Pf \in L^1(\Omega, \Sigma_0, \mu)$, (in particular, Pf is measurable with respect to Σ_0) such that*

$$\int_A f \mathrm{d}\mu = \int_A Pf \mathrm{d}\mu \quad \textit{for all} \quad A \in \Sigma_0 \ .$$

In the special case of above Pf has to be constant on each of the sets $A_1, \ldots, A_n$.

The map $P : f \mapsto Pf$ is linear, positive, and idempotent. It is called the *conditional expectation* on (Ω, Σ, μ) with respect to Σ_0. When we restrict P from $L^1(\Omega, \Sigma, \mu)$ to $L^2(\Omega, \Sigma, \mu)$ then P is the orthogonal projection from the Hilbert space $L^2(\Omega, \Sigma, \mu)$ onto the subspace $L^2(\Omega, \Sigma_0, \mu)$. Therefore, Pf is the unique element in $L^2(\Omega, \Sigma_0, \mu)$ such that $||f - Pf||$ is minimal, i.e., it is the best approximation (in the quadratic mean) to f that can be found in $L^2(\Omega, \Sigma_0, \mu)$.

An important application of conditional expectations is the following: Consider the case $\Omega = \Omega_0 \times \Omega_0$ with the product σ-algebra $\Sigma := \Sigma_0 \times \Sigma_0$. Then $\Sigma_0 \times \Omega_0 := \{A \times \Omega_0 : A \in \Sigma_0\}$ is a σ-subalgebra and we can consider the conditional expectation P on (Ω, Σ, μ) with respect to $\Sigma_0 \times \Omega_0$. If we denote by $f \otimes \mathbb{1}$ the function $\Omega_0 \times \Omega_0 \ni (\omega_1, \omega_2) \mapsto f(\omega_1)$ for a function f on Ω_0 and, similarly, $\mathbb{1} \otimes f : \Omega_0 \times \Omega_0 \ni (\omega_1, \omega_2) \mapsto f(\omega_2)$, then P maps $L^1(\Omega, \Sigma, \mu)$ into functions of the type $f \otimes \mathbb{1}$. Moreover, if T is the transition operator which we associated in Sect. 4.3.3 with the measure μ on the product space $\Omega_0 \times \Omega_0$, then an easy computation shows

$$P(\mathbb{1} \otimes f) = Tf \otimes \mathbb{1} \, ,$$

for suitable f. Therefore, conditional expectations can be used for obtaining transition operators. This idea will help to define transition operators in the quantum case (cf. Sect. 4.5.3).

4.3.6 Markov Processes

As before, consider a stochastic process $(X_t)_{t \in \mathbb{T}}$ on (Ω, Σ, μ) with values in a state space (Ω_0, Σ_0). The definition of the Markov property should reflect the

idea that what happens after time t depends on what happened at time t, but not on what happened at times before t. A convenient way to formulate this is the following: Consider times $t_1 < t_2 < \ldots < t_n < t$ and $s > 0$, and sets $A_1, \ldots, A_n, A, B \in \Sigma_0$, then

$$\mathbb{P}(X_{t+s} \in B | X_t \in A; X_{t_n} \in A_n; \ldots; X_{t_1} \in A_1) = \mathbb{P}(X_{t+s} \in B | X_t \in A) \; .$$

However, as a basis for a generalization to the quantum context the following reformulation is useful: For $t \in \mathbb{T}$, denote by Σ_t the σ-subalgebra of Σ generated by X_t, i.e., by the sets $X_t^{-1}(A)$ for $A \in \Sigma_0$. Similarly, for $I \subseteq \mathbb{T}$ the σ-algebra Σ_I is generated by $\{X_t : t \in I\}$, and we have the conditional expectations P_t and P_I on (Ω, Σ, μ) with respect to Σ_t and Σ_I. By $t]$ and $[t$ we denote the sets $\{s \in \mathbb{T} : s \leq t\}$ and $\{s \in \mathbb{T} : s \geq t\}$, respectively.

Definition 1. A stochastic process $(X_t)_{t \in \mathbb{T}}$ is a *Markov process* if for all $f \in L^1(\Omega, \Sigma_{[t}, \mu)$ and for all $t \in \mathbb{T}$ we have

$$P_{t]}(f) = P_t(f) \; .$$

As pointed out earlier, to any pair of random variables, hence to the pair (X_0, X_t), there corresponds a transition operator T_t. Now it is an easy consequence of the Markov property that

$$T_{s+t} = T_s \cdot T_t \; , \text{for} \quad s, t \geq 0 \; ,$$

hence the transition operators form a semigroup.

Conversely, if $(T_t)_{t \geq 0}$ is a strongly continuous semigroup of transition operators on a Banach space $\tilde{C}_0(\Omega_0)$, Ω_0 a nice locally compact space, then there exists a unique probability measure μ on $\Omega_0^{\mathbb{T}}$ such that the random variables $(Y_t)_{t \geq 0}$ defined as in Sect. 4.3.4 form a Markov process with transition operators $(T_t)_{t \geq 0}$. Therefore, we end up, at least in all cases of interest, with a canonical one to one correspondence between *phenomenological descriptions* of Markovian behaviour by semigroups of transition operators and descriptions by *Markov processes in their canonical realization.*

4.4 Quantum Mechanics

Our aim is to introduce quantum Markov processes. To this purpose we give in this section a formulation of traditional quantum mechanics. It then needs to be extended in the next section in order to include also the description of classical systems.

4.4.1 The Axioms of Quantum Mechanics

Following the ideas of J.v. Neumann [24] quantum mechanics can be axiomatized as follows:

To a physical system there corresponds a Hilbert space $\mathcal{H}$ such that

1. Pure states of the system are described by unit vectors in $\mathcal{H}$ (up to a phase).
2. Observables of the system are described by self-adjoint operators on $\mathcal{H}$.
3. A measurement of an observable, described by a self-adjoint operator X, on the system in a state described by the unit vector $\xi \in \mathcal{H}$, yields an expectation value $\langle X\xi, \xi\rangle$.
4. If an observable is described by the self-adjoint operator X on $\mathcal{H}$, then the observable obtained from it by changing the scale of the measurement apparatus with a measurable function f is described by the operator $f(X)$ which is obtained from X via the spectral theorem. (If f is a bounded function then $f(X)$ is bounded; therefore, from a theoretical point of view, working with bounded operators suffices.)

From these axioms one can derive large parts of the quantum mechanical formalism. How to find $\mathcal{H}, X$, and ξ, however, is a different question which is not touched in the axioms.

4.4.2 An Example: Two-Level Systems

As a concrete example, consider a quantum mechanical two-level system like a spin-1/2-particle. Then $\mathcal{H} = \mathbb{C}^2$ is two-dimensional, and a standard set of observables is given by the matrices

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

which may be interpreted as describing the measurement of some polarization in x, y, and z-direction, respectively.

Every self-adjoint matrix is a unique, real linear combination of $\mathbb{1}, \sigma_x, \sigma_y, \sigma_z$, and such a matrix

$$\Phi = \alpha \cdot \mathbb{1} + x \cdot \sigma_x + y \cdot \sigma_y + z \cdot \sigma_z = \begin{pmatrix} \alpha + z & x - \mathrm{i}y \\ x + \mathrm{i}y & \alpha - z \end{pmatrix}$$

is a density matrix for a mixed state iff,[1] by definition, $\Phi \geq 0$ and $\mathrm{tr}(\Phi) = 1$, hence iff $\alpha = 1/2$ and $x^2 + y^2 + z^2 \leq 1/4$. Therefore, the convex set of mixed states can be identified with a ball in $\mathbb{R}^3$ (of radius $1/2$ in our parametrization), and the pure states of the system correspond to the extreme points on the surface of this ball.

4.4.3 How Quantum Mechanics is Related to Classical Probability

The formalism of quantum mechanics is not so far from classical probability as it might seem at a first glance. The key for establishing this relation is the *spectral theorem* (cf. [28]):

If X is a self-adjoint operator on a separable Hilbert space then there exist

[1] Read "iff" as "if an only if".

- a probability space (Ω, Σ, μ),
- a real-valued random variable $Y : \Omega \to \mathbb{R}$,
- a unitary $u : \mathcal{H} \to L^2(\Omega, \Sigma, \mu)$,

such that $uXu^* = M_Y$, where M_Y is the multiplication operator acting on $L^2(\Omega, \Sigma, \mu)$ by multiplication with Y. It follows that the *spectrum* $\sigma(X)$ of X is equal to $\sigma(M_Y)$, hence it is given by the essential range of the random variable Y. The function Y can be composed with any further real or complex function f which is defined on the (essential) range of Y, hence on the spectrum of X, and thus we can define the operator

$$f(X) := u^* \cdot M_{f \circ Y} \cdot u$$

for any such function f.

It thus appears that a self-adjoint operator can be identified with a real-valued random variable. There is only one problem: Two self-adjoint operators may not be equivalent to multiplication operators on the same probability space with the same intertwining unitary u. Indeed, a family of self-adjoint operators on $\mathcal{H}$ admits a simultaneous realization by multiplication operators on one probability space if and only if they commute. It is only at this point, the occurence of non-commuting self-adjoint operators, where quantum mechanics separates from classical probability.

As long as only one self-adjoint operator is involved, we can proceed further as in classical probability: A state $\xi \in \mathcal{H}$ induces a probability measure μ_ξ on the spectrum $\sigma(X) \subseteq \mathbb{R}$ which is uniquely characterized by the property

$$\langle f(X)\xi, \xi \rangle = \int_{\mathbb{R}} f(\lambda) \mathrm{d}\mu_\xi(\lambda) \ ,$$

for all bounded measurable functions f on $\mathbb{R}$. The measure μ_ξ is called the *spectral measure* of X with respect to ξ but it may be viewed as the distribution of X: The function $u\xi \in L^2(\Omega, \Sigma, \mu)$ is a unit vector, therefore its pointwise absolute value squared $|u\xi|^2$ is, with respect to μ, the density of a probability measure on (Ω, Σ), and μ_ξ is the distribution of Y with respect to this probability measure. The quantum mechanical interpretation of μ_ξ is given in the next statement.

Proposition 1. *A measurement of an observable X on a system in a state ξ gives a value in $\sigma(X)$, and the probability distribution of these values is given by μ_ξ.*

This result can be deduced from the axioms as follows: Let $f := \chi := \chi_{\sigma(X)^C}$ be the characteristic function of the complement of $\sigma(X)$. By axiom 4 a measurement of $\chi(X)$ yields a value 0 or 1. Therefore, the probability that this measurement gives the value 1 is equal to the expectation of this measurement, hence equal to

$$\langle \chi(X)\xi, \xi \rangle = \langle 0\xi, \xi \rangle = 0 \ .$$

It follows that a measurement of $\chi(X)$ gives 0, hence measuring X gives a value in $\sigma(X)$. More generally, if $A \subseteq \sigma(X)$ then the probability for obtaining from a measurement of X a value in A is the probability to obtain the value 1 in a measurement of $\chi_A(X)$ (again we used the fourth axiom), which is given by

$$\langle \chi_A(X)\xi, \xi\rangle = \int_{\mathbb{R}} \chi_A \mathrm{d}\mu_\xi = \mu_\xi(A) \ .$$

It is obvious that the above proof could have been condensed. But in its present form it shows the use of the fourth axiom more clearly.

Corollary 1. *A measurement of an observable X on a system in a state ξ gives a value in a subset $A \subseteq \sigma(X)$ with certainty iff $1 = \mu_\xi(A) = \langle \chi_A(X)\xi, \xi\rangle$, hence if and only if $\chi_A(X)\xi = \xi$ (note that $\chi_A(X)$ is a spectral projection of X).*

It follows from this that after a measurement of X which gives a value in $A \subseteq \sigma(X)$, the state of the system must have changed to a vector in $\chi_A(X)\mathcal{H}$ since an immediate second measurement of X should now give a value in A with certainty. In this manner one can proceed further to deduce step by step the formalism of quantum mechanics from these axioms.

4.5 Unified Description of Classical and Quantum Systems

In this section we establish a formalism which allows to describe classical systems and quantum systems simultaneously. Additional motivation is given in [17].

4.5.1 Probability Spaces

Observables. In our formulation of the second axiom of quantum mechanics we have been a little bit vague: We left it open how many self-adjoint operators correspond to physical observables. It is this freedom which we are now going to use:

Axiom 2, improved version. There is a *–algebra $\mathcal{A}$ of bounded operators on $\mathcal{H}$ such that the (bounded) observables of the system are described by the self-adjoint operators in $\mathcal{A}$.

Here the word *–*algebra* means: If $x, y \in \mathcal{A}$, then also $x+y$, λx ($\lambda \in \mathbb{C}$), $x \cdot y$, and x^* are elements of $\mathcal{A}$. In the physics literature the adjoint of an element or operator x is frequently denoted by $x^\dagger$ instead of by x^*, but in our general context we prefer to use the mathematician's notation.

$\mathcal{A}$ is called the *algebra of observables* of the system. For simplicity we assume that $\mathcal{A}$ contains the identity $\mathbb{1}$. For mathematical convenience $\mathcal{A}$ is usually assumed to be closed either in the norm – it is then called a C^**-algebra* – or in the strong operator topology – in this case it is called a *von Neumann algebra* or W^**-algebra*.

In a truly quantum situation and when there are only finitely many degrees of freedom one would require that $\mathcal{A} = \mathcal{B}(\mathcal{H})$, the algebra of all bounded operators on $\mathcal{H}$. Indeed, von Neumann in his formulation of quantum mechanics assumed this explicitly and this assumption is known as his *irreducibility axiom* . On the other hand, if (Ω, Σ, μ) is a probability space, then bounded real-valued random variables (they are the classical pendant to observables in quantum mechanics) are functions in $L^\infty(\Omega, \Sigma, \mu)$ and any such function can be viewed as a bounded multiplication operator on $L^2(\Omega, \Sigma, \mu)$. Therefore, classical systems correspond to algebras of the type $L^\infty(\Omega, \Sigma, \mu)$, being viewed as algebras of operators. Moreover, it is a non-trivial fact (cf. [31]) that any commutative von Neumann algebra is isomorphic to some $L^\infty(\Omega, \Sigma, \mu)$ (this generalizes the spectral theorem for self-adjoint operators). Therefore, it is safe to say that classical systems correspond to commutative algebras of observables. If we do not think in probabilistic terms but want to incorporate classical mechanics, then Ω becomes the phase space of the system and the first choice for μ is the Liouville measure on Ω.

States. The next problem is to find a unified description of quantum mechanical states and classical probability measures. The idea is that both give rise to expectation values on observables and we will axiomatize this expectation value. Starting again with quantum mechanics a state given by a unit vector $\xi \in \mathcal{H}$ gives rise to the *expectation functional*

$$\varphi_\xi : \mathcal{B}(\mathcal{H}) \ni x \mapsto \langle x\xi, \xi\rangle \in \mathbb{C}\,.$$

The functional φ_ξ is linear, positive ($\varphi_\xi(x) \geq 0$ if $x \geq 0$) and normalized ($\varphi_\xi(\mathbb{1}) = 1$). More generally, if ρ is a density matrix on $\mathcal{H}$, then

$$\varphi_\rho : \mathcal{B}(\mathcal{H}) \ni x \mapsto \mathrm{tr}(\rho\, x) \in \mathbb{C}$$

still enjoys the same properties. (A *density matrix* or *density operator* ρ on $\mathcal{H}$ is a positive operator ρ such that $\mathrm{tr}(\rho) = 1$ where tr denotes the trace.)

On the other hand, if (Ω, Σ, μ) is a classical probability space, then the probability measure μ gives rise to the expectation functional

$$\varphi_\mu : L^\infty(\Omega, \Sigma, \mu) \ \ni f \mapsto \mathbb{E}(f) = \int_\Omega f d\mu \in \mathbb{C}\,.$$

Again, φ_μ is a linear, positive, and normalized functional on $L^\infty(\Omega, \Sigma, \mu)$. This leads to the following notions.

Definition 2. A *state* on an algebra $\mathcal{A}$ of observables is a positive normalized linear functional

$$\varphi : \mathcal{A} \to \mathbb{C}\,.$$

If φ is a state on $\mathcal{A}$ then the pair $(\mathcal{A}, \varphi)$ is called a *probability space*.

Instead of calling φ a state one could call it probability measure, too, but the name 'state' is widely used nowadays. In order to avoid confusion with classical probability spaces, a pair $(\mathcal{A}, \varphi)$ is sometimes called quantum probability space or non-commutative probability space, although it could describe a classical system and be commutative. Finally it should be noted that under certain continuity conditions a state on $\mathcal{B}(\mathcal{H})$ is induced by a density matrix and a state on $L^\infty(\Omega, \Sigma, \mu)$ comes from a probability measure on (Ω, Σ) (see below).

From the Vocabulary of Operator Algebras. As might become clear from the above the language of *operator algebras* is appropriate when a unified mathematical description of classical systems and quantum systems is needed. Although we reduce the use of this language to a minimum in these notes, it might be helpful to introduce the very basic notions from the vocabulary for operator algebras. For further information we refer to the books on this subject like [31].

As mentioned above, operator algebras can be viewed as *-algebras of bounded operators on some Hilbert space closed either in the operator norm (*C^*-algebra*) or on the strong operator topology (*von Neumann algebra*). Here, operators $(x_i)_{i\in I} \subseteq B(\mathcal{H})$ converge to an operator x in the *strong operator topology* if $(x_i(\xi))_{i\in I}$ converges to $x(\xi)$ for every vector $\xi \in \mathcal{H}$. Therefore, strong operator convergence is weaker than convergence in the operator norm, hence von Neumann algebras are also C*-algebras but von Neumann algebras are 'larger' than C*-algebras. There is also an abstract characterization of C*-algebras as Banach *-algebras for which $\|x^*x\| = \|x\|^2$ for all elements x. Von Neumann algebras are abstractly characterized as C*-algebras which have, as a Banach space, a predual.

A typical example of a commutative C*-algebra is $C(K)$, the algebra of continuous functions on a compact space K, and every commutative C*-algebra with an identity is of this type. A typical example of a commutative von Neumann algebra is $L^\infty(\Omega, \Sigma, \mu)$ (to be very precise, (Ω, Σ, μ) should be a localizable measure space) and every commutative von Neumann algebra is of this type. The algebras M_n of $n \times n$-matrices and more generally $B(\mathcal{H})$ of all bounded operators on a Hilbert space $\mathcal{H}$ are C*-algebras and von Neumann algebras while the algebra of all compact operators on $\mathcal{H}$ is only a C*-algebra whenever $\mathcal{H}$ is not finite dimensional. Other C*-algebras which are interesting from the point of view of physics are the C*-algebras of the canonical commutation relations (CCR) and of the canonical anticommutation relations (CAR) (cf. [6]).

Elements x with $x = x^*$ are called *self-adjoint* as they are represented by self-adjoint operators. It is less obvious that elements of the form x^*x should be called *positive*. If y is an operator on some Hilbert space then y is positive semidefinite if and only if $y = x^*x$ for some operator x. But it was not easy for mathematicians in the 1940s and 1950s to find out that also from an abstract point of view this is the right notion of positivity.

As motivated above a *state* on a C*-algebra $\mathcal{A}$ is abstractly defined as a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ which is positive (in view of the above this means that $\varphi(x^*x) \geq 0$ for all $x \in \mathcal{A}$) and normalized: If $\mathcal{A}$ has an identity and φ is already

positive then this simply means that $\varphi(\mathbb{1}) = 1$, in general it means $\|\varphi\| = 1$. Therefore, a state is an element in the Banach space dual of a C*-algebra $\mathcal{A}$. If $\mathcal{A}$ is even a von Neumann algebra and φ is not only in the dual but in the predual of $\mathcal{A}$ then it is called a *normal* state. There are other characterizations of normal states by continuity or order continuity properties. For the present it is enough to know that a state φ on a commutative von Neumann algebra $L^\infty(\Omega, \Sigma, \mu)$ is normal if and only if there is a function $f_\varphi \in L^1(\Omega, \Sigma, \mu)$ such that $\varphi(g) = \int_\Omega f_\varphi g \mathrm{d}\mu$ for all $g \in L^\infty(\Omega, \Sigma, \mu)$. A state φ on the von Neumann algebra $B(\mathcal{H})$ is normal if and only if there is a density matrix ρ_φ on $\mathcal{H}$ such that $\varphi(x) = tr(\rho_\varphi \cdot x)$ for all $x \in B(\mathcal{H})$.

The mathematical duality between states and observables has its counterpart in the description of quantum systems: Time evolutions are transformations on the space of (normal) states. The Banach space adjoint of such a transformation is a transformation on the dual space of observables. In the language of physics a description of time evolutions on the states is referred to as the *Schrödinger picture* while the *Heisenberg picture* refers to a description in terms of observables. These two descriptions are dual to each other and equivalent from a theoretical point of view. However, spaces of observables have a richer algebraic structure (e.g., observables can be multiplied); therefore, working in the Heisenberg picture can be more convenient although a discussion in the Schrödinger picture is closer to intuition.

4.5.2 Random Variables and Stochastic Processes

In view of the treatment of classical probability in Sects. 4.2 and 4.3 it seems natural to proceed now with the phenomenological description of Markovian behaviour by transition operators. However, in this general context there is no state space Ω_0 such that the system jumps on the points of Ω_0. Even if we generalized points of Ω_0 to pure states on an algebra $\mathcal{A}_0$ of observables then a more general state given by a density matrix cannot be interpreted in a unique manner as a probability on the pure states. The consequence is that there is no direct way to talk about transition probabilities and transition operators. Instead we will introduce transition operators only via conditional expectations in Sects. 4.5.3 and 4.5.4, and in generalization of our discussion in Sect. 4.3.5.

Therefore we here proceed with the introduction of random variables. Unfortunately, the notion of a general random variable seems to be the most abstract and unaccessible notion of quantum probability.

From the foregoing it should be clear that a real-valued random variable is a self-adjoint operator in $\mathcal{A}$. But what happens if one wanted to consider random variables with other state spaces? For example, when studying the behaviour of a two-level system one wants to consider polarization in all space directions simultaneously. In classical probability it is enough to change from $\Omega_0 = \mathbb{R}$ to more general versions of Ω_0 like $\Omega_0 = \mathbb{R}^3$. Now we need an algebraic description of Ω_0 and this is obtained as follows (see [1]).

If $X : (\Omega, \Sigma, \mu) \mapsto \Omega_0$ is a random variable and $f : \Omega_0 \to \mathbb{C}$ is measurable then

$$i_X(f) := f \circ X : (\Omega, \Sigma, \mu) \to \mathbb{C}$$

is measurable. Moreover, $f \mapsto i_X(f)$ is a *–homomorphism from the algebra $\mathcal{A}_0$ of all bounded measurable $\mathbb{C}$-valued functions on Ω_0 into $\mathcal{A} := L^\infty(\Omega, \Sigma, \mu)$ with $i_X(\mathbb{1}) = \mathbb{1}$. (**–homomorphism* means that i_X preserves addition, multiplication by scalars, multiplication, and involution which is complex conjugation in our case). We thus arrive at the following definition.

Definition 3. (from [1]) A *random variable* on $\mathcal{A}$ with values in $\mathcal{A}_0$ is an identity preserving *–homomorphism

$$i : \mathcal{A}_0 \mapsto \mathcal{A} \, .$$

It is confusing that the arrow seems to point into the wrong direction, but this comes from the fact that our description is dual to the classical formulation. Nevertheless our definition describes an influence of $\mathcal{A}$ onto $\mathcal{A}_0$: If the 'world' $\mathcal{A}$ is in a certain state φ then i induces the state $\varphi \circ i$ on $\mathcal{A}_0$ given by $\mathcal{A}_0 \ni x \mapsto \varphi(i(x)) \in \mathbb{C}$. If i comes from a classical random variable X then $\varphi \circ i$ comes from the distribution of X hence it should be called the *distribution* of i also in the general case.

Once having defined the notion of a random variable the definition of a stochastic process is obvious:

Definition 4. A *stochastic process* is a family

$$i_t : \mathcal{A}_0 \to (\mathcal{A}, \varphi) \quad , \quad t \in \mathbb{T} \, ,$$

of random variables (we equipped $\mathcal{A}$ already with a state φ).

Of particular importance in classical probability are stationary processes. In the spirit of our reformulations of classical concepts the following generalizes the classical notion from Sect. 4.3.4.

Definition 5. A *stochastic process* $(i_t)_{t\in\mathbb{T}} : \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ is called *stationary*, if for all $s \geq 0$

$$\varphi(i_{t_1}(x_1) \cdot \ldots \cdot i_{t_n}(x_n)) = \varphi(i_{t_1+s}(x_1) \cdot \ldots \cdot i_{t_n+s}(x_n))$$

with $n \in \mathbb{N}$, $x_1, \ldots, x_n \in \mathcal{A}_0$, $t_1, \ldots, t_n \in \mathbb{T}$ arbitrarily.

As in the classical situation this means that multiple time correlations depend only on time differences. It should also be noted that here it is not sufficient to require the above identity only for time ordered times $t_1 \leq t_2 \leq \ldots \leq t_n$.

Finally, the implementation of time translation by the shift on the path space for processes in standard representation reads as follows:

Definition 6. A process $(i_t)_{t\in\mathbb{T}} : \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ admits a *time translation* if there are *–homomorphisms $\alpha_t : \mathcal{A} \to \mathcal{A}$ $(t \in \mathbb{T})$ such that

i) $\alpha_{s+t} = \alpha_s \circ \alpha_t$, for all $s, t \in \mathbb{T}$,
ii) $i_t = \alpha_t \circ i_0$, for all $t \in \mathbb{T}$.

In most cases, in particular when the process is stationary, such a time translation exists. In the stationary case, it leaves the state φ invariant.

4.5.3 Conditional Expectations

Before we are in the position to formulate a Markov property for a stochastic process we need to talk about conditional expectations. The idea is analogous to the classical framework (cf. Sect. 4.3.5). There we were starting with a probability space (Ω, Σ, μ). Then it occured to us that we obtained some additional information on the probabilities of events in a σ–subalgebra $\Sigma_0 \subseteq \Sigma$, comprised in a probability measure ν on (Ω, Σ_0). This gave rise to improved – conditional – probabilities for all events of Σ given by a probability measure $\tilde{\nu}$ on (Ω, Σ) which extends ν on (Ω, Σ_0).

Similarly we start with a (quantum) probability space (A, φ). Thus if we perfom a measurement of a self-adjoint observable $x \in \mathcal{A}$ we expect the value $\varphi(x)$. Assume again that we gained some additional information about the expectation values of observables in a subalgebra $\mathcal{A}_0$ (for example by an observation), that is we now expect the value $\psi(x)$ for the outcome of a measurement of $x \in \mathcal{A}_0$, where ψ is a state on $\mathcal{A}_0$. As above this should change our expectation for all measurements on $\mathcal{A}$ in an appropriate way, expressed now by a state $\tilde{\psi}$ on $\mathcal{A}$.

Mathematically speaking we should have an extension map Q assigning to each state ψ on $\mathcal{A}_0$ a state $\tilde{\psi} = Q(\psi)$ on $\mathcal{A}$; the map should thus satisfy $Q(\psi)(x) = \psi(x)$ for all $x \in \mathcal{A}_0$. Moreover, if $\psi(x) = \varphi(x)$ for all $x \in \mathcal{A}_0$, that is if there is no additional information, then the state φ should remain unchanged, hence we should require $Q(\psi) = \varphi$ in this case. If we require in addition that Q is an affine map ($Q(\lambda\psi_1 + (1-\lambda)\psi_2) = \lambda Q(\psi_1) + (1-\lambda)Q(\psi_2)$ for states ψ_1 and ψ_2 on $\mathcal{A}_0$ and $0 \leq \lambda \leq 1$) and has some weak continuity properties (weak * continuous if $\mathcal{A}_0$ and $\mathcal{A}$ are C*-algebras) then one can easily show that there exists a unique linear map $P : \mathcal{A} \to \mathcal{A}$ such that $P(\mathcal{A}) = \mathcal{A}_0$, $P^2 = P$, and $||P|| \leq 1$, which has the property $Q(\psi)(x) = \psi(P(x))$ for all states ψ on $\mathcal{A}_0$ and $x \in \mathcal{A}$, hence P is the adjoint of Q. The passage from Q to P means change from a state picture (Schrödinger picture) into the dual observable picture (Heisenberg picture). If $\mathcal{A}_0$ and $\mathcal{A}$ are C*-algebras then such a map P is called a *projection of norm one* and it automatically enjoys further properties: P maps positive elements of $\mathcal{A}$ into positive elements and it has the *module property*

$$P(axb) = aP(x)b$$

for $a, b \in \mathcal{A}_0$, $x \in \mathcal{A}$ (see [31]). Therefore, such a map P is called a *conditional expectation* from $\mathcal{A}$ onto $\mathcal{A}_0$.

From the property $\varphi(P(x)) = \varphi(x)$ for all $x \in \mathcal{A}$ it follows that there is at most one such projection. Indeed, with respect to the scalar product $\langle x, y\rangle_\varphi := \varphi(y^*x)$ induced by φ on $\mathcal{A}$ the map P becomes an orthogonal projection. Therefore, we will talk about *the* conditional expectation $P : (\mathcal{A}, \varphi) \to \mathcal{A}_0$.

Typical examples for conditional expectations are conditional expectations on commutative algebras (cf. Sect. 4.3.5) and *conditional expectations of tensor type*: If $\mathcal{A}_0$ and $\mathcal{C}$ are C*-algebras and ψ is a state on $\mathcal{C}$ then

$$P_\psi : \mathcal{A}_0 \otimes \mathcal{C} \ni x \otimes y \mapsto \psi(y) \cdot x \otimes \mathbb{1}$$

extends to a conditional expectation from the (minimal) tensor product $\mathcal{A} := \mathcal{A}_0 \otimes \mathcal{C}$ onto $\mathcal{A}_0 \otimes \mathbb{1}$. If $\mathcal{A}_0$ and $\mathcal{C}$ are von Neumann algebras and ψ is a normal state on $\mathcal{C}$ then P_ψ can be further extended to a conditional expectation which is defined on the larger 'von Neumann algebra tensor product' of $\mathcal{A}_0$ and $\mathcal{C}$. Sometimes it is convenient to identify $\mathcal{A}_0$ with the subalgebra $\mathcal{A}_0 \otimes \mathbb{1}$ of $\mathcal{A}_0 \otimes \mathcal{C}$ and to call the map defined by $\mathcal{A}_0 \otimes \mathcal{C} \ni x \otimes y \mapsto \psi(y)x \in \mathcal{A}_0$ a conditional expectation, too. From its definition it is clear that P_ψ leaves every state $\varphi_0 \otimes \psi$ invariant when φ_0 is any state on $\mathcal{A}_0$.

In general the existence of a conditional expectation from $(\mathcal{A}, \varphi)$ onto a subalgebra $\mathcal{A}_0$ is a difficult subject. For example, on $\mathcal{A} = M_2$, with a state φ induced from the density matrix

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}$$

there exists P onto

$$\mathcal{A}_0 = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} : a, b \in \mathbb{C} \right\}$$

but the conditional expectation from $(\mathcal{A}, \varphi)$ onto the commutative subalgebra

$$\mathcal{A}_0 = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{C} \right\}$$

does not exist if one still insists on the invariance of φ and $\lambda \neq 1/2$. There is a general theorem due to M. Takesaki [32] which solves the problem of existence of conditional expectations, but we do not need this for the following. It suffices to note that requiring the existence of a conditional expectation can be a strong condition, but from a probabilistic point of view it nevertheless makes sense to maintain this requirement in many situations.

With the help of conditional expectations we can define transition operators as in Sect. 4.3.5: Suppose i_1, i_2: $\mathcal{A}_0 \to (\mathcal{A}, \varphi)$ are two random variables such that i_1 is injective and thus can be inverted on its range. If the conditional expectation $P : (\mathcal{A}, \varphi) \to i_1(\mathcal{A}_0)$ exists then the operator $T : \mathcal{A}_0 \to \mathcal{A}_0$ defined by

$$T(x) := i_1^{-1} P(i_2(x))$$

for $x \in \mathcal{A}_1$ should be called a transition operator. It might be instructive to check that this definition indeed generalizes the considerations at the end of Sect. 4.3.5.

4.5.4 Markov Processes

By using conditional expectations we can now formulate a Markov property which generalizes the Markov property for classical processes: Let $(i_t)_{t\in\mathbb{T}} : \mathcal{A}_0 \to (\mathcal{A}, \varphi)$ be a stochastic process. For $I \subseteq \mathbb{T}$ we denote by $\mathcal{A}_I$ the subalgebra of $\mathcal{A}$ generated by $\{i_t(x) : x \in \mathcal{A}_0, t \in \mathbb{T}\}$. In particular, subalgebras $\mathcal{A}_{t]}$ and $\mathcal{A}_{[t}$ are defined as in the classical context.

Definition 7. The process $(i_t)_{t\in\mathbb{T}}$ is a *Markov process* if for all $t \in \mathbb{T}$ the conditional expectation

$$P_{t]} : (\mathcal{A}, \varphi) \to \mathcal{A}_{t]}$$

exists,

$$\text{for all } x \in \mathcal{A}_{[t} \text{ we have } P_{t]}(x) \in i_t(\mathcal{A}_0) \ .$$

If, in particular, the conditional expectation $P_t : (\mathcal{A}, \varphi) \to i_t(\mathcal{A}_0)$ exists, then this requirement is equivalent to $P_{t]}(x) = P_t(x)$ for all $x \in \mathcal{A}_{[t}$. This parallels the classical definition from Sect. 4.3.6 (where P_t automatically exists).

Clearly, a definition without requiring the existence of conditional expectations would be more general and one can imagine several generalizations. On the other hand, the existence of $P_0 : (\mathcal{A}, \varphi) \to i_0(\mathcal{A}_0) = \mathcal{A}_{\{0\}}$ allows us to define transition operators as above: Assume again, as is almost always the case, that i_0 is injective. Then $i_0(\mathcal{A}_0)$ is an isomorphic image of $\mathcal{A}_0$ in $\mathcal{A}$ on which i_0 can be inverted. Thus we can define the *transition operator* T_t by

$$T_t : \mathcal{A}_0 \to \mathcal{A}_0 : x \mapsto i_0^{-1} P_0 i_t(x) \, .$$

From its definition it is clear that T_t is an identity preserving (completely) positive operator, as it is the composition of such operators. Moreover it generalizes the classical transition operators, and the Markov property again implies the semigroup law

$$T_{s+t} = T_s \cdot T_t \ , \text{for } s, t \geq 0 \ ,$$

while $T_0 = \mathbb{1}$ is obvious from the definition. The derivation of this semigroup law from the Markov property is sometimes called *quantum regression theorem.*

In the classical case we had a converse of this: Any such semigroup comes from a Markov process which is essentially uniquely determined. It is a natural question whether this extends to the general context. Unfortunately, it does not. But there is one good news: For a semigroup on the algebra M_n of complex $n \times n$–matrices there does exist a Markov process which can be constructed on Fock space (cf. Sect. 4.9.3). For details we refer to [25]. However, this Markov process is not uniquely determined by its semigroup as we will see in Sect. 4.7.3. Moreover, if the semigroup $(T_t)_{t\geq 0}$ on $\mathcal{A}_0$ admits a stationary state φ_0, that is, $\varphi_0(T_t(x)) = \varphi_0(x)$ for $x \in \mathcal{A}_0$, $t \geq 0$, then one should expect that it comes from a stationary Markov process as it is the case for classical processes. But here we run into severe problems. They are basically due to the fact that in a truly quantum situation interesting joint distributions – states on tensor products of algebras – do not admit conditional expectations. As an illustration of this kind of problem consider the following situation.

Consider $\mathcal{A}_0 = M_n$, $2 \leq n \leq \infty$. Such an algebra $\mathcal{A}_0$ describes a truly quantum mechanical system. Now consider any random variable $i : \mathcal{A}_0 \to (\mathcal{A}, \varphi)$.

Proposition 2. *The algebra $\mathcal{A}$ decomposes as*

$$\begin{aligned} \mathcal{A} \quad &\simeq M_n \otimes \mathcal{C} \text{ for some algebra } \mathcal{C} \text{ , such that} \\ i(x) &= x \otimes \mathbb{1} \quad \text{for all } x \in \mathcal{A}_0 = M_n \ . \end{aligned}$$

Proof: Put $\mathcal{C} := \{y \in \mathcal{A} : i(x) \cdot y = y \cdot i(x) \text{ for all } x \in \mathcal{A}_0\}$.

Moreover, the existence of a conditional expectation forces the state φ to split, too:

Proposition 3. *If the conditional expectation*

$$P : (\mathcal{A}, \varphi) \to i(\mathcal{A}_0) = M_n \otimes \mathbb{1}$$

exists, then there is a state ψ on $\mathcal{C}$ such that

$$\varphi = \varphi_0 \otimes \psi \ ,$$

i.e., $\varphi(x \otimes y) = \varphi_0(x) \cdot \psi(y)$ for $x \in \mathcal{A}_0$, $y \in \mathcal{C}$ with $\varphi_0(x) := \varphi(x \otimes \mathbb{1})$, and

$$P(x \otimes y) = \psi(y) \cdot x \otimes \mathbb{1} \ ,$$

hence P is a conditional expectation of tensor type (cf. Sect. 4.5.3).

Again, the proof is easy: From the module property of P it follows that P maps $\mathbb{1} \otimes \mathcal{C}$ into the center of M_n, hence onto the multiples of $\mathbb{1}$; thus P on $\mathbb{1} \otimes \mathcal{C}$ defines a state ψ on $\mathcal{C}$.

Therefore, the existence of the conditional expectation $P : (\mathcal{A}, \varphi) \to \mathcal{A}_0$ forces the state to split into a product state, hence the state can not represent a non-trivial joint distribution. As a consequence we have no longer a correspondence between joint distributions and transition operators like in the classical case (cf. Sect. 4.3.3).

4.5.5 Relation to Open Systems

Before we show how to elude this problem we compare our formalism of quantum probability with a standard discussion of open quantum systems. We will find many parallels, with the only difference that the discussion of open systems usually uses the Schrödinger picture, while we work in the Heisenberg picture which is dual to it. The link is that a random variable i identifies $\mathcal{A}_0$ with the observables of an open subsystem of $(\mathcal{A}, \varphi)$.

To be more specific, the description of an open system usually starts with a Hilbert space

$$\mathcal{H} = \mathcal{H}_\mathrm{s} \otimes \mathcal{H}_\mathrm{b} \ .$$

It decomposes into a Hilbert space $\mathcal{H}_\mathrm{s}$ for the open subsystem and a Hilbert space $\mathcal{H}_\mathrm{b}$ for the rest of the system which is usually considered as a bath. Correspondingly, the Hamiltonian decomposes as

$$\mathbb{H} = \mathbb{H}_\mathrm{s} + \mathbb{H}_\mathrm{b} + \mathbb{H}_\mathrm{int} \ ,$$

more precisely,

$$\mathbb{H} = \mathbb{H}_\mathrm{s} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{H}_\mathrm{b} + \mathbb{H}_\mathrm{int}$$

where $\mathbb{H}_\mathrm{s}$ is the free Hamiltonian of the system, $\mathbb{H}_\mathrm{b}$ is the free Hamiltonian of the bath and $\mathbb{H}_\mathrm{int}$ stands for the interaction Hamiltonian. Initially the bath

is usually assumed to be in an equilibrium state, hence its state is given by a density operator ρ_b on $\mathcal{H}_\mathrm{b}$ which commutes with $\mathbb{H}_\mathrm{b}$: $[\rho_\mathrm{b}, \mathbb{H}_\mathrm{b}] = 0$.

Next one can frequently find a sentence similar to "if the open system is in a state ρ_s then the composed system is in the state $\rho_\mathrm{s} \otimes \rho_\mathrm{b}$". The mapping $\rho_\mathrm{s} \mapsto \rho_\mathrm{s} \otimes \rho_\mathrm{b}$ from states of the open system into states of the composed system is dual to a conditional expectation. Indeed, if we denote by $\mathcal{A}_0$ the algebra $\mathcal{B}(\mathcal{H}_\mathrm{s})$ and by $\mathcal{C}$ the algebra $\mathcal{B}(\mathcal{H}_\mathrm{b})$, and if ψ_b on $\mathcal{C}$ is the state induced by ρ_b that is $\psi_\mathrm{b}(y) = \mathrm{tr}_\mathrm{b}(\rho_\mathrm{b} \cdot y)$ for $y \in \mathcal{C}$, then the mapping

$$A_0 \otimes \mathcal{C} \ni x \otimes y \mapsto \psi_\mathrm{b}(y) \cdot x \otimes \mathbb{1}$$

extends to a conditional expectation of tensor type $P = P_{\psi_\mathrm{b}}$ from $\mathcal{A}_0 \otimes \mathcal{C}$ to $\mathcal{A}_0 \otimes \mathbb{1}$ such that

$$\mathrm{tr}_\mathrm{s}(\rho_\mathrm{s}(P(x \otimes y))) = \mathrm{tr}(\rho_\mathrm{s} \otimes \rho_\mathrm{b} \cdot x \otimes y) \ ,$$

where we identified $\mathcal{A}_0 \otimes \mathbb{1}$ with $\mathcal{A}_0$. This duality is an example of the type of duality discussed in Sect. 4.5.3.

A further step in discussing open systems is the introduction of the *partial trace* over the bath: If the state of the composed system is described by a density operator ρ on $\mathcal{H}_\mathrm{s} \otimes \mathcal{H}_\mathrm{b}$ (which will, in general, not split into a tensor product), then the corresponding state of the open system is given by the partial trace $\mathrm{tr}_\mathrm{b}(\rho)$ of ρ over $\mathcal{H}_\mathrm{b}$. The partial trace on a tensor product $\rho = \rho_1 \otimes \rho_2$ of density matrices ρ_1 on $\mathcal{H}_\mathrm{s}$ and ρ_2 on $\mathcal{H}_\mathrm{b}$ is defined as

$$\mathrm{tr}_\mathrm{b}(\rho) = \mathrm{tr}_\mathrm{b}(\rho_1 \otimes \rho_2) := \mathrm{tr}_\mathrm{b}(\rho_2) \cdot \rho_1 \ ,$$

and is extended to general ρ by linearity. It thus has the property

$$\mathrm{tr}(\rho \cdot x \otimes \mathbb{1}) = \mathrm{tr}_\mathrm{s}(\mathrm{tr}_\mathrm{b}(\rho) \cdot x)$$

for all $x \in \mathcal{A}_0$, that is x on $\mathcal{H}_\mathrm{s}$. The partial trace is therefore dual to the map

$$i : \mathcal{B}(\mathcal{H}_\mathrm{s}) \ni x \mapsto x \otimes \mathbb{1} \in \mathcal{B}(\mathcal{H}_\mathrm{s}) \otimes \mathcal{B}(\mathcal{H}_\mathrm{b}) \ ,$$

that is to the random variable i.

The time evolution in the Schrödinger picture is given on a density matrix ρ by $\rho \mapsto u_t \rho u_t^*$ with $u_t = \mathrm{e}^{\mathrm{i}\mathbb{H}t}$. Dual to it is the time evolution in the Heisenberg picture. On an observable x it is given by

$$x \mapsto u_t^* x u_t \ ,$$

and it can be viewed as a time translation α_t of a stochastic process $(i_t)_t$ with $i_t(x) := \alpha_t \circ i(x)$.

Finally, the *reduced time evolution* on the states of the open system maps an initial state ρ_s of this system into

$$\rho_\mathrm{s}(t) := \mathrm{tr}_\mathrm{b}(u_t \cdot \rho_\mathrm{s} \otimes \rho_\mathrm{b} \cdot u_t^*) \ .$$

Thus the map $\rho_{\mathrm{s}} \mapsto \rho_{\mathrm{s}}(t)$ is the composition of the maps $\rho_{\mathrm{s}} \mapsto \rho_{\mathrm{s}} \otimes \rho_{\mathrm{b}}$, $\rho \mapsto u_t \rho u_t^*$, and $\rho \mapsto \mathrm{tr}_{\mathrm{b}}(\rho)$. Hence it is dual to the composition of the maps i, α_t, and P, that is to

$$T_t : \mathcal{A}_0 \mapsto \mathcal{A}_0 : x \mapsto P \circ \alpha_t \circ i(x) = P(i_t(x))$$

which is a transition operator of this stochastic process.

In almost all realistic models this stochastic process as it comes will not have a Markov property. But in order to make this model accessible to computation one performs in most cases a so–called 'Markovian limit'. Mathematically this turns this process into a kind of Markov process. Physically, it turns the system into another system where the heat bath has no memory, that is, it is a kind of white noise.

4.6 Constructing Markov Processes

Let us come back to the discussion of Markov processes. The discussion in Sect. 4.5 seems to indicate that there are no interesting Markov processes in the truly quantum context: On the one hand, we need a conditional expectation onto the time zero algebra $\mathcal{A}_0$ of the process, on the other hand, for $A_0 = M_n$, this forces the state to split into a tensor product, and this prevents the state from representing an interesting joint distribution. In the following we show that there is nevertheless a reasonable way of constructing Markov processes [14]. It avoids the above problem by putting the relevant information about what happens between time steps into the dynamics, instead of putting it into the state:

4.6.1 A Construction Scheme for Markov Processes

We freely use the language introduced in the previous Sect. 4.5. Given a probability space $(\mathcal{A}_0, \varphi_0)$ for the time–zero-algebra of the Markov process, and given a further probability space $(\mathcal{C}_0, \psi_0)$. Then we can form their tensor product

$$(\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}_0, \psi_0) := (\mathcal{A}_0 \otimes \mathcal{C}_0, \varphi_0 \otimes \psi_0) \ ,$$

where $\mathcal{A}_0 \otimes \mathcal{C}_0$ is the tensor product of $\mathcal{A}_0$ and $\mathcal{C}_0$ and $\varphi_0 \otimes \psi_0$ is the product state on $\mathcal{A}_0 \otimes \mathcal{C}_0$ determined by $\varphi_0 \otimes \psi_0(x \otimes y) = \varphi_0(x) \cdot \psi_0(y)$ for $x \in \mathcal{A}_0$, $y \in \mathcal{C}_0$. Finally, let α_1 be any automorphism of $(\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}_0, \psi_0)$ that means that α_1 is an automorphism of the algebra $\mathcal{A}_0 \otimes \mathcal{C}_0$ which leaves the state $\varphi \otimes \psi$ invariant. From these ingredients we now construct a Markov process:

There is also an infinite tensor product of probability spaces. In particular, we can form the infinite tensor product $\bigotimes_{\mathbb{Z}}(\mathcal{C}_0, \psi_0)$: The algebra $\bigotimes_{\mathbb{Z}} \mathcal{C}_0$ is the closed linear span of elements of the form $\cdots \otimes \mathbb{1} \otimes x_{-n} \otimes \cdots \otimes x_n \otimes \mathbb{1} \otimes \cdots$ and the state on such elements is defined as $\psi_0(x_{-n}) \cdot \ldots \cdot \psi_0(x_n)$ for $x_i \in \mathcal{C}_0$, $-n \leq i \leq n$. Then $\bigotimes_{\mathbb{Z}}(\mathcal{C}_0, \psi_0)$ is again a probability space denoted by $(\mathcal{C}, \psi)$, and the tensor right shift extends to an automorphism S of $(\mathcal{C}, \psi)$.

We form the probability space

$$(\mathcal{A}, \varphi) := (\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}, \psi) = (\mathcal{A}_0, \varphi_0) \otimes (\bigotimes_{\mathbb{Z}}(\mathcal{C}_0, \psi_0)) \ ,$$

and identify $(\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}_0, \psi_0)$ with a subalgebra of $(\mathcal{A}, \varphi)$ by identifying $(\mathcal{C}_0, \psi_0)$ with the zero factor $(n = 0)$ of $\bigotimes_{\mathbb{Z}} (\mathcal{C}_0, \psi_0)$. Thus, by letting it act as the identity on all other factors of $\bigotimes_{\mathbb{Z}} (\mathcal{C}_0, \psi_0)$, we can trivially extend α_1 from $(\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}_0, \psi_0)$ to an automorphism of $(\mathcal{A}, \varphi)$, still denoted by α_1. Similarly, S is extended to the automorphism $\mathrm{Id} \otimes S$ of $(\mathcal{A}, \varphi) = (\mathcal{A}_0, \varphi_0) \otimes (\mathcal{C}, \psi)$, acting as the identity on $\mathcal{A}_0 \otimes \mathbb{1} \subseteq \mathcal{A}$. Finally, we define the automorphism

$$\alpha := \alpha_1 \circ (\mathrm{Id} \otimes S) \ .$$

This construction may be summarized in the following picture:

$$\left.\begin{array}{r} (\mathcal{A}_0, \varphi_0) \\ \otimes \\ \cdots \otimes (\mathcal{C}_0, \psi_0) \otimes \ (\mathcal{C}_0, \psi_0) \end{array}\right\} \alpha_1 \otimes (\mathcal{C}_0, \psi_0) \otimes \cdots$$

$$\xrightarrow[S]{}$$

The identification of $\mathcal{A}_0$ with the subalgebra $\mathcal{A}_0 \otimes \mathbb{1}$ of $\mathcal{A}$ gives rise to a random variable $i_0 : \mathcal{A}_0 \to \mathcal{A}$. From i_0 we obtain random variables i_n for $n \in \mathbb{Z}$ by $i_n := \alpha^n \circ i_0$. Thus we obtain a stochastic process $(i_n)_{n \in \mathbb{Z}}$ which admits a time translation α. This process is stationary (α_1 as well as S preserve the state φ) and the conditional expectation $P_0 : (\mathcal{A}, \varphi) \to \mathcal{A}_0$ exists, cf. Sect. 4.5.3.

Theorem 9. *This stochastic process is a stationary Markov process.*

The proof is by inspection: Due to the stationarity of this process it is enough to show that for all x in the future algebra $\mathcal{A}_{[0}$ we have $P_{0]}(x) \in \mathcal{A}_0$. But the algebra $\mathcal{A}_{[0}$ is obviously contained in

$$\begin{array}{c} (\mathcal{A}_0, \varphi_0) \\ \otimes \\ \cdots \otimes \mathbb{1} \otimes \ (\mathcal{C}_0, \psi_0) \ \otimes (\mathcal{C}_0, \psi_0) \otimes \cdots \end{array}$$

while the past $\mathcal{A}_{0]}$ is contained in

$$\begin{array}{c} (\mathcal{A}_0, \varphi_0) \\ \otimes \\ \cdots \otimes (\mathcal{C}_0, \psi_0) \otimes \quad \mathbb{1} \quad \otimes \mathbb{1} \otimes \cdots \end{array}$$

Discussion. We remark that this construction can also be carried out in the special case where all algebras are commutative. It then gives a construction scheme for classical Markov processes , which is independent from its canonical realization on the space of its paths. It is not difficult to show that every classical, discrete time stationary Markov process can be obtained in this way. However, this process may not be minimal, i.e., $\mathcal{A}_{\mathbb{Z}}$ may be strictly contained in $\mathcal{A}$.

It is clear that a Markov process in this class is already determined by giving the probability space $(\mathcal{C}_0, \psi_0)$ and the automorphism α_1. In particular, the transition operator can be computed as $T(x) = P_0 \circ \alpha_1(x \otimes \mathbb{1})$ for $x \in \mathcal{A}_0$.

Conversely, given a transition operator T of $(\mathcal{A}_0, \varphi_0)$, if one wants to construct a corresponding Markov process, then it is enough to find $(\mathcal{C}_0, \psi_0)$ and α_1 as above. This makes the problem easier, compared to the originial problem of guessing the whole Markov process, but it is by no means trivial. In fact, there is no universal scheme for finding $(\mathcal{C}_0, \psi_0)$ and α_1, and there are some difficult mathematical problems associated with their existence.

We finally remark that for $\mathcal{A}_0 = M_n$ this form of a Markov process is typical. In fact there are theorems which show that in this case an arbitrary Markov process has such a structure: It is always a *coupling of* $\mathcal{A}_0$ *to a shift system*. More information on this and further references can be found in [15].

4.6.2 Other Types of Markov Processes in the Literature

Phase Space Methods. In the literature on quantum optics one can frequently find a different approach to quantum stochastic processes: If the system under observation is mathematically equivalent to a system of one or several quantum harmonic oscillators – as it is the case for one or several modes of the quantized electromagnetic field – then phase space representations are available for the density matrices of the system. The most prominent such representations are the P–representation, the Wigner–representation, and the Q–representation (there do exist other such representations, even for other quantum systems). The idea is to represent a state by a density function, a measure, or a distribution on the phase space of the corresponding classical physical system. These density functions are interpreted as classical probability distributions although they are not always positive.

This gives a tool to take advantage of the ideas of classical probability: If $(T_t)_{t\geq 0}$ on $\mathcal{A}_0$ is a semigroup of transition operators it induces a time evolution $\rho \mapsto \rho_t$ on the density operators and thus on the corresponding densities on phase space. With a bit of luck this evolution can be treated as if it were the evolution of probabilities of a classical Markov process and the machinery of partial differential equations can be brought into play (cf. Sect. 4.2.4). It should be noted, however, that a phase space representation does not inherit all properties from the quantum Markov process. It is a description of Markovian behaviour on the level of a phenomenological description in the sense of Sect. 4.2 and it can not be used to obtain a representation of the Markov process on the space of its paths.

Markov Processes with Creation and Annihilation Operators. In the literature a Markov process for an open quantum system as in Sect. 4.5.5 may also be given by certain families $(A_t^*)_t$ and $(A_t)_t$ of creation and annihilation operators. The relation to our description is the following: If the open system has an algebra $\mathcal{A}_0$ of observables which contains an annihilation operator A_0, then a Markovian time evolution α_t of the composed systems applies, in particular, to A_0 and gives A_t. Sometimes the operators $(A_t)_t$ can be obtained by solving a quantum stochastic differential equation (cf. Sect. 4.9.3).

4.6.3 Dilations

The relation between a Markov process with time translations $(\alpha_t)_t$ on $(\mathcal{A}, \varphi)$ and its semigroup $(T_t)_t$ of transition operators on $\mathcal{A}_0$ can be brought into the form of a diagram:

$$\begin{array}{ccc} \mathcal{A}_0 & \xrightarrow{T_t} & \mathcal{A}_0 \\ {\scriptstyle i_0}\downarrow & & \uparrow{\scriptstyle P} \\ (\mathcal{A}, \varphi) & \xrightarrow[\alpha_t]{} & (\mathcal{A}, \varphi) \end{array} \quad .$$

This diagram commutes for all $t \geq 0$.

From this point of view the Markovian time evolution $(\alpha_t)_t$ appears as an extension of the *irreversible* time evolution $(T_t)_t$ on $\mathcal{A}_0$ to an evolution of *-homomorphisms on the large algebra $\mathcal{A}$. Such an extension is referred to as a *dilation* of $(T_t)_t$ to $(\alpha_t)_t$. The paradimatic dilation theory is the theory of unitary dilations of contraction semigroups on Hilbert spaces defined by the commuting diagram

$$\begin{array}{ccc} \mathcal{H}_0 & \xrightarrow{T_t} & \mathcal{H}_0 \\ {\scriptstyle i_0}\downarrow & & \uparrow{\scriptstyle P_0} \\ \mathcal{H} & \xrightarrow[u_t]{} & \mathcal{H} \end{array} \quad .$$

Here $(T_t)_{t\geq 0}$ is a semigroup of contractions on a Hilbert space $\mathcal{H}_0, (U_t)_t$ is a unitary group on a Hilbert space $\mathcal{H}$, $i_0 : \mathcal{H}_0 \to \mathcal{H}$ is an isometric embedding, and P_0 is the Hilbert space adjoint of i_0 which may be identified with the orthogonal projection from $\mathcal{H}$ onto $\mathcal{H}_0$. The diagram has to commute for all $t \geq 0$.

There is an extensive literature on unitary dilations starting with the pioneering books [30] and [22], and it turned out to be fruitful to look at Markov processes and open systems from the point of view of dilations, like in [6] and [14].

4.7 An Example on M_2

In this section we discuss Markov processes for the simplest non-commutative case. It has a physical interpretation in terms of a spin-1/2-particle in a stochastic magnetic field. More information on this example can be found in [13]. A continuous time version of this example is discussed in [21].

4.7.1 The Example

We put $\mathcal{A}_0 := M_2$ and $\varphi_0 := \mathrm{tr}$, the tracial state on M_2. If $(\mathcal{C}_0, \psi_0)$ is any further probability space then the algebra $M_2 \otimes \mathcal{C}$ is canonically isomorphic to the algebra $M_2(\mathcal{C})$ of 2×2-matrices with entries in $\mathcal{C}$: The element

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \otimes \mathbb{1} \in M_2 \otimes \mathcal{C}$$

corresponds to

$$\begin{pmatrix} x_{11} \cdot \mathbb{1} & x_{12} \cdot \mathbb{1} \\ x_{21} \cdot \mathbb{1} & x_{22} \cdot \mathbb{1} \end{pmatrix} \in M_2(\mathcal{C}) \ ,$$

while the element $\mathbb{1} \otimes c \in M_2 \otimes \mathcal{C}$ $(c \in \mathcal{C})$ corresponds to

$$\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \in M_2(\mathcal{C}) \ .$$

Accordingly, the state $\mathrm{tr} \otimes \psi$ on $M_2 \otimes \mathcal{C}$ is identified with

$$M_2(\mathcal{C}) \ni \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \mapsto 1/2(\psi(c_{11}) + \psi(c_{22}))$$

on $M_2(\mathcal{C})$, and the conditional expectation P_0 from $(M_2 \otimes \mathcal{C}, \mathrm{tr} \otimes \psi)$ onto $M_2 \otimes \mathbb{1}$ reads as

$$M_2(\mathcal{C}) \ni \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \mapsto \begin{pmatrix} \psi(c_{11}) & \psi(c_{12}) \\ \psi(c_{21}) & \psi(c_{22}) \end{pmatrix} \in M_2 \ ,$$

when we identify $M_2 \otimes \mathbb{1}$ with M_2 itself.

In Sect. 4.6.1 we saw: Whenever we have a non-commutative probability space $(\mathcal{C}_0, \psi_0)$ and an automorphism α_1 of $(M_2 \otimes \mathcal{C}_0, \mathrm{tr} \otimes \psi_0)$, we can extend this to a stationary Markov process. We begin with the simplest possible choice for $(\mathcal{C}_0, \psi_0)$: Put $\Omega_0 := \{-1, 1\}$ and consider the probability measure μ_0 on Ω_0 given by $\mu_0(\{-1\}) = 1/2 = \mu_0(\{1\})$. The algebra $\mathcal{C}_0 := L^\infty(\Omega_0, \mu_0)$ is just $\mathbb{C}^2$ and the probability measure μ_0 induces the state ψ_0 on $\mathcal{C}_0$ which is given by $\psi_0(f) = 1/2 f(-1) + 1/2 f(1)$ for a vector $f \in \mathcal{C}_0$.

In this special case, there is still another picture for the algebra $M_2 \otimes \mathcal{C}_0 = M_2 \otimes \mathbb{C}^2$: It can canonically be identified with the direct sum $M_2 \oplus M_2$ in the following way: When elements of $M_2 \otimes \mathcal{C}_0 = M_2(\mathcal{C}_0)$ are written as 2×2-matrices with entries f_{ij} in $\mathcal{C}_0 = L^\infty(\Omega_0, \mu_0)$, then an isomorphism is given by

$$M_2(\mathcal{C}_0) \to M_2 \oplus M_2 : \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \mapsto \begin{pmatrix} f_{11}(-1) & f_{12}(-1) \\ f_{21}(-1) & f_{22}(-1) \end{pmatrix} \oplus \begin{pmatrix} f_{11}(1) & f_{12}(1) \\ f_{21}(1) & f_{22}(1) \end{pmatrix} .$$

Finally, we need to define an automorphism α_1: We introduce the following notation: A unitary u in an algebra $\mathcal{A}$ induces an *inner automorphism* $Ad\, u : \mathcal{A} \to \mathcal{A} : x \mapsto u^* \cdot x \cdot u$. For any real number ω we define the unitary

$$w_\omega := \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\omega} \end{pmatrix} \in M_2 \ .$$

It induces the inner automorphism

$$Ad\, w_\omega : M_2 \to M_2 : \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto \begin{pmatrix} x_{11} & x_{12}\mathrm{e}^{\mathrm{i}\omega} \\ x_{21}\mathrm{e}^{-\mathrm{i}\omega} & x_{22} \end{pmatrix} \ .$$

Now, for some fixed ω define the unitary $u := w_{-\omega} \oplus w_\omega \in M_2 \oplus M_2 = M_2 \otimes \mathcal{C}_0$. It induces the automorphism $\alpha_1 := Ad\, u$ which is given by $Ad\, w_{-\omega} \oplus Ad\, w_\omega$ on $M_2 \oplus M_2$.

To these ingredients there corresponds a stationary Markov process as in Sect. 4.6.1. From the above identifications it is immediate to verify that the corresponding one–step transition operator is given by

$$T : M_2 \to M_2 : x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto P_0 \circ \alpha_1(x \otimes \mathbb{1}) = \begin{pmatrix} x_{11} & x_{12}\rho \\ x_{21}\rho & x_{22} \end{pmatrix} ,$$

where $\rho = 1/2(\mathrm{e}^{\mathrm{i}\omega} + \mathrm{e}^{-\mathrm{i}\omega}) = \cos(\omega)$.

4.7.2 A Physical Interpretation: Spins in a Stochastic Magnetic Field

We now show that this Markov process has a natural physical interpretation: It can be viewed as the description of a spin-1/2-particle in a stochastic magnetic field. This system is basic for nuclear magnetic resonance.

Spin Relaxation. We interpret the matrices σ_x, σ_y, and σ_z in M_2 as observables of (multiples of) the spin component of a spin-1/2-particle in x, y, and z direction, respectively (cf. Sect. 4.4.2). If a probe of many spin-1/2-particle is brought into an irregular magnetic field in z-direction, one finds that the temporal behaviour of this probe is described by the semigroup of operators on M_2 given by

$$T_t : M_2 \to M_2 : x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto \begin{pmatrix} x_{11} & x_{12} \cdot \mathrm{e}^{-1/2\bar{\lambda}t} \\ x_{21} \cdot \mathrm{e}^{-1/2\lambda t} & x_{22} \end{pmatrix} ,$$

where the real part of λ is greater than zero.

When we restrict to discrete time steps and assume λ to be real (in physical terms this means that we change to the interaction picture), then this semigroup reduces to the powers of the single transition operator

$$T : M_2 \to M_2 : x = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto \begin{pmatrix} x_{11} & \rho \cdot x_{12} \\ \rho \cdot x_{21} & x_{22} \end{pmatrix} ,$$

with some ρ, $0 \leq \rho \leq 1$. This is just the operator for which we constructed the Markov process in the previous section. We see that polarization in z-direction remains uneffected, while polarization in x-direction and y-direction dissipates to zero. We want to see whether our Markov process gives a reasonable physical explanation for the observed relaxation:

A Spin-1/2-Particle in a Magnetic Field. A spin-1/2-particle in a magnetic field B in z-direction is described by the Hamiltonian $\mathbb{H} = eB \cdot \sigma_z/2m = \omega \cdot \sigma_z/2$, where e is the electric charge and m the mass of the particle. ω is called the Larmor–frequency: The time evolution, given by $\mathrm{e}^{-\mathrm{i}\mathbb{H}t}$, describes a rotation of the spin–particle around the z-axis with this frequency:

$$Ad\,\mathrm{e}^{-\mathrm{i}\mathbb{H}t}(\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}) = \begin{pmatrix} x_{11} & \mathrm{e}^{\mathrm{i}\omega t}x_{12} \\ \mathrm{e}^{-\mathrm{i}\omega t}x_{21} & x_{22} \end{pmatrix} .$$

Since we consider the situation in discrete time steps, we consider the unitary

$$\bar{w}_\omega := \mathrm{e}^{-\mathrm{i}\mathbb{H}} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\omega/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\omega/2} \end{pmatrix} .$$

It describes the effect of the time evolution after one time unit in a field of strength B. Note that $Ad\,\bar{w}_\omega = Ad\,w_\omega$ with

$$w_\omega = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\omega} \end{pmatrix}$$

as in Sect. 4.7.1.

A Spin-1/2-Particle in a Magnetic Field with Two Possible Values. Imagine now that the magnetic field is constant during one time unit, and that it has always the same absolute value $|B|$ such that $\cos\omega = \rho$, but it points into $+z$-direction and $-z$-direction with equal probability $1/2$. Representing the two possible states of the field by the points in $\Omega_0 = \{+1, -1\}$, the magnetic field is described by the probability space $(\Omega_0, \mu_0) = (\{+1, -1\}, (1/2, 1/2))$ as in the previous section. Its algebraic description leads to $(\mathcal{C}_0, \psi_0)$, as in Sect 6.1.

The spin-1/2-particle is described by the algebra of observables $\mathcal{A}_0 = M_2$, and if we assume that we know nothing about its polarization, then its state is appropriately given by the tracial state tr on M_2 (it is also called the "chaotic state").

Therefore, the system, which is composed of a spin-1/2-particle and of a magnetic field with two possible values, has just $M_2 \otimes \mathcal{C}_0$ as its algebra of observables. We use the identification of this algebra with $M_2 \oplus M_2$ as in Sect. 4.7.1.

The point $-1 \in \Omega_0$, hence also the first summand of $M_2 \oplus M_2$, corresponds to the field in $-z$-direction. Hence the time evolution on this summand is given by $Ad\,\bar{w}_{-\omega} = Ad\,w_{-\omega}$. On the second summand it is accordingly given by $Ad\,\bar{w}_\omega = Ad\,w_\omega$. Therefore, the time evolution of the whole composed system is given by the automorphism $\alpha_1 = Ad\,w_{-\omega} \oplus Ad\,w_\omega$ on $(M_2 \otimes \mathcal{C}_0, \mathrm{tr} \otimes \psi_0)$. Thus we recovered all the ingredients of the previous section, which we needed in order to extend this to a Markov process.

A Spin-1/2-Particle in a Stochastic Magnetic Field. What is the interpetation of the whole Markov process? Denote, as in Sect. 4.6.1, by $(\mathcal{C}, \psi)$ the infinite tensor product of copies of $(\mathcal{C}_0, \psi_0)$, and by S the tensor right shift on it. Then $(\mathcal{C}, \psi)$ is the algebraic description of the classical probability space (Ω, μ), whose points are two-sided infinite sequences of -1's and 1's, equipped with the product measure constructed from $\mu_0 = (1/2, 1/2)$. The tensor right shift S is induced from the left shift on these sequences. This is the classical Bernoulli–process, which describes, e.g., the tossing of a coin – or the behaviour of a stochastic magnetic field with two possible values $+B$ or $-B$, which are taken according to the outcome of the coin toss. Thus $(\mathcal{C}, \psi, S)$ is the mathematical model of such a stochastic magnetic field in time. Its time zero-component is

coupled to the spin-1/2-particle via the interaction–automorphism α_1. Finally, the whole Markov process describes the spin-1/2-particle interacting with this stochastic magnetic field.

This is precisely how one explains the spin relaxation T: The algebra M_2 of spin observables stands for a large ensemble of many spin-1/2-particles. Assume, for example, that at time zero they all point into the $-x$-direction, so one measures a macroscopic magnetic moment in this direction. Now they feel the above stochastic magnetic field in z-direction. In one time unit, half of the ensemble feels a field in $-z$-direction and starts to rotate around the z-axis, say clockwise; the other half feels a field in $+z$-direction and starts to rotate counterclockwise. Therefore, the polarization of the single spins gets out of phase and the overall polarization in x-direction after one time step reduces by a factor of ρ. Altogether, the change of polarization is appropriately described by T. After another time unit, cards are shuffeld again: another two halfs of particles, stochastically independ of the previous ones, feel the magnetic fields in $-z$-direction and $+z$-direction. The overall effect in polarization is now given by T^2, and so on. Obviously, this explanation is precisely reflected by the structure of our Markov process.

4.7.3 Further Discussion of the Example

The idea in constructing our example in Sect. 4.7.1 depended on writing the transition operator T as a convex combination of the two automorphisms $Ad\, w_{-\omega}$ and $Ad\, w_\omega$. It thus can be generalized: In fact, whenever a transition operator of a probability space $(\mathcal{A}_0, \varphi_0)$ is a convex combination of automorphisms of $(\mathcal{A}_0, \varphi_0)$ or even a convex integral of such automorphisms, a Markov process can be constructed in a similiar way [14]. There is a generalization of this idea to continuous time, which is worked out in [16].

We do not want to enter into such generality here. But it is worth going at least one small step further into this direction: There are many more ways of writing T as a convex combination of automorphisms of M_2: Let μ_0 be any probability measure on the intervall $[-\pi, \pi]$ such that $\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega} d\mu_0(\omega) = \rho$. Obviously, there are many such probability measures: When we identify the intervall $[-\pi, \pi]$ canonically with the unit circle in the complex plane, and μ_0 with a probability measure on it, this simply means that the barycenter of μ_0 is ρ. Then it is clear that $T = \int_{-\pi}^{\pi} Ad\, w_\omega d\mu_0(\omega)$, i.e., it is a convex integral of these automorphisms. To this representation of T there correspond $(\mathcal{C}_0, \psi_0)$ and α_1 as follows: Put $\mathcal{C}_0 := L^\infty([-\pi, \pi], \mu_0)$ and let ψ_0 be the state on $\mathcal{C}_0$ induced by μ_0. The function $[-\pi, \pi] \ni \omega \mapsto \mathrm{e}^{\mathrm{i}\omega}$ defines a unitary v in $\mathcal{C}_0$. It gives rise to a unitary

$$u := \begin{pmatrix} \mathbb{1} & 0 \\ 0 & v \end{pmatrix} \in M_2(\mathcal{C}_0) \cong M_2 \otimes \mathcal{C}_0 \ ,$$

and thus to an automorphism $\alpha_1 := Ad\, u$ of $(M_2 \otimes \mathcal{C}_0, \mathrm{tr} \otimes \psi_0)$. Our example of Sect. 4.7.1 is retained when choosing $\mu_0 := (\delta_{-\omega} + \delta_\omega)/2$, where δ_x is the Dirac measure in the point x (obviously, it was no restriction to assume $\omega \in [-\pi, \pi]$).

In this way for any such μ we obtain a Markov process for the same transition operator T. It is easy to see that there are uncountably many non-equivalent Markov processes of this type. This is in sharp contrast to the classical theory of Markov processes. There, up to equivalence, a Markov process is uniquely determined by its semigroup of transition operators. On the other hand, our discussion of the physical interpretation in the previous section shows that these different Markov processes are not artificial; they rather correspond to different physical situations: The probability measure μ_0 on the points ω appears as a probability measure on the possible values of the magnetic fields. It was rather artificial when we first assumed that the field B can attain only two different values of equal absolute value. Now we can desribe any stochastic magnetic field in z-direction as long as it has no memory in time.

There are even non-commutative Markov processes for a classical transition operator which are contained in these examples: The algebra M_2 contains the two-dimensional commutative subalgebra generated by the observable σ_x, and the whole Markov process can be restricted to the subalgebra generated by the translates of this observable. This gives a Markov process with values in the two-dimensional subalgebra $\mathbb{C}^2$, which still is non-commutative for certain choices of μ_0. Thus we have also non-commutative processes for a classical transition matrix. Details may be found in [14].

4.8 Completely Positive Operators

4.8.1 Complete Positivity

A physical system is described by its algebra $\mathcal{A}$ of observables. As before we assume that $\mathcal{A}$ is, at least, a C^*–algebra of operators on some Hilbert space and we can always assume that $\mathbb{1} \in \mathcal{A}$.

States were defined as linear functionals $\varphi : \mathcal{A} \to \mathbb{C}$ with $\varphi \geq 0$ (that is $\varphi(x) \geq 0$ for $x \in \mathcal{A}, x \geq 0$) and $\varphi(\mathbb{1}) = 1$. They are interpreted as either physical states of the system or as probability measures. All time evolutions and other 'operations' which we considered so far had the property of carrying states into states. This was necessary in order to be consistent with their physical or probabilistic interpretation. In the Heisenberg picture these 'operations' are described by operators on algebras. For an operator $T : \mathcal{A} \to \mathcal{B}$, $\mathcal{A}, \mathcal{B}$ C^*–algebras, the following two conditions are obviously equivalent

a) T is state preserving: For every state φ on $\mathcal{B}$ the functional

$$\varphi \circ T : \mathcal{A} \ni x \mapsto \varphi(T(x))$$

on $\mathcal{A}$ is a state, too.

b) T is positive and identity preserving: $T(x) \geq 0$ for $x \in \mathcal{A}$, $x \geq 0$, and $T(\mathbb{1}) = \mathbb{1}$.

Indeed, all operators which we considered so far did have this property.

As a matter of fact they all turn out to satisfy an even stronger notion of positivity, called *complete positivity*.

Definition 8. An operator $T : \mathcal{A} \to \mathcal{B}$ is *n–positive* if

$$T \otimes \mathrm{Id}_n : \mathcal{A} \otimes M_n \to \mathcal{B} \otimes M_n : x \otimes y \mapsto T(x) \otimes y$$

is positive. It is *completely positive* if T is n-positive for all $n \in \mathbb{N}$.

Elements of $\mathcal{A} \otimes M_n$ may be represented as $n \times n$–matrices with entries from $\mathcal{A}$, like in Sect. 4.7.1, and in this representation the operator $T \otimes \mathrm{Id}_n$ appears as the map which carries such an $n \times n$–matrix $(x_{ij})_{i,j}$ into $(T(x_{ij}))_{i,j}$ with $x_{ij} \in \mathcal{A}$. From the definition it is clear that 1–positivity is just positivity and $(n+1)$–positivity implies n–positivity (in the matrix representation $\mathcal{A} \otimes M_n$ can be identified with $n \times n$–matrices in the upper left corner of all $(n+1) \times (n+1)$–matrices).

It is a non–trivial theorem that for commutative $\mathcal{A}$ or commutative $\mathcal{B}$ positivity already implies complete positivity (cf. [31, IV. 3]). If $\mathcal{A}$ and $\mathcal{B}$ are both non-commutative algebras, this is no longer true. The simplest typical example for this is the transposition on the (complex) 2×2–matices M_2:

$$M_2 \ni \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M_2$$

is positive but not 2-positive, hence not completely positive. From this example one can proceed further to show that for all n there are maps which are n–positive but not $(n+1)$–positive. It is true, however, that on M_n n–positivity already implies complete positivity.

It is an important property of 2-positive hence of completely positive maps that they satisfy a Schwarz-type inequality:

$$T(x^*x) \geq T(x)^*T(x)$$

for $x \in \mathcal{A}$ (the property $T(x^*) = T(x)^*$ follows from positivity).

It can be shown that *–homomorphisms and conditional expectations are automatically completely positive. All operators which we considered so far are either of these types or are compositions of such operators; therefore, they are all completely positive. This is the mathematical reason why we met only completely positive operators. One could wonder, however, whether there is a physical reason for this.

4.8.2 Interpretation of Complete Positivity

In the introduction to this paragraph we argued that time evolutions should be described by positive identity preserving operators. Now suppose that T is such a time evolution on a system $\mathcal{A}$ and S is a time evolution of a different system $\mathcal{B}$. Even if these systems have nothing to do with each other we can consider them – if only in our mind – as parts of the composed system $\mathcal{A} \otimes \mathcal{B}$ whose time evolution should be given by $T \otimes S$ – there is no interaction. As the time evolution of a physical system the operator $T \otimes S$ should be identity preserving

and positive, too. This is, however, not automatic: Already for the simple case of $\mathcal{B} = M_n$ and $S = \mathrm{Id}$ there are counter-examples as mentioned above. Complete positivity is the right notion which avoids this problem: Indeed, if $T : \mathcal{A}_1 \to \mathcal{A}_2$ and $S : \mathcal{B}_1 \to \mathcal{B}_2$ are completely positive operators then $T \otimes S$ can be defined uniquely on the so called minimal tensor product $\mathcal{A}_1 \otimes \mathcal{B}_1$ and it becomes a completely positive operator from $\mathcal{A}_1 \otimes \mathcal{B}_1$ into $\mathcal{A}_2 \otimes \mathcal{B}_2$. For this theorem and related results we refer to the literature, for example ([31] IV. 3, IV. 4 and IV. 5). Rephrasing this result it means that it suffices to require that T preserves its positivity property when it is tensored with the very simple maps Id on M_n. Then T can be tensored with any other such map and the composed system has still the right positivity property.

4.8.3 Representations of Completely Positive Operators

The fundamental theorem behind almost all results on complete positivity is Stinespring's famous representation theorem for completely positive maps. Consider a map $T : \mathcal{A} \to \mathcal{B}$. Since $\mathcal{B}$ is an operator algebra it is contained in $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$ and it is no restriction to assume that T is a map $T : \mathcal{A} \to \mathcal{B}(\mathcal{H})$.

Theorem 10. *(Stinespring 1955, cf. [31]). For a map $T : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ the following conditions are equivalent:*

a) T is completely positive.
*b) There is a representation (i.e., a *-homomorphism) $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ ($\mathcal{K}$ some further Hilbert space) and a bounded linear map $v : \mathcal{H} \to \mathcal{K}$ such that*

$$T(x) = v^* \pi(x) v$$

for all $x \in \mathcal{A}$. If $T(\mathbb{1}) = \mathbb{1}$ then v is an isometry.

The triple $(\mathcal{K}, \pi, v)$ is called *Stinespring representation* for T. If it is *minimal* that is, the linear span of $\{\pi(x)v\xi \ , \ \xi \in \mathcal{H} \ , \ x \in \mathcal{A}\}$ is dense in $\mathcal{K}$, then the Stinespring representation is *unique up to unitary equivalence.*

From Stinespring's theorem it is easy to derive the following *concrete representation* for completely positive operators on M_n.

Theorem 11. *For $T : M_n \to M_n$ the following conditions are equivalent.*

a) T is completely positive.
b) There are elements $a_1, \ldots, a_k \in M_n$ for some k, such that

$$T(x) = \sum_{i=1}^{k} a_i^* x a_i \ .$$

Clearly, T is identity preserving if and only if $\sum_i a_i^* a_i = \mathbb{1}$.

Such decompositions of completely positive operators are omnipresent whenever completely positive operators occur in a physical context. These decompositions are by no means uniquely determined (see below). In a physical context they rather correspond to different physical situations (cf. Sect. 4.10.2).

The following basic facts can be derived without much difficulty from Stinespring's theorem: A concrete representation $T(x) = \sum_{i=1}^{k} a_i^* x a_i$ for T can always be chosen such that $\{a_1, a_2, \ldots, a_k\} \subseteq M_n$ is linearly independent, in particular, $k \leq n^2$. We call such a representation *minimal*. The cardinality k of a minimal representation of T is uniquely determined by T, that means two minimal representations of T have the same cardinality. Finally, all minimal representations can be obtained from the following result.

Proposition 4. *Let $T(x) = \sum_{i=1}^{k} a_i^* x a_i$ and $S(x) = \sum_{j=1}^{l} b_j^* x b_j$ be two minimal representations of completely positive operators S and T on M_n. The following conditions are equivalent:*

a) $S = T$.
b) $k = l$ *and there is a unitary* $k \times k$*–matrix* $\Lambda = (\lambda_{ij})_{i,j}$ *such that*

$$a_i = \sum_{j=1}^{k} \lambda_{ij}\, b_j \ .$$

The results on concrete representations have an obvious generalization to the case $n = \infty$. Then infinite sums may occur but they must converge (in the strong operator topology on $\mathcal{B}(\mathcal{H})$).

4.9 Semigroups of Completely Positive Operators and Lindblad Generators

4.9.1 Generators of Lindblad Form

As we saw, to a Markov process in continuous time there is always associated a semigroup $(T_t)_{t\geq 0}$ of transition operators on $\mathcal{A}_0$, which will be strongly continuous in all cases of physical interest. According to the general theory (cf. Sect. 4.2.4) it has a generator L such that

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t(x) = L(T_t(x))$$

for all x in the domain of L, formally $T_t = \mathrm{e}^{Lt}$. In the case of a classical Markov process one could say much more: L has the form of a partial differential operator of second order of a very specific form (cf. Sect. 4.2.4). It is natural to ask whether a similar characterization of generators can be given in the general noncommutative case. This turns out to be a difficult problem and much research on this problem remains to be done. A first breakthrough has been obtained in a famous paper by G. Lindblad [23] and at the same time in a special case by V. Gorini, A. Kossakowski, and E.C.G. Sudarshan [10] in 1976.

Theorem 12. *Let $(T_t)_{t\geq 0}$ be a semigroup of completely positive identity preserving operators on M_n with generator L. Then there is a completely positive operator $M : M_n \to M_n$ and a self-adjoint element $h \in M_n$ such that*

$$L(x) = \mathrm{i}[h, x] + M(x) - \frac{1}{2}(M(\mathbb{1})x + xM(\mathbb{1}))$$

where $[h, x]$ stands, as usual, for the commutator $hx - xh$. Conversely, every operator L of this form generates a semigroup of completely positive identity preserving operators.

Since we know that every such M has a concrete representation as

$$M(x) = \sum_i a_i^* x a_i$$

we obtain for L the representation

$$L(x) = \mathrm{i}[h, x] + \sum_i a_i^* x a_i - \frac{1}{2}(a_i^* a_i x + x a_i^* a_i)$$

This representation is usually called *Lindblad form* of the generator.

Lindblad was able to prove this result for norm-continuous semigroups on $\mathcal{B}(\mathcal{H})$ for infinite dimensional $\mathcal{H}$. In this situation L is still a bounded operator. If one wants to treat the general case of only strongly continuous semigroups on $\mathcal{B}(\mathcal{H})$ then one has to take into account, for example, infinite unbounded sums of bounded and unbounded operators a_i. Until today no general characterization of such generators is available which would generalize the results in Sect. 4.2.4. Nevertheless, Lindblad's characterization seems to be 'philosophically true' as in most cases of physical interest generators appear to be in Lindblad form. In typical situations the operators a_i are creation and annihilation operators.

4.9.2 Interpretation of Generators of Lindblad Form

It still may be not obvious what a generator in Lindblad form has to do with the particular form of the partial differential operators of Sect. 4.2.4. This may be clarified by the following observation.

For $h \in \mathcal{A}_0$ consider the operator D on $\mathcal{A}_0$ given by

$$D : x \mapsto \mathrm{i}[h, x] = \mathrm{i}(hx - xh) \qquad (x \in \mathcal{A}_0)\ .$$

Then

$$\begin{aligned} D(xy) &= \mathrm{i}(hxy - xyh) = \mathrm{i}(hxy - xhy + xhy - xyh) \\ &= \mathrm{i}[h, x]y + x\mathrm{i}[h, y] \\ &= D(x) \cdot y + x \cdot D(y)\ , \end{aligned}$$

hence D is a derivation.

In Lindblad's theorem h is self-adjoint and in this case D is a real derivation (i.e., $D(x^*) = D(x)^*$) and generates the time evolution $x \mapsto \mathrm{e}^{+\mathrm{i}ht} x \mathrm{e}^{-\mathrm{i}ht}$ implemented by the unitary group $(\mathrm{e}^{\mathrm{i}ht})_{t\in\mathbb{R}}$. This is all familiar from the solution

of Schrödingers equation in the Heisenberg picture. Therefore, for self-adjoint h the term $x \mapsto \mathrm{i}[h,x]$ is a 'quantum derivative' of first oder and corresponds to the drift terms in Sect. 4.2.4. For the second derivative we obtain after a short calculation

$$\begin{aligned} D^2(x) &= \mathrm{i}[h,\mathrm{i}[h,x]] \\ &= 2(hxh - \tfrac{1}{2}(h^2x + xh^2)) \ . \end{aligned}$$

This resembles the second part of a generator in Lindblad form. It shows that for self-adjoint a the term

$$axa - \frac{1}{2}(a^2x + xa^2)$$

is a second derivative and thus generates a quantum diffusion.

On the other hand for $a = u$ unitary the term $a^*xa - \frac{1}{2}(a^*ax + xa^*x)$ turns into $u^*xu - x$ which generates a jump process: If we define the *jump operator* $J(x) := u^*xu$ and

$$\begin{aligned} L(x) &:= J(x) - x = (J - \mathrm{Id})(x) \ , \text{ then} \\ \mathrm{e}^{Lt} &= \mathrm{e}^{(J-\mathrm{Id})t} = \mathrm{e}^{-t} \cdot \mathrm{e}^{Jt} \\ &= \textstyle\sum_{n=0}^{\infty} \mathrm{e}^{-t} \frac{t^n}{n!} J^n \ , \end{aligned}$$

which is a Poissonian convex combination of the jumps $\{J^n \ , \ n \in \mathbb{N}\}$. Therefore, terms of this type correspond to classical jump processes.

In a general generator of Lindblad type the sum $\sum_i a_i^* x a_i - \frac{1}{2}(a_i^* a_i x + a_i^* a_i x)$ cannot always be decomposed into summands with a_i self-adjoint and a_i unitary and there are more general types of transitions. In the context of quantum trajectories such decompositions play an important role (cf., e.g., [2]). They are closely related to *unravellings* of the time evolution T_t. The generators which allow decompositions with a_i self-adjoint or unitary have been characterized and investigated in [16].

4.9.3 Quantum Stochastic Differential Equations

We already mentioned that for a semigroup $(T_t)_{t\geq 0}$ of transition operators on $\mathcal{A}_0$ there is no canonical procedure which leads to an analogue of the canonical representation of a classical Markov process on the space of its paths. For $\mathcal{A}_0 = M_n$, however, quantum stochastic calculus allows to construct a stochastic process which is almost a Markov process in the sense of our definition. Stationarity, however, is not preserved by this construction.

Consider $T_t = \mathrm{e}^{Lt}$ on M_n and assume, for simplicity only, that the generator L has the simple Lindblad form

$$L(x) = \mathrm{i}[h,x] + b^*xb - \frac{1}{2}(b^*bx + xb^*b) \ .$$

Let $\mathcal{F}(L^2(\mathbb{R}))$ denote the symmetric Fock space of $L^2(\mathbb{R})$. For a test function $f \in L^2(\mathbb{R})$ there exist the creation operator $A^*(f)$ and annihilation operator $A(f)$ as unbounded operators on $\mathcal{F}(L^2(\mathbb{R}))$. For $f = \chi_{[0,t]}$, the characteristic function of the interval $[0,t] \subseteq \mathbb{R}$, the operators $A^*(f)$ and $A(f)$ are usually denoted by A_t^* (or $A_t^\dagger$) and A_t, respectively. It is known that the operators $B_t := A_t^* + A_t$ on $\mathcal{F}(L^2(\mathbb{R})), t \geq 0$, give a representation of classical Brownian motion by a commuting family of self-adjoint operators on $\mathcal{F}(L^2(\mathbb{R}))$ (cf. the discussion in Sect. 4.4.3). Starting from this observation R. Hudson and K.R. Parthasarathy have extended the classical Itô-calculus of stochastic integration with respect to Brownian motion to more general situations on symmetric Fock space. An account of this theory is given in [25].

In particular, one can give a rigorous meaning to the stochastic differential equation

$$du_t = u_t \left(b dA_t^* + b^* dA_t + (\mathrm{i}h - \frac{1}{2} b^* b) dt) \right)$$

where bdA_t^* stands for $b \otimes dA_t^*$ on $\mathbb{C}^n \otimes \mathcal{F}(L^2(\mathbb{R}))$ and similarly for $b^* dA_t$, while $ih - \frac{1}{2} b^* b$ stands for $(ih - \frac{1}{2} b^* b) \otimes \mathbb{1}$ on $\mathbb{C}^n \otimes \mathcal{F}(L^2(\mathbb{R}))$. It can be shown that the solution exists, is unique, and is given by a family $(u_t)_{t \geq 0}$ of unitaries on $\mathbb{C}^n \otimes \mathcal{F}(L^2(\mathbb{R}))$ with $u_0 = \mathbb{1}$. This leads to a stochastic process with random variables

$$i_t : M_n \ni x \mapsto u_t^* \cdot x \otimes \mathbb{1} \cdot u_t \in M_n \otimes \mathcal{B}(\mathcal{F}(L^2(\mathbb{R})))$$

which can, indeed, be viewed as a Markov process with transition operators $(T_t)_{t \geq 0}$. This construction can be applied to all semigroups of completely positive identity preserving operators on M_n, and to many such semigroups on $\mathcal{B}(\mathcal{H})$ for infinite dimensional $\mathcal{H}$.

4.10 Repeated Measurement and Its Ergodic Theory

In this section we describe how completely positive operators and their concrete representation occur in the description of a measurement of a quantum system. We then turn to some recent results on the ergodic theory of repeated measurement.

4.10.1 Measurement According to von Neumann

Consider a system described by its algebra $\mathcal{A}$ of observables which is in a state φ. In the typical quantum case $\mathcal{A}$ will be $\mathcal{B}(\mathcal{H})$ and φ will be given by a density matrix ρ on $\mathcal{H}$. In continuation of our discussion in Sect. 4.4.1 we consider the measurement of an observable given by a self-adjoint operator X on $\mathcal{H}$. For simplicity we assume that the spectrum $\sigma(X)$ is finite so that X has a spectral decomposition of the form $X = \sum_i \lambda_i p_i$ with $\sigma(X) = \{\lambda_1, \ldots \lambda_n\}$ and orthogonal projections $p_1, p_2, \ldots, p_n$ with $\sum_i p_i = \mathbb{1}$. According to the laws of quantum mechanics the spectrum $\sigma(X)$ is the set of possible outcomes of this measurement (cf. Sect. 4.4.1). The probability for obtaining the value $\lambda_i \in \sigma(X)$

is given by $\varphi(p_i)$ $(= \operatorname{tr}(\rho p_i))$ and after such a measurement the state of the system has changed to the state $\varphi_i : x \mapsto (\varphi(p_i x p_i))/\varphi(p_i)$ (with density matrix $p_i \rho p_i / \operatorname{tr}(p_i \rho)$), provided that $\varphi(p_i) \neq 0$. In particular, the spectral measure $\sigma(X) \ni \lambda_i \mapsto \varphi(p_i)$ defines a probability measure μ_0 on $\Omega_0 := \sigma(X)$. If we perform the measurement of X but we ignore its outcome (this is sometimes called "measurement with deliberate ignorance") then the initial state has changed to the state φ_i with probability $\varphi(p_i)$. Therefore, the state of the system after such a measurement is adequately described by the state $\varphi_X := \sum_i \varphi(p_i) \cdot \varphi_i$, hence

$$\varphi_X(x) = \sum_i \varphi(p_i x p_i) \ .$$

(It is not even necessary to single out the cases with probability $\varphi(p_i) = 0$.)

Turning to the description in the Heisenberg picture an operator x changes to $p_i x p_i / \varphi(p_i)$ if λ_i was measured. A measurement with deliberate ignorance is described by $x \mapsto \sum_i p_i x p_i$ which is a conditional expectation of $\mathcal{A}$ onto the subalgebra $\{\sum_i p_i x p_i \ , \ x \in \mathcal{A}\}$.

4.10.2 Indirect Measurement According to K. Kraus

In many cases the observables of a system are not directly accessible to an observation or an observation would lead to an undesired destruction of the system as is typically the case when measuring photons.

In such a situation one obtains information on the state φ of the system by coupling the system to another system – a measurement apparatus – and reading off the value of an observable of the measurement apparatus. A mathematical description of such measurements was first given by K. Kraus [12].

As a typical example of such a measurement one might consider the micromaser experiment [33] (from a different point of view this system is discussed in the next Section). The system to be measured is one mode of the electromagnetic field inside a cavity. Its algebra of observables is $\mathcal{A} = \mathcal{B}(\mathcal{H})$, the same as for one quantum harmonic oscillator. Two–level atoms are sent through this cavity and during their passage they can interact with the field mode according to the Jaynes–Cummings interaction. Their algebra of observables is $\mathcal{C} = M_2$, the algebra of complex 2×2–matrices. If such a two–level atom is prepared in an initial state ψ, sent through the cavity and measured afterwards this gives a typical example of such measurement.

A measurement procedure of this type can be decomposed into the following steps:

α) Couple the system $\mathcal{A}$ in its initial state φ to another system – the measurement apparatus – with observable algebra $\mathcal{C}$, which is initially in a state ψ.

β) For a certain time t_0 the composed system evolves according to a dynamics $(\alpha_t)_{t \in \mathbb{R}}$. In the Heisenberg picture, $(\alpha_t)_{t \in \mathbb{R}}$ is a group of automorphisms of $\mathcal{A} \otimes \mathcal{C}$. After the interaction time t_0 the overall change of the system is given by $T_{\text{int}} := \alpha_{t_0}$ (the index 'int' stands for 'interaction').

γ) Now an observable $X = \sum_i \lambda_i p_i \in \mathcal{C}$ is measured and changes the state of the composed system accordingly.

δ) The new state of $\mathcal{A}$ is finally obtained by restricting the new state of the composed system to the operators in $\mathcal{A}$.

Mathematically each step corresponds to a map on states and the whole measurement is obtained by composing those four maps (on infinite dimensional algebras all states are assumed to be normal):

α) The measurement apparatus is assumed to be initially in a fixed state ψ. Therefore, in the Schrödinger picture, coupling $\mathcal{A}$ to $\mathcal{C}$ corresponds to the map $\varphi \mapsto \varphi \otimes \psi$ of states on $\mathcal{A}$ into states on $\mathcal{A} \otimes \mathcal{C}$. We already saw in Sect. 4.5.5 that dual to this map is the conditional expectation of tensor type

$$P_\psi : \mathcal{A} \otimes \mathcal{C} \mapsto \mathcal{A} : x \otimes y \mapsto \psi(y) \cdot x$$

which thus describes this step in the Heisenberg picture (when identifying $\mathcal{A}$ with the subalgebra $\mathcal{A} \otimes \mathbb{1}$ of $\mathcal{A} \otimes \mathcal{C}$ we may still call P_ψ a conditional expectation).

β) The time evolution of $\mathcal{A} \otimes \mathcal{C}$ during the interaction time t_0 is given by an automophism T_{int} on $\mathcal{A} \otimes \mathcal{C}$. It changes any state χ on $\mathcal{A} \otimes \mathcal{C}$ into $\chi \circ T_{\text{int}}$.

γ) A measurement of $X = \sum_i \lambda_i p_i \in \mathcal{C}$ changes a state χ on $\mathcal{A} \otimes \mathcal{C}$ into the state $\chi(\mathbb{1} \otimes p_i \cdot \mathbb{1} \otimes p_i)/\chi(\mathbb{1} \otimes p_i)$ and this happens with probability $\chi(\mathbb{1} \otimes p_i)$. It is convenient to consider this state change together with its probability. Then it can be described by the non-normalized but linear map

$$\chi \mapsto \chi(\mathbb{1} \otimes p_i \ \cdot \ \mathbb{1} \otimes p_i) \ .$$

Dual to this is the map

$$\mathcal{A} \otimes \mathcal{C} \ni z \mapsto \mathbb{1} \otimes p_i \cdot z \cdot \mathbb{1} \otimes p_i$$

which thus describes the unnormalized state change due to a measurement with outcome λ_i in the Heisenberg picture. When turning from a measurement with outcome λ_i to a measurement with deliberate ignorance, the difference between the normalized and the unnormalized description disappears.

δ) This final step maps a state χ on the composed system $\mathcal{A} \otimes \mathcal{C}$ to the state

$$\chi|_{\mathcal{A}} : \mathcal{A} \ni x \mapsto \chi(x \otimes \mathbb{1}) \ .$$

The density matrix of $\chi|_{\mathcal{A}}$ is obtained from the density matrix of χ by a partial trace over $\mathcal{C}$. As we already saw in Sect. 4.5.5 a description of this step in the Heisenberg picture is given by the map

$$\mathcal{A} \ni x \mapsto x \otimes \mathbb{1} \in \mathcal{A} \otimes \mathcal{C} \ .$$

By composing all four maps in the Schrödinger picture and in the Heisenberg picture we obtain

$$\begin{array}{ccccccccc} \mathcal{A} & & \alpha) & \mathcal{A}\otimes\mathcal{C} & \beta) & \mathcal{A}\otimes\mathcal{C} & \gamma) & \mathcal{A}\otimes\mathcal{C} & \delta) \quad \mathcal{A} \\ \varphi & \longrightarrow & & \varphi\otimes\psi & \longrightarrow & \varphi\otimes\psi\circ T_{\mathrm{int}} & \longrightarrow & \varphi_i & \longrightarrow \varphi_i|_{\mathcal{A}} \\ P_\psi T_{\mathrm{int}}(x\otimes p_i) & \longleftarrow & & T_{\mathrm{int}}(x\otimes p_i) & \longleftarrow & x\otimes p_i & \longleftarrow & x\otimes \mathbb{1} & \longleftarrow \quad x \end{array}$$

with $\varphi_i := \varphi\otimes\psi\circ T_{\mathrm{int}}(\mathbb{1}\otimes p_i \ \cdot \ \mathbb{1}\otimes p_i)$. Altogether, the operator

$$T_i : \mathcal{A}\to\mathcal{A} : x \mapsto P_\psi T_{\mathrm{int}}(x\otimes p_i)$$

describes in the Heisenberg picture the non-normalized change of states in such a measurement with the i-th value λ_i as outcome. The probability for this to happen can be computed from the previous section as

$$\begin{aligned} \varphi\otimes\psi\circ T_{\mathrm{int}}(\mathbb{1}\otimes p_i) &= \varphi\otimes\psi(T_{\mathrm{int}}(\mathbb{1}\otimes p_i)) \\ &= \varphi(P_\psi T_{\mathrm{int}}(\mathbb{1}\otimes p_i)) \\ &= \varphi(T_i(\mathbb{1})). \end{aligned}$$

When performing such a measurement but deliberately ignoring its outcome this measurement is described (in the Heisenberg picture) by

$$T = \sum_i T_i \ .$$

Since the operators T_i were unnormalized they need not to be weighted with their probabilities.

More generally, for any subset $A\subseteq\Omega_0=\sigma(X)$ we obtain an operator

$$T_A := \sum_{i\in A} T_i \ .$$

It describes the state change (in the Heisenberg picture), if a measurement of X was performed and if it is known that the outcome of this measurement is in A ('partial deliberate ignorance'). The assignment $A\mapsto T_A$ is an example of a *positive map valued (PMV) measure* (cf. [3]).

The operator T can be computed more explicitly: for $x\in\mathcal{A}$ we obtain

$$T(x) = \sum_i P_\psi T_{\mathrm{int}}(x\otimes p_i) = P_\psi T_{\mathrm{int}}(x\otimes\mathbb{1})$$

since $\sum_i p_i = \mathbb{1}$. From their construction it is clear that all operators T and T_i are completely positive and T is, in addition, identity preserving, that is $T(\mathbb{1})=\mathbb{1}$. It should be noted that T does no longer depend on the particular observable $X\in\mathcal{C}$ but only on the interaction T_{int} and the initial state ψ of the apparatus $\mathcal{C}$. Only the particular decomposition of T reflects the particular choice of X.

4.10.3 Measurement of a Quantum System and Concrete Representations of Completely Positive Operators

Consider now a 'true quantum situation' where $\mathcal{A}$ is given by the algebra M_n of all $n\times n$–matrices and $\mathcal{C}$ is given by M_m for some m. Assume further that we

perform a kind of 'perfect measurement': In order to draw a maximal amount of information from such a measurement the spectral projection p_i should be minimal hence 1-dimensional, and the initial state ψ of the measurement apparatus should be a pure state. It then follows that there are operators $a_i \in \mathcal{A} = M_n$, $1 \leq i \leq m$, such that

$$\begin{aligned} T_i(x) &= a_i^* x a_i \ , \\ \text{and} \quad T(x) &= \textstyle\sum_i a_i^* x a_i \ . \end{aligned}$$

Indeed, every automophism T_{int} of $M_n \otimes M_m$ is implemented by a unitary $u \in M_n \otimes M_m$, such that $T_{\text{int}}(z) = Ad\, u(z) = u^* z u$ for $z \in M_n \otimes M_m$. Since $M_n \otimes M_m$ can be identified with $M_m(M_n)$, the algebra of $m \times m$–matrices with entries from M_n, the unitary u can be written as a matrix

$$u = \begin{pmatrix} & \\ & u_{ij} \\ & \end{pmatrix}_{m \times m}$$

with entries $u_{ij} \in M_n$, $1 \leq i, j \leq m$. Moreover, the pure state ψ on M_m is a vector state induced by a unit vector

$$\begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix} \in \mathbb{C}^m$$

while p_i projects onto the 1–dimensional subspace spanned by a unit vector

$$\begin{pmatrix} \xi_1^i \\ \vdots \\ \xi_m^i \end{pmatrix} \in \mathbb{C}^m \ .$$

A short computation shows that $T(x) = \sum_i T_i(x)$, where

$$\begin{aligned} T_i(x) &= P_\psi T_{\text{int}}(x \otimes p_i) = P_\psi(u^* \cdot x \otimes p_i \cdot u) \\ &= a_i^* x a_i \ , \end{aligned}$$

with

$$a_i = (\overline{\xi}_1, \ldots, \overline{\xi}_m) \cdot \begin{pmatrix} & \\ & u_{ij} \\ & \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix} \ .$$

Therefore, a completely positive operator T with $T(\mathbb{1}) = \mathbb{1}$ describes, in the Heisenberg picture, the state change of a system due to a measurement with deliberate ignorance. It depends only on the coupling to a measurement apparatus and the initial state of the apparatus. Only the measurement of a specific observable $X = \sum_i \lambda_i p_i$ leads to a decomposition $T = \sum_i T_i$, where T_i describes the (non-normalized) change of states on the occurrence of the outcome λ_i. The probability for this is given by $\psi(T_i(\mathbb{1}))$. In the special case of a perfect quantum

measurement the operators T_i are of the form $T_i(x) = a_i^* x a_i$ and the probability for an outcome λ_i is given by $\varphi(a_i^* a_i)$.

Conversely, a concrete representation $T(x) = \sum_i a_i^* x a_i$ for $T : M_n \to M_n$ with $T(\mathbb{1}) = \mathbb{1}$ may always be interpreted as coming from such a measurement: Since $T(\mathbb{1}) = \mathbb{1}$, the map

$$v := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \text{from} \quad \mathbb{C}^n \quad \text{to} \quad \mathbb{C}^n \otimes \mathbb{C}^m = \mathbb{C}^n \oplus \ldots \oplus \mathbb{C}^n$$

is an isometry, and $T(x) = v^* \cdot x \otimes \mathbb{1} \cdot v$ is a Stinespring representation for T.

Construct any unitary $u \in M_n \otimes M_m = M_m(M_n)$ which has

$$v = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

in its first column and put

$$\tilde{\psi} := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{C}^m \ .$$

Then $P_\psi(u^* \cdot x \otimes \mathbb{1} \cdot u) = v^* \cdot x \otimes \mathbb{1} \cdot v = T(x)$. With p_i the orthogonal projection onto the 1-dimensional subspace spanned by the i-th canonical basis vector

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with the 1 at the i-th place we obtain, indeed,

$$P_\psi(u^* \cdot x \otimes p_i \cdot u) = a_i^* x a_i \ .$$

4.10.4 Repeated Measurement

Consider now the case where we repeat such a measurement indefinitly. At each time step we couple the system in its present state to the same measurement apparatus which is always prepared in the same initial state. We perform a measurement thereby changing the state of the system, decouple system and apparatus, and start again the whole procedure. In order to have a concrete example in mind consider again the micromaser experiment. Sending identically prepared

atoms through the cavity, one after another, and measuring their states after they have passed the cavity, this is a typical example for such a measurement (and this is what experimentalists do with the micromaser, see also Sect. 2.5).

Continuing the previous discussion, each single measurement can have an outcome i in a (finite) set Ω_0. For simplicity assume that we perform a perfect quantum measurement. Then there is a completely positive, identity preserving operator T on an algebra M_n ($n \in \mathbb{N}$ or $n = \infty$), with a decomposition $T(x) = \sum_{i\in\Omega_0} a_i^* x a_i$. A trajectory of the repeated measurement will be an element in

$$\Omega = \Omega_0^{\mathbb{N}} = \{(\omega_1, \omega_2, \ldots) : \omega_i \in \Omega_0\} \ .$$

Given the system is initially in a state φ then the probability for measuring $i_1 \in \Omega_0$ at the first measurement is $\varphi(a_{i_1}^* a_{i_1})$, and in this case its state changes to

$$\frac{\varphi(a_{i_1}^* \cdot \ a_{i_1})}{\varphi(a_{i_1}^* a_{i_1})} \ .$$

Therefore the probability of measuring now $i_2 \in \Omega_0$ in a second measurement is given by $\varphi(a_{i_1}^* a_{i_2}^* a_{i_2} a_{i_1})$ and in this case the state changes further to

$$\frac{\varphi(a_{i_1}^* a_{i_2}^* \cdot \ a_{i_2} a_{i_1})}{\varphi(a_{i_1}^* a_{i_2}^* a_{i_2} a_{i_1})} \ .$$

Similarly, the probability for obtaining a sequence of outcomes $(i_1, \ldots, i_n) \in \Omega_0^n = \Omega_0 \times \ldots \times \Omega_0$ is given by

$$\mathbb{P}_\varphi^n((i_1, i_2, \ldots, i_n)) := \varphi(a_{i_1}^* a_{i_2}^* \cdot \ldots \cdot a_{i_n}^* a_{i_n} \cdot \ldots \cdot a_{i_2} a_{i_1})$$

which defines a probability measure $\mathbb{P}_\varphi^n$ on Ω_0^n. The identity $\sum_{i\in\Omega_0} a_i^* a_i = T(\mathbb{1}) = \mathbb{1}$ immediately implies the compatibility condition

$$\mathbb{P}_\varphi^{n+1}((i_1, i_2, \ldots, i_n) \times \Omega_0) = \mathbb{P}_\varphi^n((i_1, \ldots, i_n)) \ .$$

Therefore there is a unique probability measure $\mathbb{P}_\varphi$ on Ω defined on the σ-algebra Σ generated by cylinder sets

$$\Lambda_{i_1,\ldots,i_n} := \{\omega \in \Omega : \omega_1 = i_1, \ldots, \omega_n = i_n\} \ ,$$

such that

$$\mathbb{P}_\varphi(\Lambda_{i_1,\ldots,i_n}) = \mathbb{P}_\varphi^n((i_1, \ldots, i_n)) \ .$$

The interpretation of $\mathbb{P}_\varphi$ is obvious: For every $A \in \Sigma$, $\mathbb{P}_\varphi(A)$ is the probability of measuring a trajectory in A.

Denote, again, by σ the time shift on Ω, that is

$$\sigma((\omega_1, \omega_2, \omega_3, \ldots)) = (\omega_2, \omega_3, \omega_4, \ldots).$$

Then a short computation shows that

$$\mathbb{P}_\varphi(\sigma^{-1}(A)) = \mathbb{P}_{\varphi\circ T}(A) \ ,$$

for all sets $A \in \Sigma$. In particular, if φ is stationary for T, that is $\varphi \circ T = \varphi$, then $\mathbb{P}_\varphi$ is stationary for σ on Ω. This allows to use methods from classical ergodic theory for the analysis of trajectories for repeated quantum measurement.

Theorem 13. *Ergodic Theorem ([19]) If*

$$\lim_{n\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \varphi \circ T^n = \varphi_0$$

for all states φ, then for any initial state φ and for any set $A \in \Sigma$ which is time invariant, that is $\sigma^{-1}(A) = A$, we have either $\mathbb{P}_\varphi(A) = 0$ or $\mathbb{P}_\varphi(A) = 1$.

We illustrate this theorem by an application: How likely is it to find during a repeated measurement a certain sequence of outcomes $(i_1, \ldots, i_n) \in \Omega_0^n$? If the initial state is a T–invariant state φ_0 then the probability of finding this sequence as outcomes of the measurements $k, k+1, \ldots k+n-1$ is the same as the probability of finding it in the first n measurements. In both cases this probability is given by $\varphi_0(a_{i_1}^* \ldots a_{i_n}^* a_{i_n} \ldots a_{i_1})$. However, it is also true that this probability is identical to the relative frequency of occurences of this sequence in an arbitrary individual trajectory:

Corollary 2. *([19]) For any initial state φ and for $(i_1, \ldots, i_n) \in \Omega_0^n$,*

$$\lim_{N\to\infty} \frac{1}{N} |\{j : j < N \quad \textit{and} \quad \omega_{j+1} = i_1, \ldots, \omega_{j+n} = i_n\}|$$

$$= \varphi_0(a_{i_1}^* \cdot \ldots \cdot a_{i_n}^* a_{i_n} \cdot \ldots \cdot a_{i_1}) \ ,$$

for $\mathbb{P}_\varphi$ – almost all paths $\omega \in \Omega_0^{\mathbb{N}}$.

This is a version of of Birkhoff's ergodic theorem (cf. Sect. 4.3.4) for repeated measurement. Similarly, all statistical information can be drawn from the observation of a single trajectory of the repeated measurement process, that means that correlations can be measured as autocorrelations. This is tacitly assumed in the physics literature but it has not been proven up to now. There is also an ergodic theory for continuous time Davies processes for which we refer to [19].

Recently we have been able to prove an even stronger result:

Theorem 14. *([20]) Under the assumptions of the above ergodic theorem*

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} \frac{\varphi(a_{i_1}^* \ldots a_{i_n}^* \ \cdot \ a_{i_n} \ldots a_{i_1})}{\varphi(a_{i_1}^* \ldots a_{i_n}^* a_{i_n} \ldots a_{i_1})} = \varphi_0$$

for any initial state φ and $\omega = (i_1, i_2, \ldots)$, $\mathbb{P}_\varphi$-almost surely.

We close this discussion with some remarks. If a sequence of n measurements has led to a sequence of outcomes $(i_1, \ldots, i_n) \in \Omega_0^n$, then the operator

$$T_{i_1 i_2 \ldots i_n} : x \mapsto a_{i_1}^* \ldots a_{i_n}^* x a_{i_n} \ldots a_{i_1}$$

describes the change of the system under this measurement including the probability for this to happen. Similarly, for any subset $A \subseteq \Omega_0^n$ we may associate the operator $T_A^n := \sum_{\omega \in \Omega_0^n} T_\omega$. In particular $T_{\Omega_0^n} = T^n$. Again, the assignment $A \mapsto T_A^n$ is a positive map valued measure (cf. Sect. 4.10.2).

For subsets $A \subseteq \Omega_0^n$ and $B \subseteq \Omega_0^m$ the set $A \times B$ may be naturally identified with a subset of $\Omega_0^n \times \Omega_0^m = \Omega_0^{n+m}$, and from the definition of T_A^n we obtain

$$T_{A\times B}^{n+m} = T_A^n \circ T_B^m \ .$$

Therefore, the operators $\{T_A^n : n \in \mathbb{N}, A \subseteq \Omega_0^n\}$ form a discrete time version of the type of quantum stochastic processes which have been considered in [3] for the description of quantum counting processes. The above identity may be considered as a Markov property for positive map valued measures.

We finally remark that decompositions $T(x) = \sum_i a_i^* x a_i$ as above define so-called quantum trajectories. They are extensively used in the numerical simulation of irreversible behaviour of open quantum systems, and in particular for computing their equilibrium states (cf. [2]). The above theorems show that for many purposes it is enough to do such a simulation once along one path while it is not necessary to perform a whole sample of such simulations in order to pass to a sample average.

4.11 The Micromaser as a Quantum Markov Process

The micromaser experiment as it is carried through in the group around H. Walther [33] in Garching turns out to be an experimental realization of a quantum Markov process with all the structure as described in Sect. 4.6.1. This fact can be used to employ some general theory on such quantum Markov processes, and leads to some suggestions on how to use a micromaser for the preparation of interesting quantum states. In the following we give a description of this recent considerations. For details we refer to [34,35] for the results on the micromaser, and to [18] for some mathematical background.

4.11.1 The Experiment

In the micromaser experiment (see also Sect. 2.4.1 of this volume!) a beam of isolated rubidium atoms is prepared in highly exited Rydberg states. Only two of these states are relevant for the following and we may consider the atoms as quantum mechanical two-level systems. Thus the algebra of observables for a single atom is the algebra M_2 of 2×2-matrices. The atoms with a fixed velocity are singled out and sent through a micro-wave cavity which has small holes on both sides for the atoms to pass through this cavity. During their passage through the cavity the atoms interact with one mode of the electromagnetic field in this cavity which is in tune with the energy difference of the two levels of these atoms. One mode of the electromagnetic field is described as a quantum harmonic oscillator, hence its algebra of observables is given by $B(\mathcal{H})$, where $\mathcal{H} = L^2(\mathbb{R})$ or $\mathcal{H} = l^2(\mathbb{N})$, depending on whether we work in the position representation or in the energy representation. The atomic beam is weak enough, so there is at most one atom inside the cavity at a time, and since the atoms come all with the same velocity there is a fixed interaction time for the interaction between atom

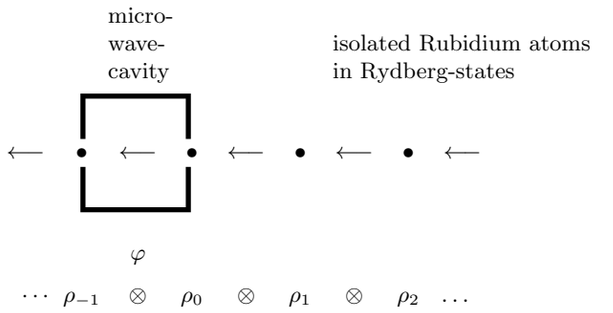


Fig. 4.1. Scheme of description for the micromaser experiment. The states of the field mode and the successive atoms are denoted as φ and ρ_i, respectively

and field for each of these atoms. To simplify the discussion further, we assume that the time between the passage through the cavity of two successive atoms is always the same, what defines our time unit. This is not realistic, but due to the particular form of the model (cf. below) the free evolution commutes with the interaction evolution and can be handled separately. Therefore it is easy to turn from this description to a more realistic description afterwards, where the arrival times of atoms in the cavity have, for example, a Poissonian distribution.

If we don't specify the algebras and interaction involved we obtain a scheme of description for the experiment which is depicted in Fig. 4.1. There, φ stands for the state (density matrix) of the field mode, and the ρ_i denote the states of the successive atoms.

4.11.2 The Micromaser Realizes a Quantum Markov Process

We consider the time evolution in the interaction picture: It decomposes naturally into two parts: During one time step one atom has passed through the cavity. If before the passage the cavity was in a state φ and the atom in a state ρ, then the state of field mode and atom is now given by $u_{\mathrm{int}} \cdot \varphi \otimes \rho \cdot u_{\mathrm{int}}^*$, where $u_{\mathrm{int}} = \mathrm{e}^{\mathrm{i}\mathbb{H}t_0}$, $\mathbb{H}$ is the Hamiltonian, and t_0 is the interaction time given by the time an atom needs to pass through the cavity. The other part of the time evolution describes the moving atoms. For one time unit, it is the tensor right shift on the tensor product of density matrices of the flying atoms. Thus the time evolution of the whole system for one time unit might be written in the following suggestive way:

$$\begin{array}{c} \varphi \\ u_{\mathrm{int}} \otimes u_{\mathrm{int}}^* \\ \text{tensor left shift } (\ \cdots\ \rho_{-1}\ \otimes\ \rho_0\ \otimes\ \rho_1 \otimes \rho_2 \cdots) \end{array}$$

We continue to use this suggestive picture for our description. Then a description of this system in the Heisenberg picture looks as follows: If $x \in B(\mathcal{H})$ is an

observable of the field mode and $(y_i)_i \in M_2$ are observables of the atoms, then a typical observable of the whole system is given by

$$\begin{array}{ccc} x & & B(\mathcal{H}) \\ \otimes & \in & \otimes \\ \cdots\, y_{-1} \otimes y_0 \otimes y_1 \cdots & & \cdots\, M_2 \otimes \ M_2 \ \otimes M_2 \cdots \end{array}$$

and arbitrary observables are (limits of) linear combinations of such observables. The dynamics of the interaction between field mode and one passing atom is given by

$$\alpha_{\text{int}} : \begin{array}{c} x \\ \otimes \\ y_0 \end{array} \mapsto \begin{array}{c} x \\ u_{\text{int}}^* \cdot \otimes \cdot u_{\text{int}} \\ y_0 \end{array}$$

while the dynamics of the chain of moving atoms is now the tensor right shift on the observables:

$$S \ : \quad \cdots y_{-1} \otimes y_0 \otimes y_1 \otimes y_2 \cdots \ \mapsto \ \cdots y_{-2} \otimes y_{-1} \otimes y_0 \otimes y_1 \cdots$$

Therefore, the complete dynamics for one time step is given by $\alpha := \alpha_{\text{int}} \circ S$, and can be written as

$$\begin{array}{c} \left.\begin{array}{c} B(\mathcal{H}) \\ \otimes \\ \cdots \otimes M_2 \otimes \ M_2 \end{array}\right\} \alpha_{\text{int}} \\ \otimes M_2 \otimes \cdots \\ \longrightarrow \\ S \end{array}$$

We see that the dynamics of this systems is a realization of the dynamics of a quantum Markov process of the type as discussed in Sect. 4.6.1.

4.11.3 The Jaynes–Cummings Interaction

Before investigating further this Markov process we need to be more specific about the nature of the interaction between field mode and two-level atoms. In the micromaser regime it is a good approximation to assume that the interaction is described by the Jaynes–Cummings model: On the Hilbert space $l^2(\mathbb{N}) \otimes \mathbb{C}^2$, we use the simplified Hamiltonian given by

$$\begin{aligned} \mathbb{H} = \ & \hbar\omega_F a^* a \otimes \mathbb{1} + \mathbb{1} \otimes \frac{\hbar}{2}\omega_A \sigma_z + g\hbar(a + a^*) \otimes (\sigma_+ + \sigma_-) \\ \rightsquigarrow \ & \hbar\omega_F a^* a \otimes \mathbb{1} + \mathbb{1} \otimes \frac{\hbar}{2}\omega_A \sigma_z + g\hbar(a \otimes \sigma_+ + a^* \otimes \sigma_-) \\ \rightsquigarrow \ & \hbar\omega \ \ a^* a \otimes \mathbb{1} + \mathbb{1} \otimes \frac{\hbar}{2}\omega \ \ \sigma_z + g\hbar(a \otimes \sigma_+ + a^* \otimes \sigma_-) \ . \end{aligned}$$

Here, the first line is the original Hamiltonian of a field-atom interaction, where ω_F is the frequency of the field mode, ω_A is the frequency for the transition

between the two levels of our atoms, and g is the coupling constant. In the second line this Hamiltonian is simplified by the rotating wave approximation (see p. 50), and in the third line we further assume $\omega_F = \omega_A =: \omega$. The operators σ_+ and σ_- are the raising and lowering operators of a two-level system. The Hamiltonian generates the unitary group

$$U(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathbb{H}t}$$

and $u_{\mathrm{int}} := U(t_0)$, where t_0 is the interaction time needed for one atom to pass through the cavity.

We denote by $|n\rangle \otimes |\downarrow\rangle$ and $|n\rangle \otimes |\uparrow\rangle$ the canonical basis vectors of the Hilbert space where $|n\rangle$ denotes the n-th eigenstate of the harmonic oscillator and $|\uparrow\rangle$ and $|\downarrow\rangle$ are the two eigenstates of the two-level atom. The Hilbert space decomposes into subspaces which are invariant under the Hamiltonian and the time evolution: $\mathcal{H}_0$ is the one-dimensional subspace spanned by $|0\rangle \otimes |\downarrow\rangle$, and the restriction of $\mathbb{H}$ to $\mathcal{H}_0$ is given by $\mathbb{H}_0 = 0$, hence the restriction of $U(t)$ to $\mathcal{H}_0$ is $U_0(t) = 1$. For $k \in \mathbb{N}$ denote by $\mathcal{H}_k$ the two-dimensional subspace spanned by the vectors $|k\rangle \otimes |\downarrow\rangle$ and $|k-1\rangle \otimes |\uparrow\rangle$. Then the restriction of $\mathbb{H}$ to $\mathcal{H}_k$ is given by

$$\mathbb{H}_k = \hbar \cdot \begin{pmatrix} \omega k & g\sqrt{k} \\ g\sqrt{k} & \omega k \end{pmatrix} ,$$

hence the restriction of $U(t)$ to $\mathcal{H}_k$ is

$$U_k(t) = \mathrm{e}^{\mathrm{i}\omega k t} \begin{pmatrix} \cos g\sqrt{k}t & -\mathrm{i}\sin g\sqrt{k}t \\ -\mathrm{i}\sin g\sqrt{k}t & \cos g\sqrt{k}t \end{pmatrix} .$$

Finally, if for some inverse temperatur β, $0 < \beta < \infty$, φ_β and ψ_β are the equilibrium states for the free Hamiltonian of the field mode and of the two-level-atom, respectively, then $\varphi_\beta \otimes \psi_\beta$ is invariant under the full time evolution generated by the Jaynes–Cummings interaction Hamiltonian $\mathbb{H}$ from above. Therefore, $\alpha_1 := Ad\, u_{\mathrm{int}}$ on $\mathcal{B}(\mathcal{H}) \otimes M_2$ leaves this state invariant and the dynamics of the micromaser is the dynamics of a full stationary Markov process as discussed in Sect. 4.6.1.

4.11.4 Asymptotic Completeness and Preparation of Quantum States

The long time behaviour of this system depends on whether or not a so-called *trapped state condition* is fulfilled. That means that, for some $k \in \mathbb{N}$, $g\sqrt{k}t_{\mathrm{int}}$ is an integer multiple $n\pi$ of π, for some $n \in \mathbb{N}$. In this case the transition

$$|k-1\rangle \otimes |\uparrow\rangle \longleftrightarrow |k\rangle \otimes |\downarrow\rangle$$

is blocked. Therefore, if the initial state of the micromaser has a density matrix with non-zero entries only in the upper left $k-1 \times k-1$ corner, then the atoms, in whichever state they are, will not be able to create a state in the micromaser

with more than $k-1$ photons. This has been used [33] to prepare experimentally two-photon number states: The initial state of the field mode is the vacuum, the two-level atoms are in the upper state $|\uparrow\rangle$, and the interaction time is chosen such that the transition from two to three photons is blocked. This forces the field-mode into the two-photon number state.

On the other hand, if no trapped state condition is fulfilled and all transitions are possible, then the state of the field-mode can be controlled by the states of the passing atoms [34,35]. Some mathematical theory is necessary in order to understand why this happens: The evolution $\alpha = \alpha_1 \circ S$ of this system is composed of the two parts α_1 and S. The automorphism α_1 is only a local interaction, and we may consider S as the free part of the evolution α. This resembles the situation discussed by Lax and Phillips in their scattering theory [22]. Following this idea a scattering theory for such Markov processes was developed in [18]: A Markov process as in Sect. 4.6.1 is called *asymptotically complete* if for all $x \in \mathcal{A}^2$

$$\begin{aligned} & \Phi_-(x) := \lim_{n\to\infty} S^{-n}\alpha^n(x) \qquad \text{ex. (stop)}, \\ \text{and} \qquad & \Phi_-(x) \in \mathbb{1}\otimes\mathcal{C}\ . \end{aligned}$$

Moreover, it is shown in [18] that if these conditions hold for $x \in \mathcal{A}_0 \otimes \mathbb{1}$ then they hold for arbitrary $x \in \mathcal{A}$. It should be noted, however, that for this part of our discussions all algebras should be von Neumann algebras, and the tensor product has to be the von Neumann infinite tensor product with respect to the product state. For $x \in \mathcal{A}_0$, however, we find that

$$\begin{aligned} \alpha^n(x\otimes\mathbb{1}) &= u_n^* \cdot x\otimes\mathbb{1}\cdot u_n \\ u_n &:= S^{n-1}(u_{\text{int}})\cdot S^{n-2}(u_{\text{int}})\cdot\ldots\cdot S(U_{\text{int}})\cdot u_{\text{int}}\ , \end{aligned}$$

and asymptotic completeness means that for $x \in \mathcal{A}_0$, and for very large $n \in \mathbb{N}$, there exists $x^n_{out} \in \mathcal{C}$ such that

$$\alpha^n(x\otimes\mathbb{1}) = u_n^*\cdot x\otimes\mathbb{1}\cdot u_n \approx \mathbb{1}\otimes x^n_{out}\ .$$

We translate this into the Schrödinger picture and use density matrices for the description of states. Then we find that if such a Markov process is asymptotically complete, then for any density matrix ϕ_n of $\mathcal{A}_0$ and large $n \in \mathbb{N}$ we find a density matrix ρ_0 of $\mathcal{C}$ such that

$$u_n\cdot\phi_0\otimes\rho_0\cdot u_n^* \approx \phi_n\otimes\rho'$$

for some density matrix ρ' of $\mathcal{C}$, and the choice of ρ_0 is independent of the initial state ϕ_0 on $\mathcal{A}_0$. This means that if we want to prepare a state ϕ_n on $\mathcal{A}_0$ (in our case of the field mode), then even without knowing the initial state ϕ_0 of $\mathcal{A}_0$ we can prepare an initial state ρ_0 on $\mathcal{C}$ such that the state $\phi_0\otimes\rho_0$ evolves

[2] Read "ex. (stop)" as "exists in the strong operator topology" hereafter.

after n time steps into the state ϕ_n on $\mathcal{A}_0$, and some other state ρ' of $\mathcal{C}$ which is, however, not entangled with $\mathcal{A}_0$.

This raises the question whether the Markov process of the Jaynes–Cummings-model is asymptotically complete and, if so, whether this can be used to prepare other states of the field mode. The first question has a positive answer (but the proof is difficult). Therefore, from a mathematical point of view it is possible to prepare an arbitrary state of the field-mode with arbitrary accuracy by sending suitably prepared atoms through the cavity. The second question has been investigated in [34,35]. These results show that already with a small number of atoms one can prepare interesting states of the field mode with a very high fidelity. Details can be found in [34,35]. As an illustration we give one concrete example: If the field mode is initially in the vacuum $|0\rangle$ and one wants to prepare the two-photon number state $|2\rangle$ with 4 incoming atoms, then when choosing an optimal interaction time t_{int} one can prepare the state $|2\rangle$ with a fidelity of 99.87%, if the four atoms are prepared in the state

$$\begin{aligned}|\psi_0\rangle = \quad & \sqrt{0.867}|\uparrow\rangle|\uparrow\rangle|\downarrow\rangle|\downarrow\rangle \\ & +\sqrt{0.069}|\uparrow\rangle|\downarrow\rangle|\uparrow\rangle|\downarrow\rangle \\ & -\sqrt{0.052}|\downarrow\rangle|\uparrow\rangle|\uparrow\rangle|\downarrow\rangle \\ & +\sqrt{0.005}|\uparrow\rangle|\downarrow\rangle|\downarrow\rangle|\uparrow\rangle \\ & -\sqrt{0.004}|\downarrow\rangle|\uparrow\rangle|\downarrow\rangle|\uparrow\rangle \\ & +\sqrt{0.003}|\downarrow\rangle|\downarrow\rangle|\uparrow\rangle|\uparrow\rangle \quad .\end{aligned}$$

References

1. L. Accardi, F. Frigerio, J. T. Lewis: Publ. RIMS **18**, 97-133 (1982)
2. H. J. Carmichael: *An Open Systems Approach to Quantum Optics* (Springer, Berlin 1993)
3. E. B. Davies: *Quantum Theory of Open Systems* (Academic Press, London 1976)
4. E. B. Davies: *One Parameter Semigroups* (Academic Press, London 1980)
5. R. Durett: *Probabilty: Theory and Examples*, 2nd edn. (Duxbury Press, Belmont 1996)
6. D. Evans, J. T. Lewis: Comm. Dublin Inst. Adv. Stud. Ser A 24 (1977)
7. W. Feller: *An Introduction to Probability Theory and its Applications, Vol. I* (John Wiley, New York 1977)
8. C. W. Gardiner: *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, Berlin 1983)
9. C. W. Gardiner, P. Zoller: *Quantum Noise*, 2nd edn. (Springer, Berlin 2000)
10. V. Gorini, A. Kossakowski, E. C. G. Sudarshan: J. Math. Phys. **17**, 821-825 (1976)
11. B. Kitchens: *Symbolic Dynamics: One-Sided, Two-Sided and Countable State Markov Shifts* (Springer, Berlin 1998)
12. K. Kraus: Ann. Phys. (New York) **64**, 311-335 (1971)

13. B. Kümmerer: 'Examples of Markov dilations over the 2×2-matrices'. In: *Quantum Probability and Applications I, Lecture Notes in Mathematics 1055* (Springer, Berlin 1984) pp. 228-244
14. B. Kümmerer: J. Funct. Anal. **63**, 139-177 (1985)
15. B. Kümmerer: 'Lectures on Stationary Processes in Quantum Probability'. In: *Quantum Probability Communications XII*, in print
16. B. Kümmerer, H. Maassen: Commun. Math. Phys. **109**, 1-22 (1987)
17. B. Kümmerer, H. Maassen: 'Elements of quantum probability'. In: *Quantum Probability Communications X* (World Scientific, Singapore 1998) pp. 73-100
18. B. Kümmerer, H. Maassen: Infinite Dimensional Analysis, Quantum Probability and Related Topics **3**(1), 161-176 (2000)
19. B. Kümmerer, H. Maassen: *An Ergodic Theorem for Quantum Counting Processes*. Preprint (2001)
20. B. Kümmerer, H. Maassen, to be published.
21. B. Kümmerer, W. Schröder: Comm. Math. Phys. **90**, 251-262 (1983)
22. P. D. Lax, R. S. Phillips: *Scattering Theory* (Academic Press, New York 1967)
23. G. Lindblad: Commun. Math. Phys. **48**, 119-130 (1976)
24. J. von Neumann: *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin 1932)
25. K. R. Parthasarathy: *An Introduction to Quantum Stochastic Calculus* (Birkhäuser Verlag, Basel 1992)
26. K. Petersen: *Ergodic Theory* (Cambridge University Press, Cambridge 1983)
27. S. M. Ross: *Probability Models*, 7th edn. (Academic Press, San Diego 2000)
28. M. Reed, B. Simon: *Methods of Modern Mathematical Physics. Vol. I: Functional Analysis* (Academic Press, New York 1972)
29. E. Seneta: *Non-negative Matrices and Markov Chains* (Springer, New York 1981)
30. B. Sz.-Nagy, C. Foias: *Harmonic Analysis of Operators on Hilbert Space* (North Holland, Amsterdam 1970)
31. M. Takesaki: *Theory of Operator Algebras I* (Springer, New York 1979)
32. M. Takesaki: J. Funct. Anal. **9**, 306-321 (1971)
33. B. T. H. Varcoe *et al.*: Nature **403**, 743 (2000)
34. T. Wellens, A. Buchleitner and B. Kümmerer, H. Maassen: Quantum state preparation via asymptotic completeness. Phys. Rev. Lett. **85**, 3361 (2000)
35. T. Wellens: Entanglement and Control of Quantum States. Dissertation, Ludwig-Maximilians-Universität, München (2002)

5 Decoherence in Quantum Physics

Walter T. Strunz

Theoretische Quantendynamik, Fakultät für Physik, Universität Freiburg,
Hermann-Herder-Str. 3, D-79104 Freiburg, Germany

5.1 Introduction

In quantum mechanics, two states $|\psi_1\rangle$ and $|\psi_2\rangle$ of a system may be superposed to give a new state $|\psi\rangle = (|\psi_1\rangle + |\psi_2\rangle)/\sqrt{2}$. Sometimes, due to environmental influences, such a superposition is not dynamically robust and decays into a mixture $\rho = \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)$. These lectures are concerned with the nature of robust states and the time scales involved during the transition from the coherent superposition to the mixture, i.e the time scale of decoherence.

Coherence in a quantum system may be lost due to several mechanisms: the system under study may be subjected to the influence of some fluctuating, external classical fields such that the dynamics of a state ψ_t is unitary, yet stochastic. Certain interferences will be washed out by averaging over these fluctuations, a manifestation of decoherence. Closely related, uncertainty in initial conditions will also lead to a suppression of interference phenomena resulting from the average over a distribution of initial states. None of these reasons for decoherence is truly quantum and may also be encountered as the origin of decoherence in classical systems, as in light interferometry, for instance. A profoundly quantum cause for decoherence is entanglement. A quantum system interacting with its environment will become entangled and will thus no longer be described by a single state vector, even if both, system and environment are in a single pure state. The reduced density operator of the open system will generally be a mixture due to entanglement. It is this latter, genuinely quantum origin of decoherence that will be our major concern in these lectures, see also [1,2,3,4,5].

5.1.1 Open Quantum Systems

Standard quantum theory applies to closed systems only. If we are to describe an open quantum system, a standard approach [6,7] is to specify a total Hamiltonian consisting of the Hamiltonian of the open "system", its "environment", and their interaction:

$$H_{\rm tot} = H_{\rm sys} + H_{\rm env} + H_{\rm int} \ . \tag{5.1}$$

In order to determine the dynamics, ideally, we would like to solve Schrödinger's equation for system and environment,

$$\mathrm{i}\hbar\partial_t \Psi_{\rm tot}(t) = H_{\rm tot}\Psi_{\rm tot}(t) \ , \tag{5.2}$$

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 199–233, 2002.

for a pure initial total state Ψ_0, or equivalently, von Neumann's equation

$$\mathrm{i}\hbar\partial_t \rho_{\mathrm{tot}}(t) = [H_{\mathrm{tot}}, \rho_{\mathrm{tot}}(t)]$$

which applies to a mixed initial total state ρ_0 too.

In almost all cases, however, the determination of the solution of these equations is beyond reach. As we are usually interested in properties of the open "system" only, we may concentrate on the time evolution of the reduced density operator

$$\rho(t) = \mathrm{tr}_{\mathrm{env}}\left(\rho_{\mathrm{tot}}(t)\right) = \mathrm{tr}_{\mathrm{env}}\left(|\Psi_{\mathrm{tot}}(t)\rangle\langle\Psi_{\mathrm{tot}}(t)|\right) \ , \tag{5.4}$$

where the last equality holds for a pure total state of system and environment. The reduced density operator, whose time dependence we often abbreviate by the notation $\rho_t \equiv \rho(t)$ will in general be a mixed state due to *entanglement* between "system" and "environment", even if $\rho_{\mathrm{tot}} = |\Psi_{\mathrm{tot}}\rangle\langle\Psi_{\mathrm{tot}}|$ is pure.

Open system dynamics [6,7] refers to the time evolution of ρ_t. In many cases, notably in quantum optics applications when the environment is the quantized electromagnetic field, it is possible to derive certain *master equations* for the time evolution of the open system's density operator,

$$\partial_t \rho_t = \mathcal{L}\rho_t, \tag{5.5}$$

with $\mathcal{L}$ denoting the generator of this dissipative, non-unitary dynamics (see Chapt. 2). Very often, this master equation is of the so-called Lindblad form, ensuring that general properties of density operators are preserved under time evolution.

Let us expand the reduced density operator in a certain basis $|\phi_n\rangle$,

$$\rho = \sum_{nm} \rho_{nm} |\phi_n\rangle\langle\phi_m| \ . \tag{5.6}$$

The diagonal elements $\rho_{nn} = \langle\phi_n|\rho|\phi_n\rangle = p_n$ may be interpreted as probabilities, namely the probability that the system is in state $|\phi_n\rangle$.

The off-diagonal elements $\rho_{nm} = \langle\phi_n|\rho|\phi_m\rangle, n \neq m$ are called *coherences* and indicate that the state of the quantum system has contributions from a coherent superposition $c_m|\phi_m\rangle + c_n|\phi_n\rangle$. Note that only the probabilities p_n have a "classical" meaning.

Decoherence is the dynamical loss of coherences, $\rho_{nm}(t) \to 0$ for $n \neq m$ and $t \to \infty$, here due to the coupled dynamics of the open system and its environment. Obvious questions arise that will be addressed during these lectures: What determines the basis $|\phi_n\rangle$ with respect to which off-diagonal elements in the reduced density operator disappear? What is the time scale of their disappearance? In answering these questions we will find that it proves difficult to preserve coherences in quantum systems that interact with a many-degree of freedom reservoir. In particular, we will see how decoherence explains why we do not encounter quantum superpositions at the macroscopic level.

5.1.2 Decoherence: Two Simple Examples

Damped Electromagnetic Field Mode. A simple example with the nice feature of experimental access is the decoherence of the state of an electromagnetic field mode inside a cavity. Due to losses through the cavity mirrors, photons may leak to the field modes outside the cavity which play the role of the "environment". This standard model is discussed in Chapt. 2 and in all quantum optics text books [6], and will be used here to reveal important aspects of decoherence that will later on be addressed in a more general framework. The system is a single mode of the electromagnetic field with Hamiltonian $H_{\rm sys} = \Omega a^\dagger a$. The loss of photons to the outside world leads to an exponential decay of the number of photons $\langle N(t)\rangle = N_0 \mathrm{e}^{-\gamma t}$ which introduces a relaxation time constant γ^{-1} in addition to the free system time scale Ω^{-1}.

Master Equation. In standard Born-Markov approximation, and at zero temperature $T = 0$, the master equation of the damped field mode takes the Lindblad form [6] (see (2.1) with $\omega \to \Omega$, $A \to \gamma$, $\nu = 0$)

$$\partial_t \rho = -\mathrm{i}[\Omega a^\dagger a, \rho] + \frac{\gamma}{2}\left([a, \rho a^\dagger] + [a\rho, a^\dagger]\right) \ . \tag{5.7}$$

With a pure initial state $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and at zero temperature, entanglement between "system"and "environment" is the only cause of decoherence. Indeed, in almost all circumstances, one expects that such dissipative dynamics leads to a mixed state ρ_t, even if the initial state ρ_0 is pure.

Yet in this simple example, we make a remarkable observation concerning *coherent states* $|z\rangle = e^{-\frac{1}{2}|z|^2 + z a^\dagger}|0\rangle$. Coherent states are eigenstates of the annihilation operator $a|z\rangle = z|z\rangle$, with $z \in \mathbb{C}$.

Let us choose $\rho_0 = |z_0\rangle\langle z_0|$ to be a pure coherent state. Under time evolution according to (5.7), we find that $\rho_t = |z(t)\rangle\langle z(t)|$ remains a pure state solution with $z(t) = \mathrm{e}^{-\mathrm{i}\Omega t}\mathrm{e}^{-\gamma t/2} z_0$ describing the damped, oscillating motion of the field amplitude. We conclude that the master equation (5.7) has very specific pure state solutions.

An initial mixture of two pure coherent states $\rho_0 = \frac{1}{2}|z_1\rangle\langle z_1| + \frac{1}{2}|z_2\rangle\langle z_2|$, remains that mixture due to linearity; we find $\rho_t = \frac{1}{2}|z_1(t)\rangle\langle z_1(t)| + \frac{1}{2}|z_2(t)\rangle\langle z_2(t)|$, with $z_1(t), z_2(t)$ again describing the damped, oscillating motion of a field amplitude as above.

Let us now investigate the fate of an initial *coherent* superposition of coherent states, $\rho_0 = |\psi_0\rangle\langle\psi_0|$ with $|\psi_0\rangle = \left(|z_1\rangle + |z_2\rangle\right)/\sqrt{2}$ (strictly speaking, the normalization factor is more involved than just $1/\sqrt{2}$ - we will, however always assume $|z_1 - z_2| \gg 1$ so that the deviation from our choice is negligible). In its full beauty, the initial reduced density operator consists of four terms, $\rho_0 = \frac{1}{2}|z_1\rangle\langle z_1| + \frac{1}{2}|z_2\rangle\langle z_2| + \frac{1}{2}|z_1\rangle\langle z_2| + \frac{1}{2}|z_2\rangle\langle z_1|$.

Time evolution according to the master equation (5.7) leads to

$$\begin{aligned}\rho_t = &\frac{1}{2}|z_1(t)\rangle\langle z_1(t)| + \frac{1}{2}|z_2(t)\rangle\langle z_2(t)| \\ &+\frac{1}{2}f(t)|z_1(t)\rangle\langle z_2(t)| + \frac{1}{2}f^*(t)|z_2(t)\rangle\langle z_1(t)| \ ,\end{aligned} \tag{5.8}$$

where $z_1(t)$ and $z_2(t)$ again follow the oscillating, damped motion. The cross terms (the "coherences") are weighted by an amplitude $f(t)$ with

$$|f(t)| = \exp\left(-\frac{1}{2}|z_1(0) - z_2(0)|^2 \left(1 - \mathrm{e}^{-\gamma t}\right)\right) . \tag{5.9}$$

Obviously, whenever the initial "distance" between the superposed coherent states is large, $D^2 \equiv |z_1(0) - z_2(0)|^2 \gg 1$, their mutual coherence vanishes since $|f(t)| \to 0$. Thus, the initially fully coherent reduced density operator becomes the mixed state $\rho_t = \frac{1}{2}|z_1(t)\rangle\langle z_1(t)| + \frac{1}{2}|z_2(t)\rangle\langle z_2(t)|$ under time evolution, if only $D^2 = |z_1(0) - z_2(0)|^2 \gg 1$. The larger D, the more rapidly the coherences will disappear.

Of particular interest is the time scale of their disappearance: Assume $\gamma t \ll 1$, so that almost no "damping" or even "motion" has occurred yet. Still, decoherence with $|f(t)| \sim \mathrm{e}^{-\gamma D^2 t/2}$ will be effective as long as the initial distance D is large. For macroscopically distinct coherent states $\{|z_1\rangle, |z_2\rangle\}$ a huge separation of time scales between decoherence and damping results:

$$\frac{\text{decoherence rate}}{\text{relaxation rate}} \sim D^2.$$

Upon the identification $z = (q + \mathrm{i}p)/\sqrt{2}$, the distance $D = |z_1 - z_2|$ is just a "phase space distance" $D^2 = (|q_1 - q_2|^2 + |p_1 - p_2|^2)/2$. The square D^2 of this distance can easily be much, much larger than unity for macroscopically distinct states, since the phase space coordinates (q, p) are measured in quantum oscillator units involving Planck's constant $\hbar$. Thus values of astonishing magnitude like $D^2 \sim 10^{40}$ may appear (which nevertheless have to be taken very cautiously, as we will explain in more detail later).

Thus, decoherence for superpositions of macroscopically distinct states ($D \gg 1$) may be incredibly effective, even though damping for one of the components of the superposition is hardly noticeable. This separation of time scales between damping and decoherence is the reason why decoherence deserves special attention in open quantum system dynamics. For two-level systems, in contrast, the decay of the off-diagonal terms differs by a factor of two only from the decay of the diagonal terms and no special distinction between decoherence and dissipation seems necessary. As soon as a quantum system allows for a meaningful classical limit, however, superpositions of then macroscopically distinct states decohere towards mixtures almost instantaneously. This may be seen as the reason why it is so difficult to extend coherent quantum phenomena to the macroscopic domain, and is the reason for the absence of quantum superpositions ("Schrödinger cat states") from the macroworld.

Wigner Function. It is instructive to display decoherence using a phase space representation of the density operator in terms of the Wigner function $W(q, p)$.

The Wigner function is the integral transform

$$W(q,p) = \frac{1}{2\pi\hbar} \int \mathrm{d}x \, \langle q - \frac{x}{2}|\rho|q + \frac{x}{2}\rangle \, \mathrm{e}^{\mathrm{i}px/\hbar} \, . \tag{5.11}$$

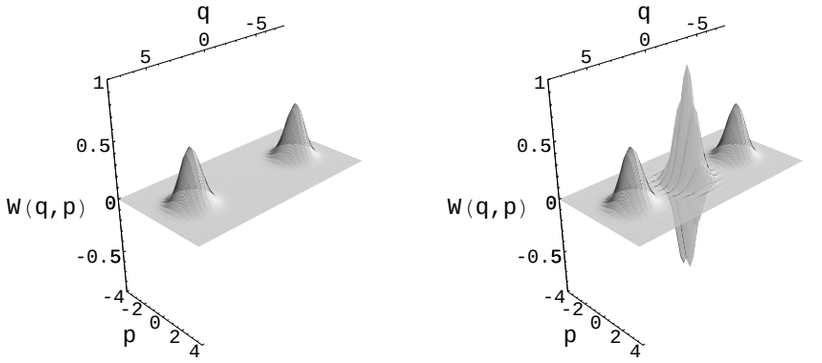


Fig. 5.1. (**left**) Wigner function of a mixture of two coherent states. (**right**) Wigner function of a coherent superposition of two coherent states

Upon integration over one of the variables, we get the probability distribution of the other, $\int \mathrm{d}p W(q,p) = \langle q|\rho|q\rangle$ and $\int \mathrm{d}q W(q,p) = \langle p|\rho|p\rangle$. (From now on we use phase space units such that $\hbar = 1$.)

The Wigner function of a coherent state $\rho = |z\rangle\langle z|$ with $z = (\bar{q} + i\bar{p})/\sqrt{2}$ is a Gaussian, centered at $(\bar{q}, \bar{p})$,

$$W(q,p) = \frac{1}{\pi} \exp\Big(- \Big((q-\bar{q})^2 + (p-\bar{p})^2 \Big) \Big) \,. \tag{5.12}$$

A mixture of coherent states with $z_1 = z = -z_2$, i.e. $\rho = \frac{1}{2}|z\rangle\langle z| + \frac{1}{2}|-z\rangle\langle -z|$ consists of two Gaussians with opposite centroids, $W(q,p) = \frac{1}{2}(W_+ + W_-)$, as displayed in the left part of Fig. 5.1, for $z = 5/\sqrt{2}$.

By contrast, the Wigner function of a coherent superposition of two coherent states, $|\psi\rangle = (|z\rangle + |-z\rangle)/\sqrt{2}$ (with $|z| \gg 1$) shows an additional oscillatory pattern, centered between the two Gaussians (right part of Fig. 5.1, again $z = 5/\sqrt{2}$).

Naturally, the oscillating pattern originates from the crossterms in the reduced density operator and thus from the coherence between the superposed wave packets. We find $W = \frac{1}{2}(W_+ + W_-) + W_{+-}$ with

$$W_{+-} = \frac{1}{\pi} \exp\Big(- (q^2 + p^2) \Big) \cos(D \cdot p) \tag{5.13}$$

where, as before, $D = |z_2 - z_1| = 2|z|$. Coherence between the superposed wave packets results in an oscillatory (sometimes negative) Wigner function near $(q,p) \approx 0$. The oscillations are the more rapid, the larger D, i.e. the further the superposed coherent states are apart.

If these Wigner functions represent states of the electromagnetic field mode as discussed previously, we see that decoherence in that system is nothing but the transition from the right to the left in Fig. 5.1, i.e. the disappearance of

the oscillating pattern, as a result of quantum dynamics of both, system and environment. The decay of the oscillating pattern is governed by the exponential $\mathrm{e}^{-\gamma D^2 t/2}$ for $\gamma t \ll 1$. This can be a very, very short scale if $D \gg 1$. So short in fact, that the Gaussians W_+ and W_- themselves do not move visibly.

We emphasize again that the underlying reason for decoherence is entanglement formation between system and environment in the evolution of the total state $\Psi(t)$.

Schrödinger Equation for "System" and "Environment", and Robust States. Further insight is gained by discussing decoherence in this simple example from the point of view of the full dynamics of system and environment.

The underlying master equation (5.7) may be derived from the total Hamiltonian

$$\begin{aligned} H_{\mathrm{tot}} &= \Omega a^\dagger a + \sum_\lambda g_\lambda \left(a b_\lambda^\dagger + a^\dagger b_\lambda \right) + \sum_\lambda \omega_\lambda b_\lambda^\dagger b_\lambda \\ &= \Omega a^\dagger a + a B^\dagger + a^\dagger B + \sum_\lambda \omega_\lambda b_\lambda^\dagger b_\lambda \\ &= H_{\mathrm{sys}} + H_{\mathrm{int}} + H_{\mathrm{env}} \ , \end{aligned} \tag{5.14}$$

with the bath part of the interaction $B = \sum_\lambda g_\lambda b_\lambda$. In order to arrive at (5.7), one has to assume that the bath correlation function is delta-like,

$$\alpha(t-s) = \langle B(t) B^\dagger(s) \rangle \approx \gamma \delta(t-s) \ , \tag{5.15}$$

where $B(t) = \mathrm{e}^{\mathrm{i}H_{\mathrm{env}}t} B \mathrm{e}^{-\mathrm{i}H_{\mathrm{env}}t} = \sum_\lambda g_\lambda b_\lambda \mathrm{e}^{-\mathrm{i}\omega_\lambda t}$. With our choice of the initial condition as before, i.e. assuming the special coherent state $|z\rangle$ as system part,

$$|\Psi_0\rangle = \underbrace{|z\rangle}_{\mathrm{sys}} \underbrace{|0\rangle|0\rangle|0\rangle|0\rangle \cdots}_{\mathrm{env}} , \tag{5.16}$$

the total Schrödinger equation $\mathrm{i}\hbar \partial_t \Psi_t = H_{tot} \Psi_t$ has a product state solution

$$|\Psi_t\rangle = |z(t)\rangle |\beta_1(t)\rangle \cdots |\beta_\lambda(t)\rangle \cdots , \tag{5.17}$$

where all states are coherent states of the oscillators.

Remarkably, (5.17) represents a particular solution of Schrödinger's equation *without* entanglement between system and environment, for all times. Here, as before, $z(t) = \mathrm{e}^{(-\mathrm{i}\Omega - \gamma/2)t} z(0)$, and

$$\beta_\lambda(t) = -\mathrm{i} g_\lambda \int_0^t \mathrm{d}s \, \mathrm{e}^{-\mathrm{i}\omega_\lambda (t-s)} z(s) = -\mathrm{i} z(0) g_\lambda \int_0^t \mathrm{d}s \, \mathrm{e}^{-\mathrm{i}\omega_\lambda (t-s) - \mathrm{i}\Omega s - \gamma s/2} \ . \tag{5.18}$$

The second expression highlights how the motion of the environmental oscillators depends on the initial state $|z(0)\rangle$ of the central oscillator.

Linearity demands that an initial superposition of coherent states,

$$|\Psi_0\rangle = \underbrace{\frac{1}{\sqrt{2}}(|z_1\rangle + |z_2\rangle)}_{\text{sys}} \cdot \underbrace{|0\rangle|0\rangle|0\rangle|0\rangle\cdots}_{\text{env}}, \tag{5.19}$$

evolves into an entangled state,

$$\begin{aligned} |\Psi_t\rangle = \frac{1}{\sqrt{2}}|z_1(t)\rangle|\beta^1_\lambda(t)\rangle\cdots \\ +\frac{1}{\sqrt{2}}|z_2(t)\rangle|\beta^2_\lambda(t)\rangle\cdots\ . \end{aligned} \tag{5.20}$$

We conclude that some initial states of the open system lead to entanglement between system and environment, while some (very special ones) do not. This observation leads to the notion of robust (or *preferred*) states of the open system: robust states are those system states that lead to little (or no) entanglement.

From the total state (5.20) we may determine the reduced density operator,

$$\begin{aligned} \rho_{\text{red}}(t) = \frac{1}{2}|z_1(t)\rangle\langle z_1(t)| + \frac{1}{2}|z_2(t)\rangle\langle z_2(t)| + \frac{1}{2}\underbrace{\langle\beta^2(t)|\beta^1(t)\rangle}_{f(t)}|z_1(t)\rangle\langle z_2(t)| \\ +\frac{1}{2}\langle\beta^1(t)|\beta^2(t)\rangle|z_2(t)\rangle\langle z_1(t)|\ , \end{aligned} \tag{5.21}$$

which allows to identify the coefficient $f(t)$ in (5.8) as the overlap of the corresponding environmental states $|\beta^1_\lambda(t)\rangle$ and $|\beta^2_\lambda(t)\rangle$. As they evolve differently according to (5.18) (since $|z_1(0) - z_2(0)| \gg 1$), the overlap $f(t) = \langle\beta^2(t)|\beta^1(t)\rangle$ approaches zero rapidly and the coherences in the reduced density operator disappear. We see that as soon as there is enough information in the environment to be able to deduce the system state, i.e. $\langle\beta^1|\beta^2\rangle \approx 0$, the coherence is lost.

A Remark on Robust States. The damped harmonic oscillator model singles out coherent states $|z(t)\rangle$ of the system because the total state $|\Psi_t\rangle = |z(t)\rangle|\beta(t)\rangle$ remains a product state under time evolution. The amplitude evolves according to the classical expectation, $z(t) = \mathrm{e}^{(-\mathrm{i}\Omega-\gamma/2)t}z(0)$. Interestingly, there is a closed evolution equation for the system state $|z(t)\rangle = \mathrm{e}^{-\frac{1}{2}|z(t)|^2+z(t)a^\dagger}|0\rangle$. With $|z|^2(t) = \mathrm{e}^{-\gamma t}|z|^2(0)$, we get $\partial_t|z|^2 = -\gamma|z|^2$ and

$$\partial_t|z(t)\rangle = \left(-\mathrm{i}\Omega - \frac{\gamma}{2}\right)a^\dagger \underbrace{z(t)}_{a}|z(t)\rangle + \frac{\gamma}{2}\underbrace{|z(t)|^2}_{\langle z(t)|a^\dagger a|z(t)\rangle}|z(t)\rangle. \tag{5.22}$$

Apparently, these particular states $|\psi_t\rangle = |z(t)\rangle$ are solutions of a *nonlinear* Schrödinger equation (see also Sect. 5.3).

$$\partial_t|\psi_t\rangle = -\mathrm{i}\Omega a^\dagger a|\psi_t\rangle - \frac{\gamma}{2}\left(a^\dagger a - \langle a^\dagger a\rangle\right)|\psi_t\rangle, \tag{5.23}$$

where $\langle a^\dagger a\rangle = \langle\psi_t|a^\dagger a|\psi_t\rangle$. Not surprisingly, robust states satisfy a *nonlinear* evolution equation since, as we have seen, an initial superposition of robust states leads to entanglement and can thus no longer be a robust state.

The question arises why in this case coherent states $|z\rangle$ are robust. Clearly, "robustness" of states depends on system Hamiltonian $H_{\rm sys}$ and coupling $H_{\rm int}$ to the environment. Being a model of linearly coupled harmonic oscillators, coherent states are singled out.

Dephasing. As a second simple example consider a quantum system with Hamiltonian $H_{\rm sys}$ that is coupled to the environment through its energy $H_{\rm sys}$. Then the system energy is a conserved quantity, and there is no dissipation, i.e. no energy lost to the environment. Nevertheless, phase coherences between different energy eigenstates are destroyed. The total Hamiltonian of system and environment reads

$$H_{\rm tot} = H_{\rm sys} + H_{\rm sys}\sum_\lambda g_\lambda\left(b_\lambda + b_\lambda^\dagger\right) + \sum_\lambda \omega_\lambda b_\lambda^\dagger b_\lambda\ . \tag{5.24}$$

We may either determine the total state as in the first example, or derive a master equation for the evolution of the reduced density operator using standard methods (ie, the Markov approximation, see Chapt. 2). One arrives at

$$\partial_t\rho = -\frac{\rm i}{\hbar}\Big[H,\rho\Big] - \kappa\Big[H,[H,\rho]\Big]. \tag{5.25}$$

With $H_{\rm sys}|\phi_n\rangle = E_n|\phi_n\rangle$, energy eigenstates of the system Hamiltonian will turn out to be robust here: if $\rho_0 = |\phi_n\rangle\langle\phi_n|$, then $[H_{\rm sys},\rho_0]=0$ and $\rho_t = \rho_0 = |\phi_n\rangle\langle\phi_n|$ for all times. The energy eigenstates remain pure for all times, these states do not lead to entanglement between system and environment. Thus eigenstates $|\phi_n\rangle$ of $H_{\rm sys}$ are robust. In general, we can write $\rho_0 = \sum_{nm}\rho_{nm}|\phi_n\rangle\langle\phi_m|$, and find

$$\rho_t = \sum_{nm}\rho_{nm}{\rm e}^{-\kappa(E_n-E_m)^2t/\hbar}\ {\rm e}^{-{\rm i}(E_n-E_m)t}\ |\phi_n\rangle\langle\phi_m|\ . \tag{5.26}$$

As in the first example of the damped cavity mode, coherences between robust states are lost: $|\rho_{nm}(t)| = |\langle\phi_n|\rho(t)|\phi_m\rangle| = {\rm e}^{-\kappa(E_n-E_m)^2t} \to 0$ rapidly, if $D^2 = (E_n-E_m)^2 \gg 1$. Again, the decoherence time scale is proportional to the square of a "distance" $D = |E_n - E_m|$ between the superposed robust states. The further they are apart with respect to this distance, the more rapidly their coherence is lost, just as in the case of the damped harmonic oscillator. Here, however, due to the difference in the coupling between system and environment, energy eigenstates rather than coherent states are singled out to be robust.

5.1.3 First Conclusions

In general, open quantum system dynamics leads to ever growing entanglement between "system" and "environment". However, there may be very particular

system states $|\psi_i\rangle$, such that upon time evolution, there is only little (or no) entanglement between "system" and "environment". These states are singled out and called *robust* (or *preferred*) states of the open system. They are determined through the system part in the interaction Hamiltonian and the system Hamiltonian: In the case of an interaction coupling to energy $H_{\rm sys}$, energy eigenstates were robust. In the case of the damped harmonic oscillator with coupling to annihilation and creation operators $a, a^\dagger$ and $H_{\rm sys} = \Omega a^\dagger a$, coherent states were identified to be robust. Superpositions $|\psi\rangle = \big(|\psi_1\rangle + |\psi_2\rangle\big)/\sqrt{2}$ of such robust states decay towards the mixture $\rho = \frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2|$ on a decoherence time scale that it proportional to an inverse "distance" squared. The decoherence time can be very, very, very short for macroscopic distance D (see Sect. 5.2), since D is measured with respect to a quantum reference scale. For macroscopic D, such decoherence appears instantaneous. For the damped harmonic oscillator, we saw that $\frac{\text{decoherence rate}}{\text{relaxation rate}} \sim D^2$. An arbitrary initial state ρ_0, after the decoherence time scale, will be given by a mixture of these robust states: $\rho(t) \approx \sum_i p_i |\psi_i(t)\rangle\langle\psi_i(t)|$, with time independent probabilities p_i.

5.1.4 Decoherence and the "Measurement Problem"

Decoherence is often employed to explain the absence of pointer superpositions in a quantum measurement. Assume an initial product state of the system to be measured in state ψ_n and the pointer of the apparatus in its ground state ϕ_0,

$$\Psi_0 = |\psi_n\rangle\,|\phi_0\rangle. \tag{5.27}$$

A measurement apparatus should be such that upon time evolution,

$$\Psi_t = |\psi_n\rangle\,|\phi_n\rangle, \tag{5.28}$$

meaning that the state of the system is unchanged, yet the pointer now points to result "n", i.e. is in state $|\phi_n\rangle$.

Linearity demands that if the system is initially in a superposition of microstates,

$$\Psi_0 = \left(\sum_n a_n|\psi_n\rangle\right)|\phi_0\rangle, \tag{5.29}$$

time evolution leads to

$$\Psi_t = \sum a_n|\psi_n\rangle|\phi_n\rangle, \tag{5.30}$$

i.e. a coherent superposition involving different pointer states.

According to the standard measurement formulation, von Neumann collapse demands that in fact, after the measurement, the state of the system and the apparatus is described by the mixture of the measurement outcomes $|\phi_n\rangle\langle\phi_n|$, each correlated with the appropriate system state $|\psi_n\rangle\langle\psi_n|$, i.e.

$$\rho_t = \sum_n |a_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |\phi_n\rangle\langle\phi_n| \neq |\Psi_t\rangle\langle\Psi_t|\ , \tag{5.31}$$

This is clearly different from the coherent superposition Ψ_t.

The transition from the superposition Ψ_t to the mixture ρ_t can be explained with decoherence. A pointer of a measurement apparatus is a macroscopic degree of freedom, thus unavoidably coupled to a many degree of freedom environment, initially in state $|E_0\rangle$. The total state of system, pointer and environment is therefore

$$\tilde{\Psi}_0 = \left(\sum_n a_n |\psi_n\rangle \right) |\phi_0\rangle |E_0\rangle \tag{5.32}$$

and evolves into

$$\tilde{\Psi}_t = \sum_n a_n |\psi_n\rangle |\phi_n\rangle |E_n\rangle, \tag{5.33}$$

Now if $\langle E_n | E_m \rangle = \delta_{nm}$ (which will be the case whenever pointer states are *robust*), the reduced density operator of system and pointer is

$$\rho_{\mathrm{SP}} = \mathrm{tr}_{\mathrm{env}} \Big(|\Psi_t\rangle\langle\Psi_t| \Big) = \sum_n |a_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |\phi_n\rangle\langle\phi_n| \,, \tag{5.34}$$

which is the expected mixture of pointer states. Thus, for the system–apparatus–subsystem, decoherence has the same effect as the "collapse postulate" of textbook quantum measurement theory. The term "collapse" is indeed justified, since the transition from the coherent superposition to the mixture is instantaneous for a macroscopic distance D between robust pointer states (as a result of Schrödinger's equation). We emphasize, however, that the mixture ρ_{SP} arises from a reduction of the possible observations. The total state Ψ still contains the full coherences, whereas the measurement mixture (5.31) is supposed to describe genuine alternatives. Sometimes this latter is called a "real" mixture, while ρ_{SP} is referred to as an "apparent" mixture. Physically, as long as observations are restricted to system and apparatus, these mixtures cannot be distinguished, of course. Within the framework of the linear Schrödinger theory alone, which is the basis for environment-induced decoherence discussed here, individual, alternative outcomes in measurement situations cannot be obtained dynamically, and there remains room for heated debates. This situation has caused experimenters to investigate decoherence in measurement-type situations in detail, as will be explained next.

5.1.5 The Paris and Boulder Experiments

The first experiment to investigate decoherence dynamics in a very controlled way has been performed by Haroche's group at the ENS in Paris [8], in an article entitled "Observing the progressive decoherence of the 'Meter' in a Quantum measurement". They prepared a coherent initial state $|z\rangle$ of an electromagnetic field mode inside a microwave cavity and sent single Rydberg atoms through the cavity in order to manipulate the field state inside. The cavity mirrors are not perfect and thus the dynamics of the field mode is nothing but a realization of the damped harmonic oscillator discussed earlier.

Two Rydberg levels ($|g\rangle, |e\rangle$) of the atoms are relevant. The interaction between atoms and field mode is made purely dispersive such that upon passing the cavity, the atomic state does not change, yet the field is rotated by an angle $\pm\phi$, depending on the state of the atom:

$$\begin{aligned} |e\rangle|z\rangle &\longrightarrow |e\rangle|z\mathrm{e}^{\mathrm{i}\phi}\rangle| \,, \\ |g\rangle|z\rangle &\longrightarrow |g\rangle|z\mathrm{e}^{-\mathrm{i}\phi}\rangle. \end{aligned} \tag{5.35}$$

The angle $\phi = \Omega^2 t_i/\delta$ is determined through the Rabi frequency Ω, the detuning δ between cavity frequency and transition frequency between $|g\rangle$ and $|e\rangle$ and the interaction time t_i, which can be adjusted through the velocity of the atoms. Before and after the atoms pass the cavity, they may interact with a $\pi/2$ pulse such that superpositions of the two atomic states may be produced: $|e\rangle \longrightarrow (|g\rangle + |e\rangle)/\sqrt{2}$, or $|g\rangle \longrightarrow (|g\rangle - |e\rangle)/\sqrt{2}$.

The experiment first produces a coherent superposition of two coherent states of the field. For that, an atom is subjected to a $\pi/2$ pulse (R_1), sent through the cavity (C), subjected to the second $\pi/2$ pulse (R_2), and finally detected in the ground state $|g\rangle$, in half of the cases. The state of atom and field is thus transformed according to

$$\begin{aligned} |e\rangle|z\rangle &\overset{R_1}{\rightarrow} \frac{1}{\sqrt{2}}\Big(|g\rangle + |e\rangle\Big)|z\rangle \\ &\overset{C}{\rightarrow} \frac{1}{\sqrt{2}}\left(|g\rangle|z\mathrm{e}^{-\mathrm{i}\phi}\rangle + |e\rangle|z\mathrm{e}^{\mathrm{i}\phi}\rangle\right) \\ &\overset{R_2}{\rightarrow} \frac{1}{2}|e\rangle\left(|z\mathrm{e}^{\mathrm{i}\phi}\rangle - |z\mathrm{e}^{-\mathrm{i}\phi}\rangle\right) + \frac{1}{2}|g\rangle\left(|z\mathrm{e}^{\mathrm{i}\phi}\rangle + |z\mathrm{e}^{-\mathrm{i}\phi}\rangle\right) \,. \end{aligned} \tag{5.36}$$

After detection of the atom in state "g", we know that the field is indeed in a "cat state" superposition of two coherent states, $\psi_{\mathrm{field}} = \left(|z\mathrm{e}^{\mathrm{i}\phi}\rangle + |z\mathrm{e}^{-\mathrm{i}\phi}\rangle\right)/\sqrt{2}$. with distance

$$D = |z_1 - z_2| = 2\sqrt{\bar{n}}\sin\left(\frac{\Omega^2 t_i}{\delta}\right) \,, \tag{5.37}$$

where $\bar{n}$ is the mean initial photon number of the field state $|z\rangle$.

Now that such a superposition is created, it will decohere on a time scale $\tau_{\mathrm{dec}} = \gamma^{-1} D^{-2}$, i.e. the more rapid, the larger D. This has indeed been confirmed experimentally by sending a second atom through the cavity after a controllable time delay τ and going through a similar sequence of manipulations as for the first atom. Measuring the difference between conditional detection probabilities in the $|g\rangle$ and $|e\rangle$ states as a function of the time delay τ between the two atoms, it it possible to monitor the decay of the coherence between the two superposed field states and its dependence on the size of the superposition D.

More recently, using an ion in a trap (i.e a material oscillator), and the role of the decohering environment played by fluctuating classical fields, related experiments have been performed in D. Wineland's group in Boulder [9]. They were not only able to observe decoherence between coherent states of a damped harmonic oscillator as in Paris, but in a further part of their experiment they

managed to couple the ion's state in the trap to its energy, such that energy eigenstates rather than coherent states turn robust. In this way they were indeed able to confirm the quadratic dependence of the decoherence time scale on the "distance" D, which was either a phase space distance for a damped oscillator, or the energy difference for the coupling to the ion's energy.

5.2 Decoherence in Quantum Brownian Motion

The archetype of an open system is a Brownian particle suspended in some environment at temperature T. Classically, we can observe the erratic trajectory of such a Brownian particle. Quantum mechanically, such a classical trajectory of a particle ought to be some (semi-)classical limit of the dynamics of a localized wave packet. As we have seen, it is the type of interaction between system and environment that determines the nature of robust open system states. For the damped oscillator, indeed, localized wave packets, i.e. "classical" states were robust. Similar findings will be made here in the general framework of Quantum Brownian motion. For that we have to identify the robust states of a quantum Brownian particle, i.e. those system states that lead to little (or no) entanglement with the environment. As the interaction of a Brownian particle with its environment depends on position, it comes as no surprise that coherent superpositions of states with largely different positions will decohere rapidly. Localization in position space (a "classical" property) thus follows from robustness and a position dependent interaction between the particle and its environment. This is how decoherence may explain why macroscopic objects appear localized.

5.2.1 Classical Brownian Motion

Langevin vs. Fokker–Planck Equation. Before discussing the quantum version, let us briefly recall the classical theory of Brownian motion. Newton's equation of motion of a Brownian particle subjected to an external force $(-V'(q))$ and friction reads

$$m\ddot{q} + m\gamma\dot{q} + V'(q) = F(t) \tag{5.38}$$

with γ a friction or relaxation rate, and $F(t)$ a force due to thermal fluctuations with statistics

$$\begin{aligned} \mathcal{M}\Big[F(t)\Big] &= 0 \,, \\ \mathcal{M}\Big[F(t)F(s)\Big] &= 2m\gamma k_{\mathrm{B}} T\delta(t-s) \,. \end{aligned} \tag{5.39}$$

We use the notation $\mathcal{M}\big[\,\ldots\big]$ to denote such classical ensemble means in order to clearly distinguish those from quantum expectation values, denoted by angular brackets, $\langle\ldots\rangle$. We may transform this stochastic equation of motion (Langevin equation) to phase space equations. If we denote the Hamiltonian of the isolated

particle with $H = p^2/2m + V(q)$, we find the Brownian dynamics expressed in (q, p):

$$\dot{q} = \frac{\partial H}{\partial p} \,,$$
$$\dot{p} = -\frac{\partial H}{\partial q} - \gamma p + F(t) \,. \tag{5.40}$$

Instead of individual particles following stochastic trajectories, we may look at an ensemble of Brownian particles with phase space probability distribution: $\rho(q, p, t)$. Its time evolution is governed by the Fokker–Planck equation:

$$\partial_t \rho = \{H, \rho\} + \gamma \frac{\partial}{\partial p}(p\rho) + m\gamma k_{\mathrm{B}} T \frac{\partial^2}{\partial p^2}\rho \,. \tag{5.41}$$

Both descriptions, the stochastic trajectories and the ensemble evolution are equivalent in the sense that the latter may be recovered from the former as the ensemble mean $\rho(q, p, t) = \mathcal{M}\left[\delta(q - q_t)\delta(p - p_t)\right]$ over stochastic trajectories (q_t, p_t). The fluctuation-dissipation relation (see p. 48) ensures that the thermal state $\rho \simeq \mathrm{e}^{-H(q,p)/k_{\mathrm{B}}T}$ is a stationary solution of (5.41), as is easily verified. The quantization of such dissipative dynamics may be based on a microscopic model for system and environment, as developed next.

Microscopic Model. The classical dynamics of the Brownian particle may be derived from a closed model of "system" and "environment". Let the "system" coordinates and momenta be denoted by (q, p) and, accordingly, the "environment" canonical pairs be (q_λ, p_λ) with $\lambda = 1, 2, 3 \cdots, \infty$.
As the "system" Hamiltonian we choose $H(q, p) = p^2/2m + V(q)$. The environment is modeled by a collection of harmonic oscillators, coupled bilinearly to the system through their positions (see Chap. 1 for details). The full microscopic model thus reads

$$\begin{aligned} H_{\mathrm{tot}}(q, \rho, q_\lambda, \rho_\lambda) &= H(q, p) + \sum_\lambda \left\{ \frac{p_\lambda^2}{2m_\lambda} + \frac{1}{2} m_\lambda \omega_\lambda^2 \left(q_\lambda - \frac{g_\lambda}{m_\lambda \omega_\lambda^2} q \right)^2 \right\} \\ &= \underbrace{\tilde{H}(q, p)}_{H_{\mathrm{sys}}} + \underbrace{q \sum_\lambda g_\lambda q_\lambda}_{H_{\mathrm{int}}} + \underbrace{\sum_\lambda \frac{p_\lambda^2}{2m_\lambda} + \frac{1}{2} m_\lambda \omega_\lambda^2 q_\lambda^2}_{H_{\mathrm{env}}} \,. \end{aligned} \tag{5.42}$$

We have introduced an effective $\tilde{H}(q, p)$ to be able to write the total Hamiltonian in standard form. The analytical solution of the classical Hamiltonian equation of motion for the environmental degrees may be employed to express the system dynamics in the form

$$m\ddot{q}_s + \int_0^t \mathrm{d}s\, \kappa(t - s)\, \dot{q}_s + V'(q) = F(t) \quad + \text{transient term} \,. \tag{5.43}$$

Here,

$$\kappa(t-s) = \sum_\lambda \frac{g_\lambda^2}{m_\lambda \omega_\lambda^2} \cos\omega_\lambda (t-s) \tag{5.44}$$

denotes the classical damping kernel (cf. (1.127)) and the right hand-side of the equation reads

$$F(t) \equiv \sum_\lambda g_\lambda \left(q_\lambda(0)\cos\omega_\lambda t + \frac{p_\lambda(0)}{m_\lambda\omega_\lambda}\sin\omega_\lambda t \right) . \tag{5.45}$$

Assuming a thermal initial distribution $\simeq \mathrm{e}^{-H_\lambda(q,p)/k_\mathrm{B}T}$ of initial conditions $(q_\lambda(0), p_\lambda(0))$, $F(t)$ plays the role of a stochastic force with

$$\mathcal{M}\Big[F(t)\Big] = 0 \quad \text{and} \quad \mathcal{M}\Big[F(t)F(s)\Big] = k_\mathrm{B}T\kappa(t-s) , \tag{5.46}$$

see the discussion in Sect. 1.3.2. Apparently, we recover the desired Brownian Langevin equation from the microscopic model if the damping kernel is replaced by $\kappa(t-s) = 2m\gamma\delta(t-s)$ (ohmic damping). Having a model of system and environment at hand, it is now straightforward to quantize Brownian motion [10].

5.2.2 High Temperature Limit

For the quantized Brownian motion with H_tot from the last section we solve von Neumann's equation for the total density operator with an initial total state $\rho_\mathrm{tot}(0) = \rho_\mathrm{sys}(0) \otimes \rho_\mathrm{thermal}$. As shown in Chapt. 1 the propagator of the reduced density operator

$$\rho_\mathrm{red}(x,x',t) = \int \mathrm{d}x_0 \int \mathrm{d}x_0' \, J(x,x',t;x_0,x_0',0)\rho_\mathrm{red}(x_0,x_0',0) \tag{5.47}$$

can be written as double path integral

$$J = \int_{(x_0,0)}^{(x,t)} \mathcal{D}[q] \int_{(x_0',0)}^{(x',t)} \mathcal{D}[q'] \mathrm{e}^{\mathrm{i}(S[q]-S[q'])/\hbar} \mathcal{F}[q,q'] \tag{5.48}$$

with the Feynman-Vernon influence functional

$$\mathcal{F}[q,q'] = \exp\Big(-\frac{1}{\hbar^2} \int_0^t \mathrm{d}s \int_0^s \mathrm{d}\tau (q_s - q_s')\left(\alpha(s-\tau)q_\tau - \alpha^*(s-\tau)q_\tau'\right) \Big) . \tag{5.49}$$

Here the thermal quantum correlation function is

$$\begin{aligned} \alpha(t-s) &= \Big\langle \hat{F}(t)\hat{F}(s) \Big\rangle \\ &= \hbar \sum_\lambda \frac{g_\lambda^2}{2m_\lambda\omega_\lambda} \left[\coth\left(\frac{\hbar\omega_\lambda}{2k_\mathrm{B}T}\right)\cos\omega_\lambda(t-s) - \mathrm{i}\sin\omega_\lambda(t-s) \right] , \end{aligned} \tag{5.50}$$

see (1.135). In all generality, the non-Markovian nature of the dynamics, visible through the memory integrals in the influence functional, prevents us from deriving a time-local, simple master equation for the reduced density operator. We therefore concentrate on the high-temperature case.

Master Equation for $\rho(x, x', t)$. In the high temperature limit, $k_{\mathrm{B}}T \gg \hbar\Lambda$, where Λ is a high frequency cutoff for the environmental oscillators ($\omega_\lambda < \Lambda$), we may use $\coth(x) \sim 1/x + \mathcal{O}(x)$ for $x \to 0$ and get the high temperature quantum correlation function

$$\alpha(t-s) = k_{\mathrm{B}}T \sum_\lambda \frac{g_\lambda^2}{m_\lambda \omega_\lambda^2} \cos\omega_\lambda(t-s) - \frac{\mathrm{i}\hbar}{2} \sum_\lambda \frac{g_\lambda^2}{m_\lambda \omega_\lambda} \sin\omega_\lambda(t-s)\,. \quad (5.51)$$

Employing the damping kernel $\kappa(t-s)$ of the classical Brownian motion, the quantum correlation function may be expressed in the form

$$\alpha(t-s) = k_{\mathrm{B}}T\kappa(t-s) + \frac{\mathrm{i}\hbar}{2}\partial_t\kappa(t-s) \quad (5.52)$$

valid at high temperature, $k_{\mathrm{B}}T \gg \hbar\Lambda$. We had to choose $\kappa(t-s) = 2m\gamma\delta(t-s)$ (ohmic damping) in order to describe standard classical Brownian motion. In the quantum case that choice leads to the influence functional

$$\begin{aligned} F[q,q'] = \exp\Big(-\frac{m\gamma k_{\mathrm{B}}T}{\hbar^2}\int_0^t \big(q(s)-q'(s)\big)^2 \mathrm{d}s\Big) \\ \times \exp\Big(-\frac{\mathrm{i}m\gamma}{2\hbar}\int_0^t \big(q(s)-q'(s)\big)\big(\dot q(s)+\dot q'(s)\big)\,\mathrm{d}s\Big)\,. \end{aligned} \quad (5.53)$$

Now the overall action in the path integral for the density operator propagator is time-local and the corresponding evolution equation for $\rho(x, x', t)$ takes the form

$$\begin{aligned} \partial_t\rho(x,x',t) = \frac{1}{\mathrm{i}\hbar}\left(-\frac{\hbar^2}{2m}\Big(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial x'^2}\Big) + \Big(V(x)-V(x')\Big)\right)\rho(x,x',t) \\ -\frac{\gamma}{2}(x-x')\Big(\frac{\partial}{\partial x}-\frac{\partial}{\partial x'}\Big)\rho(x,x') - \frac{m\gamma k_{\mathrm{B}}T}{\hbar^2}(x-x')^2\rho(x,x',t)\,, \end{aligned} \quad (5.54)$$

the high-temperature master equation for Quantum Brownian motion.

Operator- and Wigner Representation. With the standard operator replacement ($-\mathrm{i}\hbar\partial_x \to \hat p$) and noting that $\lambda_{\mathrm{dB}}^2 = \hbar^2/mk_{\mathrm{B}}T$ is the thermal de Broglie wave length, the above result may be written independently of the representation as

$$\partial_t\rho = \frac{1}{\mathrm{i}\hbar}\Big[H,\rho\Big] - \frac{\mathrm{i}\gamma}{2\hbar}\Big[q,\{p,\rho\}\Big] - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\Big[q,[q,\rho]\Big]. \quad (5.55)$$

Alternatively, expressed in terms of the Wigner function (5.11) we find

$$\partial_t W = \{H, W\}_{\text{Moyal}} + \gamma \frac{\partial}{\partial p}(pW) + m\gamma k_{\text{B}} T \frac{\partial^2}{\partial p^2} W \tag{5.56}$$

Here, $\{H, W\}_{\text{Moyal}}$ is the Moyal bracket (Poisson-bracket plus higher order $\hbar$-terms):

$$\{H, W\}_{\text{Moyal}} = \{H, W\}_{\text{Poisson}} - \frac{\hbar^2}{24} \frac{\partial^3 W}{\partial q^3} + \frac{\hbar^4}{2^4 \cdot 5} \frac{\partial^5 V}{\partial q^5} \frac{\partial^5 W}{\partial p^5} + \cdots . \tag{5.57}$$

We see that the terms arising from the coupling to the environment in this quantum master equation are identical to those in the classical Fokker–Planck equation for a phase space density $\rho(q, p)$. In particular, the term $\big[q, \{p, \rho\}\big]$ describes damping, whereas the double commutator term $[q, [q, \rho]]$ describes diffusion in the classical case.

5.2.3 Decoherence

The crucial term for decoherence is the double commutator term. Let us, for simplicity, forget the remaining terms for a moment and write

$$\partial_t \rho = -\frac{\gamma}{\lambda_{\text{dB}}^2} [q, [q, \rho]]. \tag{5.58}$$

In position representation we get

$$\partial_t \rho(x, x', t) = -\gamma \frac{(x - x')^2}{\lambda_{\text{dB}}^2} \rho(x, x', t) \ , \tag{5.59}$$

and thus

$$\begin{aligned} \rho(x, x', t) &= \exp\left(-\gamma \frac{(x - x')^2}{\lambda_{\text{dB}}^2} t \right) \rho(x, x', 0) \\ &= \mathrm{e}^{-\gamma t D^2} \rho(x, x', 0) \ . \end{aligned} \tag{5.60}$$

Here

$$D = \frac{|x - x'|}{\lambda_{\text{dB}}} \tag{5.61}$$

is the distance in position expressed in terms of the quantum reference length $\lambda_{\text{dB}} = (\hbar^2 / m k_{\text{B}} T)^{1/2}$. For macroscopic values of $|x - x'|$, D may assume huge values due to the appearance of Planck's constant in λ_{dB}. As an example, at $T = 300\,\text{K}$ and with $m = 1\,\text{g}$, we find $\lambda_{\text{dB}} = 10^{-22}\,\text{m}$ and hence for a separation of $\delta x = 1\,\text{cm}$ the incredible $D = 10^{20}$. Remarkably, even without noticeable relaxation (take, for instance $\gamma = (\text{age of universe})^{-1} \approx 10^{-17}\text{sec}^{-1}$), decoherence has been effective after a time

$$\tau_{\text{dec}} = \frac{1}{\gamma D^2} = \frac{10^{17}}{10^{40}}\,\text{sec} = 10^{-23}\,\text{sec} \ . \tag{5.62}$$

After this very short amount of time, coherences between $|x\rangle$ and $|x + 1\,\mathrm{cm}\rangle$ are lost, yet no relaxation of quantities with a classical limit has taken place.

In a phase space representation the crucial double commutator term in the master equation is the "diffusion" term: $m\gamma k_\mathrm{B}T\frac{\partial^2}{\partial p^2}W$. Applied to a classical probability distribution, it describes diffusion, and thus no spectacular time scales occur. If applied to a quantum Wigner function, however, as may happen for the quantum Brownian motion master equation, it can have dramatic effects as we will explain next:

Consider the case of a superposition of two coherent states with distance δx, $(|-\delta x/2\rangle + |\delta x/2\rangle)/\sqrt{2}$. We saw earlier that the corresponding Wigner function consists of three terms

$$W(q,p) = \frac{1}{2}(W_- + W_+) + W_{+-} \tag{5.63}$$

with the coherences leading to the cross term

$$W_{+-}(q,p) = \frac{1}{\pi}W_s \cos\left(\frac{\delta x\, p}{\hbar}\right) , \tag{5.64}$$

which takes on both, positive and negative values. The effect of the "diffusion" operator on such a non-classical phase space function is

$$m\gamma kT\frac{\partial^2}{\partial p^2}W_{+-}(q,p) \approx -\gamma\left(\frac{\delta x}{\lambda_\mathrm{dB}}\right)^2 W_{+-}(q,p) = -\gamma D^2 W_{+-}(q,p) , \tag{5.65}$$

and therefore

$$W_{+-}(q,p,t) \sim \mathrm{e}^{-\gamma D^2 t}W_{+-}(q,p,0) . \tag{5.66}$$

Thus, decoherence manifests itself in the rapid destruction of oscillatory (negative) regions of the Wigner function on the fast decoherence time scale. After decoherence, the Wigner function resembles a classical phase space distribution, with phase space structures involving Planck's constant being leveled out. The quantum evolution may then be approximated by the classical Fokker–Planck evolution for a phase space density $\rho(q,p)$, see also [11].

In the discussion of this section we neglected completely the effect of the Hamiltonian and the damping term and only investigated the effect of the decohering, diffusive term. In the general Brownian motion case, there will be a competition between these influences. For large separations $|x - x'|$, however, decoherence will be the most important mechanism for short times.

As soon as the ensemble evolution may be approximated by the classical Fokker–Planck evolution, we know that we could equally well use an ensemble of classical trajectories, following the stochastic Langevin equation. We may therefore hope that starting from the quantized version of Brownian motion, after decoherence has been effective, the reduced density operator may be expressed as an ensemble of robust wave packets $\psi_i(t)$ following the classical Langevin trajectories, $\rho_t \approx \sum_i p_i|\psi_i(t)\rangle\langle\psi_i(t)|$, with time independent probabilities p_i. It is in fact possible to construct such an ensemble $\psi_i(t)$, employing a stochastic Schrödinger equation, as we will explain soon. Before doing so, however, we want to state general criteria how to identify robust states.

5.3 Robust States

Robust states of an open system are those system states leading to "little" (or ideally no) entanglement with the environment. In this section this statement is formulated in a more quantitative way, enabling us to find a criterion to identify robust states. Here, we follow closely the ideas developed in [12,13]. One fairly obvious approach is based on *entropy*: robust states remain (approximately) pure under time evolution of system and environment and thus entropy production in the open system is small (this entropy criterion to identify robust states has been termed "predictability sieve" by Zurek [5]).

Motivated by an earlier example, we here follow a somewhat different line of reasoning, eventually leading to similar results. We saw above in Sect. 5.1.2 that for the case of a damped cavity field mode with master equation

$$\partial_t \rho = -\mathrm{i}[\Omega a^\dagger a, \rho] + \frac{\gamma}{2}\Big([a, \rho a^\dagger] + [a\rho, a^\dagger]\Big) \,, \tag{5.67}$$

we were able to find pure coherent state solutions

$$\rho_t = |z(t)\rangle\langle z(t)| \qquad \text{with} \qquad z(t) = \mathrm{e}^{-(i\Omega + \frac{\gamma}{2})t} z(0) \,. \tag{5.68}$$

through an educated guess. These states remain pure and thus no entropy is produced in the open system – these states are robust. We were also able to identify an evolution equation for those robust states. Time derivation leads to a *nonlinear* evolution equation for $|\psi_t\rangle = |z(t)\rangle$:

$$\partial_t |\psi_t\rangle = -i\Omega a^\dagger a |\psi_t\rangle - \frac{\gamma}{2}\Big(a^\dagger a - \langle\psi_t| a^\dagger a |\psi_t\rangle\Big)|\psi_t\rangle. \tag{5.69}$$

While the total product state of system *and* environment, $|\Psi_t\rangle = |\psi_t\rangle|\beta_t\rangle$ is a solution of the linear Schrödinger equation, (5.69) is a nonlinear, closed evolution equation for the system part $|\psi_t\rangle$ of this product state.

Clearly, we expect an evolution equation for robust states to be nonlinear: as we saw earlier, almost by definition, a superposition of different robust states is no longer robust. In general, an open system will not allow exact robust states as in the simple example of a damped light mode above, where coherent states remain pure for all times. We expect, however, to find approximately robust states, i.e. states that remain close to being pure under time evolution. The next section helps to find them for a large class of open systems.

5.3.1 Robustness in Terms of Hilbert–Schmidt Norm

Imagine we have an evolution equation $\partial_t \rho = \mathcal{L}[\rho_t]$ for the reduced density operator of the open system. Here we show how to identify robust states and, closely related, how we can find an evolution equation for robust states.

Suppose that at some time t the state of the open system is given by a robust pure state $P_t = |\psi_t\rangle\langle\psi_t|$. After a short time δt, the dynamics of the open system will in general move such a pure system state away from the set of pure states

to $\rho_{t+\delta t} = P_t + \mathcal{L}[P_t]\delta t$. We try to find the pure state $P_{t+\delta t} = P_t + \delta P_t$ that approximates the true mixed state $\rho_{t+\delta t}$ as closely as possible. If it happened that $\delta P_t = \mathcal{L}[P_t]\delta t$, the pure P_t is obviously robust.

We use the Hilbert–Schmidt norm to measure the deviation

$$|\delta P_t - \mathcal{L}[P_t]\delta t|^2 \equiv \mathrm{tr}\Big(\Big(\delta P_t - \mathcal{L}[P_t]\delta t\Big)^\dagger \Big(\delta P_t - \mathcal{L}[P_t]\delta t\Big)\Big) \tag{5.70}$$

between a propagated pure state $P_{t+\delta t} = P_t + \delta P_t$ and the true mixed $\rho_{t+\delta t} = P_t + \mathcal{L}[P_t]\delta t$. The requirement that this distance should be minimal defines the new pure state $P_{t+\delta t}$, and thus an evolution equation for robust states ψ_t. Clearly, if P_t is exactly robust, this construction identifies pure state solutions of the master equation.

Let us find δP_t from the condition of minimal deviation in the sense of the Hilbert–Schmidt norm. Since $P_{t+\delta t} = P_t + \delta P_t$ is pure, i.e.

$$\begin{aligned} &\text{(i)} \quad P_{t+\delta t} = P_{t+\delta t}^\dagger \,, \\ &\text{(ii)} \quad P_{t+\delta t}^2 \;= P_{t+\delta t} \,, \end{aligned} \tag{5.71}$$

we see that

$$\begin{aligned} &\text{(i')} \qquad\qquad \delta P_t = \delta P_t^\dagger \,, \\ &\text{(ii')} \quad P_t\,\delta P_t + \delta P_t P_t = \delta P_t \,. \end{aligned} \tag{5.72}$$

From these conditions one easily concludes that there are Hermitian operators $\tilde{X}, \tilde{Y}$, such that $\delta P_t = \big[P_t, [P_t, \tilde{X}]\big]\delta t - i[\tilde{Y}, P_t]\delta t$. With the Hermitian $X = \tilde{X} - \mathrm{i}[\tilde{Y}, P_t]$ we find that there must be a Hermitian operator X such that the time evolution of the pure state projectors P_t may be written in the form

$$\delta P_t = \Big[P_t, [P_t, X]\Big]\delta t. \tag{5.73}$$

It is easily verified that with this expression, indeed, $\delta P = \delta P^\dagger$ and

$$\begin{aligned} P_t\,\delta P_t + \delta P_t\, P_t &= \Big(P_t X + P_t X P_t - 2P_t X P_t + P_t X P_t + X P_t - 2P_t X P_t\Big)\delta t \\ &= \Big(P_t X + X P_t - 2P_t X P_t\Big)\delta t \\ &= \delta P_t \,, \end{aligned} \tag{5.74}$$

so that evolution according to (5.73) indeed satisfies (5.72). What remains is the determination of the operator X in (5.73). For brevity, we write $\mathcal{L}[P_t] = Z$ for the generator of the master equation evolution, drop the time argument for the same reason and try to find that operator X that corresponds to the minimum of $\mathrm{tr}\big((Z - \big[P, [P, X]\big])^2\big)$. Simple algebra shows that

$$\mathrm{tr}\Big(\Big(Z-\Big[P,[P,X]\Big]\Big)^2\Big) = \mathrm{tr}\Big(Z^2 - 2Z\,(PX - 2PXP + PX) + \underbrace{(PX - 2PXP + XP)\,(PX - 2PXP + XP)}_{2PX^2 - 2PXPX}\Big)$$

$$= \mathrm{tr}\Big(\underbrace{Z^2 - 2\left(Z^2P - (ZP)^2\right)}_{\text{independent of } X} + 2\underbrace{\left((Z-X)^2P - [(Z-X)P]^2\right)}_{\langle((Z-X)-\langle Z-X\rangle)^2\rangle}\Big)\,. \tag{5.75}$$

We see that the norm is minimal for those operators X that satisfy $(Z-X)P_t = (\mathcal{L}[P_t] - X)P_t = \mathrm{const}\cdot P_t$. Therefore

$$\delta P_t = \Big[P_t, \Big[P_t, \mathcal{L}[P_t]\Big]\Big]\delta t, \tag{5.76}$$

or

$$\partial_t P_t = P_t\mathcal{L} + \mathcal{L}P_t - 2P_t\mathcal{L}P_t = \Big(\mathcal{L} - P_t\mathcal{L}\Big)P_t + P_t\Big(\mathcal{L} - \mathcal{L}P_t\Big)\,. \tag{5.77}$$

In terms of the pure state $|\psi_t\rangle$ with $P_t = |\psi_t\rangle\langle\psi_t|$ and with $\mathcal{L} \equiv \mathcal{L}[|\psi_t\rangle\langle\psi_t|]$ we find the evolution equation for Hilbert–Schmidt–robust states:

$$\partial_t|\psi_t\rangle = \Big(\mathcal{L} - \langle\psi_t|\mathcal{L}|\psi_t\rangle\Big)|\psi_t\rangle. \tag{5.78}$$

As expected, this is a non-linear evolution equation for robust states ψ_t.

5.3.2 Example: Lindblad Master Equation

The general Lindblad master equation is the most relevant example of this class of evolution equations:

$$\begin{aligned}\partial_t\rho &= -\frac{\mathrm{i}}{\hbar}[H,\rho] + \frac{1}{2}\Big([L\rho, L^\dagger] + [L, \rho L^\dagger]\Big)\\ &= -\frac{\mathrm{i}}{\hbar}\Big(H\rho - \rho H\Big) + L\rho L^\dagger - \frac{1}{2}L^\dagger L\rho - \frac{1}{2}\rho L^\dagger L \equiv \mathcal{L}[\rho].\end{aligned} \tag{5.79}$$

The corresponding evolution equation for Hilbert–Schmidt–robust states reads

$$\begin{aligned}\partial_t|\psi\rangle &= \Big(\Big\{ -\frac{\mathrm{i}}{\hbar}H|\psi\rangle\langle\psi| + \frac{\mathrm{i}}{\hbar}|\psi\rangle\langle\psi|H + L|\psi\rangle\langle\psi|\langle L^\dagger - \frac{1}{2}L^\dagger L|\psi\rangle\langle\psi|\\ &\quad -\frac{1}{2}|\psi\rangle\langle\psi|L^\dagger L\Big\} - \langle\psi|\{\}|\psi\rangle\Big)|\psi\rangle\\ &= \Big(-\frac{\mathrm{i}}{\hbar}H + \frac{\mathrm{i}}{\hbar}\langle H\rangle + \langle L^\dagger\rangle L - \frac{1}{2}L^\dagger L - \frac{1}{2}\langle L^\dagger L\rangle - \langle L^\dagger\rangle\langle L\rangle\\ &\quad +\frac{1}{2}\langle L^\dagger L\rangle + \frac{1}{2}\langle L^\dagger L\rangle\Big)|\psi\rangle.\end{aligned} \tag{5.80}$$

Dropping the irrelevant phase arising from the expectation value of the Hamiltonian, (5.80) may be simplified to

$$\partial_t|\psi_t\rangle = -\frac{\mathrm{i}}{\hbar}H|\psi_t\rangle - \frac{1}{2}\Big(L^\dagger L - \langle L^\dagger L\rangle - 2\langle L^\dagger\rangle\Big(L - \langle L\rangle\Big)\Big)|\psi_t\rangle. \tag{5.81}$$

This nonlinear evolution equation turns out to be of central importance for the propagation of robust states. We turn to two simple examples.

Damped Electromagnetic Field Mode. Our first example is the evolution equation of the damped harmonic oscillator of Lindblad type with $L = \sqrt{\gamma}a$ and $H = \hbar\Omega a^\dagger a$. For this case, the general evolution equation (5.81) for robust states takes the form

$$\partial_t|\psi_t\rangle = -\mathrm{i}\Omega a^\dagger a|\psi_t\rangle - \frac{\gamma}{2}\Big(a^\dagger a - \langle a^\dagger a\rangle - 2\langle a^\dagger\rangle\underbrace{\big(a - \langle a\rangle\big)}_{=0 \text{ for coherent states}}\Big)|\psi_t\rangle. \tag{5.82}$$

We see that (5.82) preserves coherent states for which the last term in the bracket is equal to zero. Nicely, we confirm the nonlinear evolution equation (5.23) for the coherent state solutions of the master equation mentioned earlier. Thus, the general Hilbert–Schmidt criterion for robust state dynamics recovers the coherent states $|z(t)\rangle$ we found earlier by an educated guess.

5.3.3 Quantum Brownian Motion (Simplified)

In a simplified version of Brownian motion one sometimes neglects the damping term and only retains diffusion. Then, the high temperature quantum Brownian motion master equation takes the Lindblad form

$$\partial_t\rho = -\frac{\mathrm{i}}{\hbar}[H,\rho] - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\Big[q,[q,\rho]\Big]. \tag{5.83}$$

The corresponding evolution equation (5.81) for robust states reads

$$\partial_t\psi_t = -\frac{\mathrm{i}}{\hbar}H\psi_t - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\Big(\,(q-\langle q\rangle)^2 - \Big\langle(q-\langle q\rangle)^2\Big\rangle\,\Big)\psi_t\,. \tag{5.84}$$

Without system Hamiltonian H, robust states would clearly be position eigenstates $|q\rangle$. As we will see soon, the nonlinear term tends to localize an arbitrary wave packet to such position eigenstates.

Taking into account the system dynamics through the Hamiltonian H, however, leads to a competition between unitary evolution (spreading of wave packets) and this localization, resulting in robust wave packets with a specific width.

Consider, for simplicity, the case of a free Brownian particle with $H = p^2/2m$. The pure state evolution optimally mimicking the true density operator evolution according to our Hilbert–Schmidt criterion (5.81) is

$$\partial_t\psi_t = -\frac{\mathrm{i}}{\hbar}\frac{p^2}{2m}\psi_t - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\Big[(q-\langle q\rangle)^2 - \Big\langle(q-\langle q\rangle)^2\Big\rangle\Big]\psi_t. \tag{5.85}$$

As in the case of the damped harmonic oscillator we try to find stable Gaussian wave packet solutions of this nonlinear evolution equation. In position representation, the evolution of Hilbert–Schmidt robust states for Quantum Brownian motion reads

$$\partial_t \psi(x,t) = \frac{1}{\mathrm{i}\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\right)\psi(x,t) - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\left[\Big(x-\langle x\rangle\Big)^2 - \left\langle\Big(x-\langle x\rangle\Big)^2\right\rangle\right]\psi(x,t)\,. \tag{5.86}$$

We can already read off some general properties of this equation. First, it is nonlinear and norm preserving, $\int \mathrm{d}x|\psi(x,t)|^2 = 1$. Next, with $\sigma = \sqrt{\langle (x-\langle x\rangle)^2\rangle}$ the spread of the wave packet, we see that for positions $x \in \big(\langle x\rangle - \sigma; \langle x\rangle + \sigma\big)$ near the expectation value, $(x-\langle x\rangle)^2 - \sigma^2 < 0$ and therefore $\psi(x)$ increases for these x. In contrast, if $x \notin \big(\langle x\rangle - \sigma; \langle x\rangle + \sigma\big)$, i.e. for those positions x far away from $\langle x\rangle$, $(x-\langle x\rangle)^2 - \sigma^2 > 0$, and $\psi(x)$ decreases. Thus, the time evolution due to the nonlinear term tends to localize wave packets in position space. The kinetic energy, on the other hand, tends to increase the spread of the wave packet. There are indeed soliton-like Gaussian solutions that are effected equally by the localizing influence of the nonlinear and the dispersing influence of the kinetic term, such that they preserve their shape for all times. We make the ansatz

$$\psi(x,t) = \left(\frac{C_R}{2\hbar}\right)^{1/4} \exp\left(-\frac{C(x-\bar{x}_t)^2}{4} + \frac{\mathrm{i}\bar{p}(x-\bar{x}_t)}{\hbar} + i\phi_t\right) \tag{5.87}$$

with $\bar{x}_t = \bar{x}_0 + \bar{p}t/m$. It turns out that there is a single value $C = C_R + \mathrm{i}C_I = (1-\mathrm{i})\big(4m\gamma/\hbar\lambda_{\mathrm{dB}}^2\big)^{1/2}$ for which the Gaussian ansatz satisfies equation (5.86). The corresponding robust width is

$$\begin{aligned}\sigma_0^2 &\equiv \big\langle (x-\langle x\rangle)^2\big\rangle = \frac{1}{C_R} = \sqrt{\frac{\hbar\lambda_{\mathrm{dB}}^2}{4m\gamma}} = \lambda_{\mathrm{dB}}^2\sqrt{\frac{k_{\mathrm{B}}T}{4\hbar\gamma}} \\ \sigma_0 &= \lambda_{\mathrm{dB}}\left(\frac{k_{\mathrm{B}}T}{4\hbar\gamma}\right)^{1/4}\,.\end{aligned} \tag{5.88}$$

We may interprete this σ_0 as the scale beyond which no coherences will be maintained under time evolution - i.e. σ_0 is nothing but the coherence length of the Brownian particle.

As a concrete example, consider again a mass $m = 1\,\mathrm{g}$, at $T = 300\,\mathrm{K}$, and $\gamma = (\text{age of universe})^{-1}$. We find $\sigma \approx 10^{-15}\,\mathrm{m}$. These pure Gaussian states of just this spread σ_0 approximate the true time evolution of the master equation (5.55) well and thus lead to little entanglement between system and environment. Entropy production is minimal. This may be confirmed by a direct entropy argument.

5.3.4 Robustness Based on Entropy

There is an obvious approach to robustness based on entropy. This being the approach most often employed, we devote a section to some results that may be

found in the literature [1,2,5]. For a given reduced density operator evolution $\dot{\rho} = \mathcal{L}[\rho_t]$, robust states are those that remain as pure as possible under time evolution. A measure for purity is entropy. In the decoherence literature, one often encounters a simplified measure for purity, sometimes also called "linear entropy":

$$S(t) = \mathrm{tr}\left(\rho_t - \rho_t^2\right) \ . \tag{5.89}$$

As long as ρ_t remains pure we have $\rho = \rho^2$, and therefore $S(t) = 0 = \mathrm{const}$ for all times. This pure state describes maximal information about the system. If, however, ρ_t becomes mixed, $S(t)$ increases, corresponding to a loss of information.

To identify the robust states, consider all possible pure initial states for the system and compute the linear entropy $S(t)$ associated with the reduced density matrix after some time t. The robust states are the ones that become only very little entangled with the environment and thus minimize entropy production.

As an example we consider once again the simplified version of quantum Brownian motion (i.e. neglecting friction) with

$$\partial_t \rho = \frac{1}{\mathrm{i}\hbar}\Big[H, \rho\Big] - \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\Big[q, [q, \rho]\Big]. \tag{5.90}$$

The initial linear entropy production is $\partial_t S(0) = -2\,\mathrm{tr}(\rho\partial_t\rho) = 4\gamma/\lambda_{\mathrm{dB}}^2 \times \left\langle (q - \langle q\rangle)^2 \right\rangle$ for a pure initial state. Obviously, the smaller the spread in position, the less the state of the system will be affected by the coupling to the environment. However, this reasoning applies for very early times only, since the effect of the system Hamiltonian H has been entirely neglected. For a harmonic oscillator system, for instance, the role of q and p are exchanged after a quarter of the period, which is why in the literature time-averaged entropy production is considered [14].

The linear entropy production is given by

$$\begin{aligned}
\partial_t S(t) &= -2\,\mathrm{tr}(\rho_t \partial_t \rho_t) = \frac{4\gamma}{\lambda_{\mathrm{dB}}^2}\left(\rho^2 x^2 - \rho x \rho x\right) \\
&\approx \frac{4\gamma}{\lambda_{\mathrm{dB}}^2}\left\langle (x - \langle x\rangle)^2 \right\rangle_t \\
&= \frac{4\gamma}{\lambda_{\mathrm{dB}}^2}\,(\Delta x)_t^2\,.
\end{aligned} \tag{5.91}$$

The crucial approximation here is that ρ is still considered approximately pure which requires very weak coupling.

Consider a harmonic oscillator with $H = p^2/2m + m\omega^2 q^2/2$. We integrate entropy production over one period of the oscillation, where the time evolution is based on H_{sys} only, again an approximation based on weak coupling. The entropy produced over a whole period of the oscillator becomes

$$S(t) \approx \frac{4\gamma}{\lambda_{\rm dB}^2} \int\limits_0^T {\rm d}t \left\langle \left((x - \langle x \rangle) \cos\omega t + \frac{(p - \langle p \rangle)}{m\omega} \sin\omega t \right)^2 \right\rangle$$
$$= \frac{2\gamma}{\lambda_{\rm dB}^2} \left((\Delta x)^2 + \frac{(\Delta p)^2}{m^2\omega^2} \right)$$
$$= \frac{4\gamma}{m\omega^2\lambda_{\rm dB}^2} \left(\frac{1}{2} m\omega^2 (\Delta x)^2 + \frac{\Delta p^2}{2m} \right) . \qquad (5.92)$$

This expression should be minimized. Employing Heisenberg's uncertainty relation $\Delta x \Delta p \geq \hbar/2$, we find that the time averaged entropy production is minimal for coherent states with $\Delta x \Delta p = \hbar/2$ and $(\Delta x)^2 = \hbar/2m\omega$. Thus, for a Brownian particle bound harmonically, the robust states are indeed the minimum uncertainty coherent states, here derived from the entropy argument.

A free Brownian particle with $H = p^2/2m$ is another common example [1]. In order to simplify equations, we introduce a length scale $\ell = (\hbar\lambda_{\rm dB}^2/m\gamma)^{1/4} = \lambda_{\rm dB}(k_{\rm B}T/\hbar\gamma)^{1/4} = \sqrt{2}\ell_{\rm HS}$ and a time scale $t_0 = (m\lambda_{\rm dB}^2/\hbar\gamma)^{1/2}$. Here, $\ell_{\rm HS}$ denotes the length obtained from the robustness criterion based on the Hilbert–Schmidt norm argument.

The quantum Brownian master equation (again without friction), expressed in these dimensionless units $x = q/\ell, \quad \tau = t/t_0$ reads

$$\partial_\tau \rho(x, x', \tau) = \frac{\rm i}{2} \left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial x'^2} \right) \rho(x, x', \tau) - (x - x')^2 \rho(x, x', \tau) . \qquad (5.93)$$

The propagator is easily found [1]: $G(x, x', t', x_0, x_0', 0) =$

$$\frac{1}{2\pi\tau} \exp\left(\frac{\rm i}{2\tau} \Big[(x - x_0)^2 - (x' - x_0')^2 \Big] - \frac{\tau}{3} \Big[(x - x')^2 + (x_0 - x_0')^2 + (x - x')(x_0 - x_0') \Big] \right) . \qquad (5.94)$$

For an initial pure Gaussian,

$$\rho_0(x, x') = \frac{1}{\sqrt{\pi b^2}} \, {\rm e}^{-\frac{x^2}{2b^2}} \cdot {\rm e}^{-\frac{x'^2}{2b^2}} \qquad (5.95)$$

with width b, we can evaluate the Gaussian integrals and find the density operator explicitely. We are thus in the position to determine the linear entropy: $S = {\rm tr}(\rho - \rho^2)$ and find [1]

$$S(b, \tau) = 1 - \left(\frac{3b^2}{4b^2\tau^4 + 2\tau^3 + 24b^4\tau + 3b^2} \right)^{1/2} . \qquad (5.96)$$

We see that $S(b, \tau)$ increases from zero to one for scaled times $\tau \gg 1$. For a "typical" time, $\tau \approx 1$, it easily shown that $S(b, \tau)$ assumes its smallest value for a scaled size b of the order of one (recall that the length scale was essentially the coherence length $\ell_{\rm HS}$ of the Hilbert–Schmidt-robust states). Thus, we do indeed confirm the predicted minimum (averaged) entropy production for those wave packets whose width is of the order of $\ell_{\rm HS} = \lambda_{\rm dB}(k_{\rm B}T/4\hbar\gamma)^{1/4}$.

Thus, robust states determined through linear entropy production confirm the findings based on the Hilbert–Schmidt-robustness criterion. However, we had to rely on very simple, soluble models to be able to determine the robust states. How can we find them in more complicated situations? In particular, in the Brownian motion case we always expect robust states to be well localized wave packets when we approach the macroscopic $\hbar \to 0$ limit, irrespectively of any particular system Hamiltonian H. If we base our investigation on the concept of Hilbert–Schmidt-robustness we can let the computers do the work and determine robust states numerically, as will be explained next.

5.3.5 Stochastic Schrödinger Equations and Robust States

Again, for concreteness, we consider an open system whose dynamics is described by an evolution equation of Lindblad type:

$$\partial_t \rho = -\frac{\mathrm{i}}{\hbar}[H,\rho] + \frac{1}{2}\Big([L\rho, L^\dagger] + [L, \rho L^\dagger]\Big)\,. \tag{5.97}$$

The Hilbert–Schmidt-robustness criterion led to an evolution equation for pure states mimicking the exact evolution:

$$\partial_t|\psi_t\rangle = -\frac{\mathrm{i}}{\hbar}H|\psi_t\rangle - \frac{1}{2}\Big(L^\dagger L - \langle L^\dagger L\rangle - 2\langle L^\dagger\rangle\Big(L - \langle L\rangle\Big)\Big)|\psi_t\rangle. \tag{5.98}$$

Next we add some multiplicative noise (Stratonovich) to this evolution equation and consider the stochastic Schrödinger *quantum state diffusion* equation [12,15]

$$\begin{aligned}\partial_t|\psi\rangle = &-\frac{\mathrm{i}}{\hbar}H|\psi_t\rangle - \frac{1}{2}\Big(L^\dagger L - \langle L^\dagger L\rangle - 2\langle L^\dagger\rangle\Big(L - \langle L\rangle\Big)\Big)|\psi_t\rangle \\ &+\Big(L - \langle L\rangle\Big)z_t|\psi_t\rangle\end{aligned} \tag{5.99}$$

with a white, complex stochastic process z_t, whose correlations are given by

$$\begin{aligned}\mathcal{M}\,[z_t] \quad &= 0\,, \\ \mathcal{M}\,[z_t z_s] \;&= 0\,, \\ \mathcal{M}\,[z_t^* z_s] &= \delta(t-s)\,.\end{aligned} \tag{5.100}$$

Here we use again $\mathcal{M}[\ldots]$ to denote the mean over this stochastic process in order to clearly distinguish such an ensemble mean from quantum expectation values denoted by angular brackets, $\langle\ldots\rangle$. The introduction of the noise appears somewhat ad hoc at this stage but will be motivated below. Using stochastic calculus, as explained in the next section, it may be shown that for an initial pure state $\rho_0 = |\psi_0\rangle\langle\psi_0|$ the stochastic terms lead to an astonishing identity [15],

$$\rho_t = \mathcal{M}\Big[|\psi_t\rangle\langle\psi_t|\Big], \tag{5.101}$$

which is valid for all times t, with ψ_t the individual solutions of the quantum state diffusion equation (5.99). Thus the reduced density operator of the system is expressed without any approximation as an ensemble of pure states that may be determined (numerically) for an arbitrary system Hamiltonian H and coupling operator L to the environment. We thus succeded in writing the reduced density operator as an ensemble of states ψ_t (with time independent probabilities) that will turn out to be robust states for times $t \gg t_{\rm dec}$, i.e. after decoherence took place and the fluctuations became small. We see that the stochastic Schrödinger equation enables us to express ρ_t dynamically as a mixture of robust states.

Let us discuss this mechanism in more detail for the case of quantum Brownian motion neglecting friction, i.e. with $L = \sqrt{\gamma} q/\lambda_{\rm dB}$ and a general Hamiltonian $H = p^2/2m + V(q,t)$. The quantum state diffusion stochastic Schrödinger equation reads for this case

$$\partial_t|\psi\rangle = -\frac{\mathrm{i}}{\hbar}H|\psi\rangle - \frac{\gamma}{\lambda_{\rm dB}^2}\left(\Big(q-\langle q\rangle\Big)^2 - \Big\langle (q-\langle q\rangle)^2\Big\rangle\right)|\psi\rangle + \frac{\sqrt{2\gamma}}{\lambda_{\rm dB}}\,(q-\langle q\rangle)\;\; z_t|\psi\rangle\;. \tag{5.102}$$

A few remarks about this equation are in order. First, we see that the states remain normalized, $\langle\psi_t|\psi_t\rangle = 1$, which is obviously true for the general equation (5.99). Next, as we saw in the last Section, we know that the deterministic term arising from the environmental influence tends to localize ψ_t in position space, balanced by the spreading influence of the kinetic term in the Hamiltonian. The new fluctuating term is relevant as long as the state ψ_t extends over a large region and leads to localization at a stochastic position $\langle q\rangle$ in each run. Once the wave packet $\psi(x,t)$ has localized to its robust width $\sigma_{\rm HS}$, i.e. after the decoherence time $t \gg t_{\rm dec}$, the fluctuations become negligible and the wave packet follows closely a classical orbit with localizing influence of the additional deterministic term and with tiny fluctuations.

The stochastic Schrödinger equation (5.102) describes both: the early decoherence towards robust states for $t \leq t_{\rm dec}$ and the following near-classical motion of robust states as given by the nonlinear evolution equation (5.81) from the Hilbert–Schmidt criterion for robustness.

5.3.6 Some Remarks about Stochastic Schrödinger Equations

Stochastic Schrödinger equations were introduced in the beginning of the nineties independently in various communities, most notably in quantum optics and continuous quantum measurement circles, see [16,17]. There are several variants, some involving white noise such as the quantum state diffusion equation mentioned above, some driven by discrete, Poisson noise, modeling jump processes as in photon-counting. Depending on the circumstances and authors who introduced them, several names for stochastic Schrödinger equations are used. Very often a solution of a stochastic Schrödinger equation is termed "quantum trajectory" in reference to the stochastic classical Langevin equation. Also, the

terms "wave function Monte-Carlo" or "quantum filtering equation" have been used or even the simple "quantum stochastic differential equation". In quantum optics, diffusive type stochastic equations similar to the quantum state diffusion equation employed here are related to so-called heterodyne (and also homodyne) detection. We restrict ourselves to the continuous, diffusive equation (5.99) driven by white noise.

Itô vs Stratonovich Calculus. In stochastic calculus it is important to state which convention (Stratonovich or Itô) is used. The Stratonovich equation (5.99), which more formally should be written as

$$\mathrm{d}\psi = -\frac{\mathrm{i}}{\hbar}H\psi\mathrm{d}t - \frac{1}{2}\Big(L^\dagger L - \langle L^\dagger L\rangle - 2\langle L^\dagger\rangle\left(L - \langle L\rangle\right)\Big)\psi\mathrm{d}t + \left(L - \langle L\rangle\right)\psi \circ \mathrm{d}\xi, \tag{5.103}$$

where the circle $\circ$ denotes Stratonovich calculus, may be translated to its Itô-version through the rule

$$A \circ \mathrm{d}\xi \;=\; A\mathrm{d}\xi + \frac{1}{2}\mathrm{d}A\mathrm{d}\xi. \tag{5.104}$$

The Itô version of the quantum state diffusion equation (5.99) reads

$$\mathrm{d}\psi = -\mathrm{i}H\psi\mathrm{d}t - \frac{1}{2}\left(L^\dagger L - 2\langle L^\dagger\rangle L + |\langle L\rangle|^2\right)\psi\mathrm{d}t + \left(L - \langle L\rangle\right)\psi\mathrm{d}\xi\ , \tag{5.105}$$

with

$$\mathrm{d}\xi\mathrm{d}\xi^* = \mathrm{d}t\ , \quad \mathrm{d}\xi^2 = 0\ .$$

Using Itô stochastic calculus is somewhat easier since the stochastic increments are forward directed and thus independent of the current state. Mean values are easily evaluated. Care must be taken to always expand to second order in the infinitesimals to ensure the correct order in $\mathrm{d}t$. For instance, it is easy to see that $\rho_t = \mathcal{M}[|\psi\rangle\langle\psi|]$ does indeed satisfy Lindblad's master equation through the little calculation $\mathrm{d}|\psi\rangle\langle\psi| = |\psi\rangle\langle\mathrm{d}\psi| + |\mathrm{d}\psi\rangle\langle\psi| + |\mathrm{d}\psi\rangle\langle\mathrm{d}\psi|$ with $\mathrm{d}\psi$ from (5.105) and ensemble averaging.

Sketch of a Derivation. We saw the nice properties of the stochastic Schrödinger equation. Here we want to sketch briefly some more of the underlying physics that helps to understand why the stochastic Schrödinger equation (5.99) gives robust states in open system dynamics. One way to achieve that goal is to derive it from the standard microscopic model of system and environment of harmonic oscillators,

$$H_{\mathrm{tot}} = H + \sum_\lambda g_\lambda\Big(La_\lambda^\dagger + L^\dagger a_\lambda\Big) + \sum_\lambda \omega_\lambda a_\lambda^\dagger a_\lambda\ . \tag{5.107}$$

with an initial total state $\Psi(0) = \underbrace{|\psi_0\rangle}_{\mathrm{sys}}\ \underbrace{|0\rangle|0\rangle\cdots}_{\mathrm{env}}$.

Assuming a delta-correlated bath response function this model implies a Lindblad equation for the reduced density operator evolution:

$$\partial_t \rho = -\frac{\mathrm{i}}{\hbar}[H\rho] + \frac{1}{2}\Big([L\rho, L^\dagger] + [L, \rho L^\dagger]\Big) \tag{5.108}$$

with $\rho_0 = |\psi_0\rangle\langle\psi_0|$.

In our very first example of a damped harmonic oscillator the crucial step towards robust states was the identification of total system product solutions $|\Psi_t\rangle = |\psi_t\rangle\,|z(t)\rangle$ of the Schrödinger equation, with both the system part $|\psi_t\rangle$ and the environmental part $|z(t)\rangle$ being coherent states. In general, however, no such exact product solutions will exist – what we are after are product states $|\psi_t\rangle\,|z(t)\rangle$ that solve the total Schrödinger equation approximately.

The first crucial requirement towards this goal is to keep the environmental part moving coherent states $|z(t)\rangle$, since, as before, the environment is a collection of harmonic oscillators. We saw that coherent states are robust for harmonic oscillators and we expect them to lead to the least entanglement. In order to identify the corresponding system state ψ_t, our strategy is as follows:

We start from the Schrödinger equation $\mathrm{i}\hbar\partial_t\Psi_t = H_{\mathrm{tot}}\Psi_t$ for system and environment and expand the total state in a (fixed) (Bargmann) coherent state basis with $\|z\rangle = \mathrm{e}^{za^\dagger}|0\rangle$ (these are unnormalized coherent states). The resolution of the identity reads

$$\mathbb{1} = \int \frac{\mathrm{d}^2 z}{\pi}\mathrm{e}^{-|z|^2}\ \|z\rangle\langle z\| \tag{5.109}$$

so that

$$|\Psi_t\rangle = \int \frac{\mathrm{d}^2 z}{\pi}\mathrm{e}^{-|z|^2}\underbrace{|\psi_t(z^*)\rangle}_{\mathrm{sys}}\underbrace{\|z\rangle}_{\mathrm{env}} \tag{5.110}$$

where $|\psi_t(z^*)\rangle = \langle z\|\ \Psi_t\rangle$. Be aware that the coherent state label z should be understood as the vector $z = (z_1, z_2, \ldots, z_\lambda, \ldots)$ consisting of the labels of the coherent states of each environmental oscillator. Note that in this expansion $|\psi_t(z^*)\rangle$ also carries the information about the amplitude which we neglect for the sake of simplicity – this is why the states $|\psi_t(z^*)\rangle$ are unnormalized. In standard Born-Markov approximation (which leads to the Lindblad equation for ρ_t) it may be shown that there is a closed evolution equation for the system part $\psi_t(z^*)$ in the expansion, it reads

$$\partial_t\psi_t(z^*) = -\frac{\mathrm{i}}{\hbar}H\psi_t(z^*) - \frac{1}{2}L^\dagger L\psi_t(z^*) + Lz_t\psi_t(z^*) \tag{5.111}$$

with

$$z_t = -\mathrm{i}\sum_\lambda g_\lambda z_\lambda^* \mathrm{e}^{\mathrm{i}\omega_\lambda t}. \tag{5.112}$$

Remarkably, each term $|\psi_t(z^*)\rangle\ \|z\rangle$ constituting the entangled total state may be determined individually. Moreover, for the reduced density operator we find:

$$\rho_t = \mathrm{tr}_{\mathrm{env}}\left(|\Psi(t)\rangle(\Psi(t)|\right) = \int \frac{\mathrm{d}^2 z}{\pi}\mathrm{e}^{-|z|^2}|\psi_t(z^*)\rangle\langle\psi_t(z^*)|\ . \tag{5.113}$$

We may use the Gaussian weight over the coherent state labels z_λ as a probability distribution $\mathcal{M}[\ldots]$ and do a Monte-Carlo integration. Thus, equation (5.111) may be seen as a "stochastic" equation with, as it turns out, z_t white noise with

$$\begin{aligned} \mathcal{M}[z_t] &= 0 \,, \\ \mathcal{M}[z_t z_s] &= 0 \,, \\ \mathcal{M}[z_t^* z_s] &= \delta(t-s) \,. \end{aligned} \tag{5.114}$$

The product states $|\psi_t(z^*)\rangle \, \|z\rangle$ determine the exact, entangled state Ψ_t according to its expansion. Crucially, we may replace the fixed environmental state $\|z\rangle$ by a moving one $\|z(t)\rangle$, where the evolution of $z(t)$ is determined according to the Ehrenfest dynamics induced by the total Hamiltonian. This way we are led to consider the product states

$$\underbrace{|\psi_t\,(z(t))\rangle}_{\text{sys}} \, \underbrace{|z(t)\rangle}_{\text{env}} = |\tilde{\psi}_t\rangle \, |z(t)\rangle, \tag{5.115}$$

where it is easy to show that the system part $|\tilde{\psi}_t\rangle$, corresponding to a *moving* environmental state $|z(t)\rangle$ is indeed a solution of the nonlinear quantum state diffusion equation (5.99). Moreover, the reduced density operator is an ensemble of these pure states, $\rho_t = \mathcal{M}[|\tilde{\psi}_t\rangle\langle\tilde{\psi}_t|]$. The exact, entangling dynamics of the total state is thus approximated by a product state evolution of a robust system state and a coherent environmental state. We have found a way to identify robust states of an open system for any Hamiltonian H or coupling operator L to the environment as solutions of the nonlinear, stochastic quantum state diffusion equation (5.99).

5.4 Decoherence of Macroscopic Superpositions

The results obtained so far indicate that in an open quantum system the coherences between "robust" states $|s\rangle$ decay according to

$$\langle s|\rho_t|s'\rangle \sim \mathrm{e}^{-\gamma t D^2} \langle s|\rho_0|s'\rangle; \quad D \sim |s - s'| \,. \tag{5.116}$$

Here D is some dimensionless indicator of a "distance" between the superposed robust states, measured in units of a quantum reference length. As soon as the physical distance approaches macroscopic values, D may assume enormous values, like $\sim 10^{20}$ as we have seen in some examples. Thus, for such "macroscopic" superpositions, decoherence happens on incredibly short time scales. These results should not be taken too seriously. The reason is that in the typical description of open system dynamics (as explained in previous sections) one relies on the Markov approximation, which assumes all relevant system time scales to be long with respect to environmental correlation times. The replacement of the environmental correlation function by a delta function amounts to some version of long-time limit, or time course-graining: $\sum_\lambda |g_\lambda|^2 \mathrm{e}^{-\mathrm{i}\omega_\lambda (t-s)} \sim \delta(t-s)$.

Physically, there is a certain memory or bath correlation time scale $\tau_{\rm mem}$ (very often the inverse of some bath cutoff frequency Λ) that is assumed to be short compared to all other time scales, $T_{\rm sys} \gg \tau_{\rm mem} \simeq \Lambda^{-1}$. Therefore, the master equation approach cannot be trusted if it predicts a decoherence time of the order of $\tau_{\rm dec} \leq \tau_{\rm mem}$ which can easily happen for macroscopic superpositions $D \gg 1$. We may state this observation more drastically: Markov master equations are irrelevant for decoherence of macroscopic superpositions and thus fail to be the starting point for explaining the classical behavior of the macroworld. Nevertheless, decoherence is rather obviously a feature of the classical world: it is the argument that needs to be changed. We recall at this point that the experiments [8,9] resolving decoherence dynamics were performed under such "slow" decoherence conditions that the quadratic dependence of the decay on the distance D resulting from a master equation description was indeed valid. Yet the decoherence dynamics for more macroscopic superpositions may no longer be derived from master equations and new approaches have to be established. This will be the subject of the remaining part of this section. First, we state results for an exactly solvable model consisting of harmonic oscillators as "system" and "environment", coupled bilinearly through position. As valuable as the oscillator model is, we want to reveal decoherence under more general conditions, which can be achieved by making use of the very fact that decoherence is fast for macroscopic superpositions. Then, as will be shown in the final part of this section, a certain short time expansion of the dynamics reveals universal decoherence. For simplicity, we will for the reminder of this section restrict our attention to the loss of coherence between two superposed wave packets with largely different locations (x, x'), whose temporal fate is sufficiently captured by the evolution of the matrix elements $\rho(x, x', t)$ of the reduced density operator.

5.4.1 Soluble Model

It is possible to investigate rigorously [19,20,21,23] the dynamics of a harmonically bound Brownian particle with

$$H = \frac{p^2}{2m} + \frac{1}{2} m \Omega^2 q^2. \tag{5.117}$$

The total Hamiltonian of system and environment reads

$$H_{\rm tot} = H + q \underbrace{\sum_\lambda g_\lambda q_\lambda}_{Q} + \sum_\lambda \frac{p_\lambda^2}{2m_\lambda} + \frac{1}{2} m_\lambda \omega_\lambda^2 q_\lambda^2 \tag{5.118}$$

(we absorb the usual "counter term" in H, it will be irrelevant for what we are going to investigate). The usual quantum bath correlation function $\alpha(t-s) = \langle Q(t) Q(s) \rangle$ at finite temperature is microscopically given by (5.50). The equations of motion are linear, the propagator is Gaussian, and we can even write

down the exact equation of motion for the reduced density operator without any approximation [19,20],

$$\partial_t \rho = -\frac{\mathrm{i}}{\hbar}\,[H,\rho] - i\frac{A(t)}{\hbar}\,\big[q^2,\rho\big] + \frac{B(t)}{\hbar^2}\Big[q,\big[p,\rho\big]\Big] \\ -i\frac{C(t)}{\hbar}\Big[q,\big[p,\rho\big]\Big] - \frac{D(t)}{\hbar^2}\Big[q,\big[q,\rho\big]\Big]\,. \tag{5.119}$$

The real, time dependent coefficients $A(t),\ldots,D(t)$ can be expressed in terms of the two fundamental solutions $\{x_1(t),x_2(t)\}$ of the classical equation of motion

$$m\ddot{x}(t) + m\Omega^2 x(t) + 2\int_0^t \mathrm{d}s\,\mathrm{Im}\,[\alpha(t-s)]\,x(s) = 0 \tag{5.120}$$

with

$$x_1(0) = 1, \qquad \dot{x}_1(0) = 0\,, \\ x_2(0) = 0, \qquad \dot{x}_2(0) = 1\,. \tag{5.121}$$

For instance,

$$A(t) = \int_0^t \mathrm{d}s\,\mathrm{Im}[\alpha(t-s)]\Big\{\frac{x_1(s)\dot{x}_2(t) - x_2(s)\dot{x}_1(t)}{x_1(t)\dot{x}_2(t) - x_2(t)\dot{x}_1(t)}\Big\}\,, \tag{5.122}$$

$B(t)$ looks similar and $C(t)$, $D(t)$ are somewhat more involved. It should be appreciated that this particular model allows the derivation of a convolutionless, i.e. time-local, master equation for the reduced density operator ρ_t in spite of the generally non-Markovian dynamics of the open system (in contrast to the usual Nakayima-Zwanzig-formalism).

Decoherence of macroscopic superpositions (i.e. $|x-x'|$ of macroscopic magnitude) is fast. Therefore, only the short-time behavior of the coefficients is relevant, we find

$$A\,(t) = 0 + \mathcal{O}(t^2)\,, \\ B\,(t) = 0 + \mathcal{O}(t^3)\,, \\ C\,(t) = 0 + \mathcal{O}(t^2)\,, \\ D\,(t) = \alpha(0)\,t + \mathcal{O}(t^3)\,. \tag{5.123}$$

We conclude that it is the double commutator term involving $D(t)$ only that is initially relevant. Also, the system Hamiltonian H does not contribute in any way to this initial decay of coherences. We know the effect of the relevant "diffusion" term from earlier investigations (see Sect. 5.2.3) and find from (5.123) and (5.119)

$$\rho(x,x',t) = \exp\left(-\frac{1}{2}\frac{(x-x')^2}{\hbar^2}\,\alpha(0)t^2\right)\rho(x,x',0) \qquad \text{for} \qquad t\to 0. \tag{5.124}$$

Rather than the common exponential decay, we obviously encounter a Gaussian fall-off of the initial coherences. If the distance $|x-x'|$ is large enough, this Gaussian decay regime will be efficient enough to destroy all coherences, long before the usual long-time (Markovian) exponential decay will be a valid description.

What remains to be determined is the actual decoherence time scale. For simplicity, let us restrict ourselves to the case of high temperature T. When we wrote previously

$$\alpha(t-s) \approx 2m\gamma k_{\mathrm{B}} T\delta(t-s)\ , \tag{5.125}$$

for the Brownian kernel (neglecting damping again), we really have in mind an expression of the kind

$$\alpha(t-s) = m\gamma k_{\mathrm{B}} T\Lambda \mathrm{e}^{-\Lambda|t-s|} \tag{5.126}$$

where Λ denotes the high frequency cutoff of the bath oscillators. Then,

$$\alpha(0) = 2m\gamma k_{\mathrm{B}} T\frac{\Lambda}{2} = m\gamma k_{\mathrm{B}} T\Lambda, \tag{5.127}$$

and

$$D(t) \sim \hbar^2 \frac{\gamma}{\lambda_{\mathrm{dB}}^2}\,(\Lambda t) \qquad \text{for} \qquad t \to 0, \tag{5.128}$$

for short times. The reduced density operator looses coherence according to the quadratic law

$$\rho\,(x,x',t) = \exp\left(-\frac{(\gamma t)(\Lambda t)}{2}D^2\right)\rho\,(x,x',0) \tag{5.129}$$

with $D = |x-x'|/\lambda_{\mathrm{dB}}$ the dimensionless distance in units of the thermal de Broglie wave length.

For appreciable distance D, i.e. for macroscopic distance $|x-x'|$, all coherences disappear on the memory time scale Λ^{-1}, i.e. long before any Markov master equation can be valid. Still, we observe the dependence on D^2: Macroscopic superpositions will disappear rapidly, the decoherence time scaling only linearly with D rather than quadratically as for "slow" decoherence when the Markovian master equation description is valid.

These results indicate that the Markov master equation approach is irrelevant for decoherence of macroscopic superpositions, when the decoherence time scale can easily be the smallest time scale involved, even smaller than bath correlation time scales. It would be somewhat disappointing to have to rely on such exactly soluble models in order to see this behavior. As we will see next, an approach based on a short time expansion reveals this behavior under very general conditions.

5.4.2 Universality of Decoherence

We here present a somewhat simplified and shortened version of a general approach to short time decoherence dynamics as developed in [22,23] to which we

refer for details. Decoherence for macroscopic superpositions may be so fast that no motion due to the system Hamiltonian has a chance to be of relevance. In a first approximation, the system part of the total Hamiltonian may be discarded completely, and we assume $H_{\rm int} = S \cdot B$ with S a system operator (e.g. $S = q$, the position in the quantum Brownian motion model), and B a bath operator, typically consisting of additive contributions of many degrees of freedom, $B = \sum_\lambda B_\lambda$. The total Hamiltonian becomes

$$\begin{aligned} H_{\rm tot} &= \cancel{H_{\rm sys}} + H_{\rm int} + H_{\rm env} \\ &= S \cdot B + H_{\rm env} \ . \end{aligned}$$

The bath correlation function is $\alpha(t-s) = \langle B(t) B^\dagger(s) \rangle$, with $B(t) = \mathrm{e}^{\mathrm{i}H_{\rm env}t/\hbar}$ $B\ \mathrm{e}^{-\mathrm{i}H_{\rm env}t/\hbar}$.

Discarding the effect of the system Hamiltonian, we determine the initial dynamics of the reduced density operator,

$$\rho_t = \mathrm{tr}_{\rm env} \left\{ \mathrm{e}^{-\mathrm{i}(SB+H_{\rm env})t/\hbar} \left(\rho_0 \otimes \exp\left(-\frac{H_{\rm env}}{k_{\rm B}T}\right)\right) \mathrm{e}^{\mathrm{i}(SB+H_{\rm env})t/\hbar} \right\} \tag{5.130}$$

by employing the factorization

$$\mathrm{e}^{-\mathrm{i}(SB+H_{\rm env})t/\hbar} = \mathrm{e}^{-\mathrm{i}H_{\rm env}t/2\hbar}\ \mathrm{e}^{-\mathrm{i}\left(SBt+\mathcal{O}(t^3)\right)/\hbar}\, \mathrm{e}^{-\mathrm{i}H_{\rm env}t/2\hbar}\ , \tag{5.131}$$

such that

$$\rho_t = \mathrm{tr}_{\rm env} \left\{ \mathrm{e}^{-\mathrm{i}\left(SBt+\mathcal{O}(t^3)\right)/\hbar} \left(\rho_0 \otimes \exp\left(-\frac{H_{\rm env}}{k_{\rm B}T}\right)\right) \mathrm{e}^{\mathrm{i}\left(SBt+\mathcal{O}(t^3)\right)/\hbar} \right\} , \tag{5.132}$$

due to the cyclic property of the trace over the environment. Coherences between different s-eigenstates decay according to

$$\langle s|\rho_t|s'\rangle = \left\langle \mathrm{e}^{-\mathrm{i}(s-s')(Bt+\mathcal{O}(t^3))/\hbar} \right\rangle \langle s|\rho_0|s'\rangle, \tag{5.133}$$

where, as indicated, we neglect higher order terms in the short time expansion of the logarithm of the interaction Hamiltonian. We use $\langle \ldots \rangle$ to denote the expectation value with respect to the initial bath state, $\mathrm{tr}_{\rm env} \left\{ \ldots \mathrm{e}^{-H_{\rm env}/k_{\rm B}T} \right\}$.

We next make use of the fact that the bath coupling operator B is a sum of independent bath operators, and thus, according to the central limit theorem, may be considered to obey Gaussian statistics,

$$\left\langle \mathrm{e}^{-\mathrm{i}(s-s')Bt/\hbar} \right\rangle = \exp\left(-\frac{(s-s')^2}{2\hbar^2} \langle B^2 \rangle t^2 \right) . \tag{5.134}$$

We conclude that

$$\begin{aligned} \langle s|\rho_t|s'\rangle &= \exp\left(-\frac{(s-s')^2}{2\hbar^2} \langle B^2 \rangle t^2 \right) \langle s|\rho_0|s'\rangle \\ &= \exp\left(-\frac{(s-s')^2}{2\hbar^2} \alpha(0) t^2 \right) \langle s|\rho_0|s'\rangle, \end{aligned} \tag{5.135}$$

which is just the result (5.124),(5.129) we obtained for the exactly soluble model with $S = q$.

The only assumption concerning the nature of the bath is that B may be considered Gaussian, motivated by the observation that it is typically of the additive form $B = \sum_\lambda B_\lambda$ with operators B_λ of independent degrees of freedom. There is no need to assume an oscillator bath. We further neglected the system Hamiltonian $H_{\rm sys}$. This approximation may be verified through self consistency: the times we derive for the decoherence dynamics may indeed be so short that the system Hamiltonian has no effect. This will always be the case for large enough separations $D \simeq |s - s'|$, i.e. for macroscopically distinct states.

The quadratic decay of coherences for macroscopic separations D may thus be assigned a certain degree of universality, as it is largely independent of the nature of the system and the bath. Decoherence is still accelerated by a distance squared, D^2, yet a quadratic (t^2) time dependence of the decay in contrast to the Markovian master equation results that gave an exponential decay.

For completeness we mention that there are subtleties concerning the decay of coherences between eigenstates of a canonically conjugated variable, momentum in the case $S = q$. In the lowest order expansion shown here, neglecting the effect of the system Hamiltonian entirely, such superpositions would appear to be stable. It turns out, however, that they too decay rapidly, which requires the next-to-leading order, resulting in a fast decay of p-coherences according to a $\mathrm{e}^{-(t/\tau_{\rm dec})^4}$-law (for details, see [23]). Superpositions with respect to p need a little bit of time and free motion to turn into a superposition with respect to q. This explains the slower t^4-decay.

5.5 Conclusions

Decoherence in quantum dynamics leads to classical behavior, interference phenomena no longer occur. Certain robust states are singled out in open quantum system dynamics. They are those system states that are the least effected by the coupling to the environment, i.e. lead to only little entanglement between system and environment under time evolution. The loss of the ability to interfere is the more drastic the more macroscopically distinct the superposed robust states are. While the coupling to the environment may be so weak as to not notice any damping at all for quantities with a classical limit, coherences may nevertheless be destroyed rapidly. What states are considered robust depends crucially on the nature of the coupling between system and environment. Such interactions typically being dependent on position, localized wave packets, and thus "classical" states turn out to be robust. We may identify three regimes that lead to classical dynamics. First, for a highly non-classical initial state, there will be very fast decoherence towards eigenstates of the system part of the coupling S according to the quadratic law mentioned in the last section. This phase of decoherence is largely independent of the system Hamiltonian and the nature of the bath. After this very short universal initial period, there will be a competition between the effects of the interaction between system and environment and the system

Hamiltonian, singling out certain robust states, very often wave packets. As we have shown, such robust states finally follow closely the classical dynamics of point particles, yet with tiny fluctuations (partly of quantum origin). As experiments to follow the temporal course of decoherence are improving rapidly, it will be fascinating to see how far quantum superpositions may be extended further and further into the macroscopic domain.

Acknowledgement

I would like to thank M. Motzko for invaluable help during the preparation of this manuscript.

References

1. D. Giulini *et al.*: *Decoherence and the Appearance of a Classical world in Quantum Theory* (Springer, Heidelberg 1996)
2. Ph. Blanchard *et al.* (Eds.): *Decoherence: Theoretical, Experimental, and Conceptual Problems*, Lecture Notes in Physics 538 (Springer, Heidelberg 2000)
3. M. Tegmark, J. A. Wheeler: Scientific American (February 2001). Also: Spektrum der Wissenschaft (April 2001), with an additional comment by H. D. Zeh
4. W. H. Zurek: Physics Today (October 1991)
5. J. P. Paz, W. H. Zurek: 72nd Les Houches Summer School, July-August 1999. See also preprint arXiv:quant-ph/0010011
6. D. F. Walls, G. J. Milburn: *Quantum Optics*, 2nd edn. (Springer, Berlin 1995)
7. U. Weiss, *Quantum Dissipative Systems*, 2nd edition (World Scientific, Singapore 2000)
8. M. Brune *et al.*: Phys. Rev. Lett. **77**, 4887 (1996)
9. C. J. Myatt *et al.*: Nature **403**, 269 (2000)
10. A. O. Caldeira, A. J. Leggett: Physica A **121**, 587 (1983)
11. S. Habib *et al.*: Phys. Rev. Lett. **80**, 4361 (1998)
12. N. Gisin, M. Rigo: J. Phys. A **28**, 7375 (1995)
13. L. Diósi, C. Kiefer: Phys. Rev. Lett. **85**, 3552 (2000)
14. W. H. Zurek *et al.*: Phys. Rev. Lett. **70**, 1187 (1993)
15. N. Gisin, I. C. Percival: J. Phys. A **25**, 5677 (1992); **26**, 2245 (1993); **26**, 2245 (1993)
16. J. Dalibard, Y. Castin, K. Mölmer: Phys. Rev. Lett. **68**, 580 (1992)
17. H. Carmichael: *An open system approach to Quantum Optics* (Springer, Berlin 1994)
18. W. T. Strunz *et al.*: Phys. Rev. Lett. **83**, 4909 (1999)
19. F. Haake, R. Reibold: Phys. Rev. A **32**, 2462 (1985)
20. B. L. Hu, J. P. Paz, Y. Zhang: Phys. Rev. D **45**, 2843 (1992)
21. J. P. Paz, S. Habib, W. H. Zurek: Phys. Rev. D **47**, 488 (1993)
22. D. Braun, F. Haake, W. T. Strunz: Phys. Rev. Lett. **86**, 2913 (2001)
23. W. T. Strunz, F. Haake, D. Braun: Preprint arXiv:quant-ph/0204129; W. T. Strunz, F. Haake: Preprint arXiv:quant-ph/0205108

6 Quantum Communication and Decoherence

Hans Aschauer and Hans J. Briegel

Theoretische Physik, Ludwig-Maximilians-Universität,
Theresienstr. 37, D-80333 München

6.1 Introduction

The problem of decoherence is an integral part of the theory of quantum computation and communication. The potential of a quantum computer lies in its ability to process information in the form of a coherent superposition of quantum mechanical states. Quantum algorithms such as Shor's algorithm [1] make use of the interference of different "computational paths", which can strongly enhance their efficiency compared to classical algorithms. Because quantum coherence and interference play a central role in a quantum computer, *de*coherence is a major threat to its proper functioning.

A similar situation prevails in quantum communication. The central problem of quantum communication is how to faithfully transmit unknown quantum states through a noisy quantum channel. While quantum information is sent through a channel such as an optical fiber, the carriers of the information (e.g. photons) interact with the channel and get entangled with its many degrees of freedom, which gives rise to the phenomenon of decoherence on the state space of the information carriers. An intially pure state becomes a mixed state when it leaves the channel. For quantum communication purposes, it is however essential that the transmitted qubits retain their genuine quantum properties, for example in form of an entanglement with qubits on the other side of the channel.

To deal with the problem of decoherence, two methods have been developed, known as *quantum error correction* [2,3,4] and *entanglement purification* [5,6,7], respectively. In quantum error correction, which will be discussed in the next section, quantum information is encoded in the joint state of several two-state particles, forming a so-called quantum error correcting code, before it is sent through the channel. By measuring certain joint observables of the particles (the so-called stabilizer of the code), it is thereby possible to "reset" the state of the information carriers after a given time, by projecting their joint state onto certain subspaces of their Hilbert space without destroying the coherence of the encoded information. Even though quantum error correction can be used, in principle, to send quantum information through a noisy channel, it has been primarily developed to stabilize a quantum computer against the effect of decoherence. Entanglement purification, on the other hand, together with the method of teleportation [8], is a powerful tool that is particularly suitable for quantum communication. The idea of entanglement purification is to "distill" from an ensemble of low-fidelity Einstein-Podolsky-Rosen (EPR) [9] pairs,

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 235–261, 2002.

which have been distributed through some noisy channel, a smaller ensemble of high-fidelity EPR pairs which may then be used for faithful teleportation [5] or for quantum cryptography [10,11]. This distillation process requires certain unitary operations and measurements to be performed on the qubits at each side of the channel, and a process of postselection, which also requires classical communication between the parties.

Both methods, quantum error correction and entanglement purification, fight decoherence by a process of *controlled disentanglement* of the information carriers from the quantum channel. This process involves the action of some apparatus that is used to transform and measure the state of the particles, for example via tunable interactions of the particles with each other and with external fields. Real apparatuses are themselves sources of noise, which complicates the situation considerably. From a general perspective, the apparatuses used by Alice and Bob must themselves be considered as part of the noisy communication channel. Under realistic circumstances, the information carriers will thus *always* become entangled with other degrees of freedom and therefore suffer from a certain amount of decoherence. The question is therefore not whether decoherence can be avoided at all, but whether its influence can be kept on a tolerable level.

What "tolerable" means depends on the context. In quantum computation, for example, the effect of decoherence may be tolerable as long as the fidelity of the output of a quantum algorithm is above a certain value, allowing one to extract the desired result with the corresponding probability. In quantum cryptography the effect of the channel cannot, in principle, be distinguished from an intelligent third party who manipulates the transmitted quantum systems to gain information about their state. All noise of a channel is therefore attributed – this is the pessimistic attitude of the cryptologist – to an adversary. Decoherence is thereby considered due to entanglement of the information carriers with degrees of freedom controlled by an adversary. As we will show in the later part of this review, the *security* of quantum cryptography is in fact closely connected to the *disentanglement* of the degrees of freedom of the information carriers, on one side, and the channel, on the other side. Even though we cannot avoid all residual entanglement with the channel, we can distinguish between residual entanglement with the apparatus, which is harmless, and residual entanglement with the part of the channel accessible to an eavesdropper, which is potentially harmful.

In the following, we will give a brief introduction to the methods of quantum error correction and entanglement purification, and to the basic protocols of quantum cryptography.[1] We will then discuss a recent security proof [13] for entanglement-based quantum communication through noisy channels, which explicitly takes into account the role of noisy apparatus. We try to pay particular attention to conceptual issues but skip some of the technical details, which can be found in the literature.

[1] For a more comprehensive introduction into these fields of quantum information theory see, for example, [12].

6.2 Quantum Error Correction

Quantum mechanical entanglement is exploited in quantum algorithms and in many protocols for quantum communication such as teleportation or entanglement-based quantum key distribution. It also plays a fundamental role in quantum error correction, where the coding operations are themselves simple quantum algorithms. (For quantum error correction and the related topic of channel capacities see also Chap. 7.)

Let us illustrate the basic principles at the example of the first quantum error correcting code found by Peter Shor in 1995 [2]. To protect quantum information that is represented by the state of a particle (central qubit in Fig. 6.1) against decoherence, the information is first distributed or delocalized over several particles. In Fig. 6.1 this is done with the help of the network ENC, which realizes the following mapping:

$$ENC: \ (\alpha|0\rangle + \beta|1\rangle)|0\rangle|0\rangle\cdots|0\rangle \longmapsto \alpha|0\rangle_S + \beta|1\rangle_S \tag{6.1}$$

in which the states

$$\begin{aligned} |0\rangle_S &= 2^{-3/2}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \ , \\ |1\rangle_S &= 2^{-3/2}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \ . \end{aligned} \tag{6.2}$$

denote the so-called *code words* of the (9-bit) Shor code. The encoding transformation thus corresponds to an embedding $\mathcal{H} \ni |\phi\rangle = \alpha|0\rangle + \beta|1\rangle \longmapsto \alpha|0\rangle_S + \beta|1\rangle_S = |\phi\rangle_S \in \mathcal{H}_S \subset \mathcal{H}^{\otimes 9}$ of the two-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^2$ of the central qubit into the higher-dimensional Hilbert space of all 9 qubits. After the transformation the quantum information lies in a two-dimensional subspace $\mathcal{H}_S$ of a 2^9 dimensional Hilbert space. The code words $|0\rangle_S$ and $|1\rangle_S$ are tensor products of entangled three-qubit states of the form $|000\rangle \pm |111\rangle$, the so-called Greenberger-Horne-Zeilinger (GHZ) states [14], which play a prominent role for the interpretation of quantum mechanics [14,15]. One can easily check that after the encoding (see the dotted line in Fig. 6.1) the reduced density operator of each of the qubits is totally mixed; that is, the *individual* state of the particles carries no information about $|\phi\rangle$.[2]

For simplicity let us consider an error model where random rotations are applied to the individual qubits with a certain "error rate". This model is more general than it appears to be at first sight but it needs a justification to which we shall return below. Suppose that, after the encoding circuit of Fig. 6.1, one of

[2] Quantum error correcting codes are indeed constructed in such a way that the state of individual qubits in a codeword becomes completely undetermined. As was shown by DiVincenzo and Peres [16], the codewords satisfy generalized Mermin relations [15] that exclude the possibility of consistently assigning a predetermined value to complementary observables of each qubit. From the measurement of an individual qubit one can thus not gain any information about $|\phi\rangle$. In the positive sense this means that an uncontrolled interaction of the environment with one of the qubits does not (necessarily) lead to an irreversible *loss* of information.

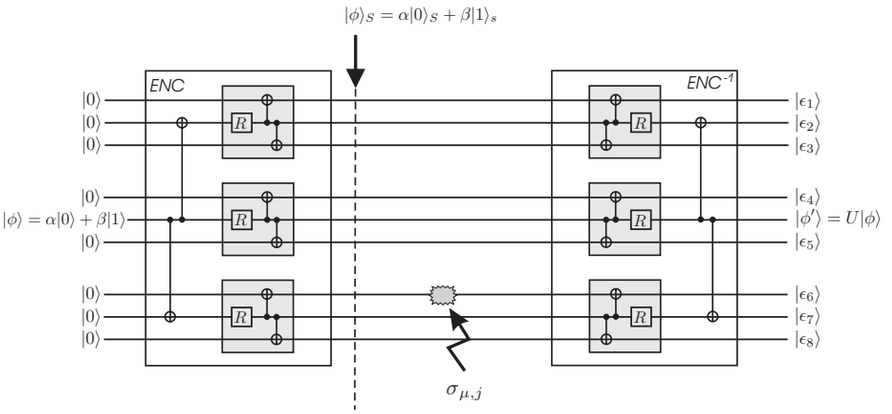


Fig. 6.1. Quantum logic network of the Shor code and quantum error correction. A "random rotation" $\sigma_{\mu,j}$ on qubit j in the encoded state translates into a certain error syndrome $\varepsilon_1, \ldots, \varepsilon_8$ and a corresponding unitary operation $U = U(\varepsilon_1, \ldots, \varepsilon_8)$ on the central qubit (see text). The network uses the Hadamard-Rotation $R_j = 1/\sqrt{2}(\sigma_{x,j} + \sigma_{z,j})$ and the CNOT gate, $\overset{\cdot}{\oplus}$ $= \mathrm{CNOT}_{i,j} = (1 + \sigma_{z,i})/2 + (1 - \sigma_{z,i})\,\sigma_{x,j}/2$

the four Pauli-Rotations $\sigma_{\mu,j}$ ($\mu = 0, 1, 2, 3$) is applied to one of the nine qubits, where both j and μ are random and unknown to us. The question is this: Can we still extract $|\phi\rangle$ from the joint state of the particles? If such a random rotation is applied to the particle *before* the encoding, it is clear that the information is lost, for $\frac{1}{4}\sum_\mu \sigma_\mu|\phi\rangle\langle\phi|\sigma_\mu = \frac{1}{2}\mathbb{1}$. If it is applied to one of the particles of the encoded state, however, the information can still be rescued from the joint state of all 9 particles. A possibility to do this is shown in Fig. 6.1. There the inverse network of ENC is applied to the code which transforms the corrupt state $\sigma_{\mu,j}|\phi\rangle_S$, for arbitrary $\sigma_{\mu,j}$, back to a product state:

$$ENC^{-1}:\ \sigma_{\mu,j}|\phi\rangle_S \longmapsto |\phi'\rangle|\varepsilon_1\rangle|\varepsilon_2\rangle\ldots|\varepsilon_8\rangle\,. \tag{6.3}$$

After this decoding transformation, the state of the neighboring qubits carries the so-called *error syndrome* $\varepsilon_1 \ldots \varepsilon_8$. The central qubit is in state $|\phi'\rangle = U(\varepsilon_1 \ldots \varepsilon_8)|\phi\rangle$, where the unitary transformation $U(\varepsilon_1 \ldots \varepsilon_8) \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$, is uniquely determined by the error syndrome.

By reading off the error syndrome, i.e. measuring the state of all neighboring qubits, and subsequently applying the correction operation $U^{-1}(\varepsilon_1 \ldots \varepsilon_8)$, the central qubit is transformed back to its initial state. Please note that the central qubit remains unmeasured, and no information about the state $|\phi\rangle$ is obtained at any step of the protocol. By iteration of the sequence *decoding* $\rightarrow$ *syndrome measurement* & *correction* $\rightarrow$ *encoding* [17] an unknown quantum state can thus be protected against decoherence over a time significantly longer than the decoherence time.

The effect of the random rotations $\sigma_{\mu,j}$ is to map the code space $\mathcal{H}_S$ to a set of orthogonal error spaces $\sigma_{\mu,j}\mathcal{H}_S \perp \mathcal{H}_S$. The images of the code words

thereby satisfy the following orthogonality relations ${}_S\langle 0|\sigma_{\mu,j}\sigma_{\nu,k}|1\rangle_S = 0$ and ${}_S\langle 0|\sigma_{\mu,j}\sigma_{\nu,k}|0\rangle_S = \langle 1|\sigma_{\mu,j}\sigma_{\nu,k}|1\rangle_S$ for all j,k,μ,ν. Theses relations ensure [6,18] that all errors $\sigma_{\mu,j}$ can, in fact, be corrected. The Shor code was the first quantum error correcting code found that can correct all of the four errors (spin flip, phase flip, spin&phase flip, identity) on any one of the qubits. Independent of Shor, Steane [4] found a code that achieves the same task using only 7 qubits. Later, the theory of quantum error correcting codes was further developed [3,19], establishing in particular the connection with classical coding theory. A number of other codes were found, among them a so-called 'perfect' code using a minimum number of only 5 qubits [17,6]. One can also construct codes that are able to correct more than a single qubit error. These satisfy a similar set of orthogonality relations of the form given above, and the code words are entangled states of an increasing number of qubits. An introduction to the theory of quantum error correction can be found in the articles by Steane [20,21] and by Gottesman [22], for example.

In quantum error correction we exploit, as in quantum algorithms, the possibility to manipulate superpositions of states of a quantum register and to measure joint observables which describe joint properties of several qubits. The operation ENC^{-1} for example, followed by one-qubit measurements, corresponds to the measurement of the "parity" of different qubits, as was shown by Gottesman [23]. The joint observables are here[3]

$$\begin{aligned}
M_1 &= \sigma_{z,1}\sigma_{z,2}\,,\\
M_2 &= \sigma_{z,2}\sigma_{z,3}\,,\\
M_3 &= \sigma_{z,4}\sigma_{z,5}\,,\\
M_4 &= \sigma_{z,5}\sigma_{z,6}\,,\\
M_5 &= \sigma_{z,7}\sigma_{z,8}\,,\\
M_6 &= \sigma_{z,8}\sigma_{z,9}\,,\\
M_7 &= \sigma_{x,1}\sigma_{x,2}\sigma_{x,3}\sigma_{x,4}\sigma_{x,5}\sigma_{x,6}\,,\\
M_8 &= \sigma_{x,4}\sigma_{x,5}\sigma_{x,6}\sigma_{x,7}\sigma_{x,8}\sigma_{x,9}\,.
\end{aligned} \tag{6.4}$$

The error spaces $\sigma_{\mu,j}\mathcal{H}_S$ are eigenspaces of these observables with eigenvalues ± 1. The observable $M_1 = \sigma_{z,1}\sigma_{z,2}$, for example, tells us whether on qubit 1 or 2 a spin flip has occurred without revealing any information on which of the qubits: $M_1(|000\rangle + |111\rangle) = +(|000\rangle + |111\rangle)$, $M_1(|100\rangle + |011\rangle) = -(|100\rangle + |011\rangle)$, $M_1(|010\rangle + |101\rangle) = -(|010\rangle + |101\rangle)$. By measuring both observables $M_1 = \sigma_{z,1}\sigma_{z,2}$ and $M_2 = \sigma_{z,2}\sigma_{z,3}$ one can find out whether, and on which of the qubits 1, 2, 3 a spin flip has taken place.

The measurement of the eigenvalues of these joint observables can be realized, as described in Fig. 6.1, by the method "decode and subsequently measure the individual state of the surrounding qubits". This strategy has the disadvantage that the decoding leaves the logical qubit in an unprotected state, exposing

[3] These observables generate an Abelian group, the so-called *stabilizer* of the Shor code [23].

it directly to the influence of decoherence. There are different methods (or networks, respectively) which use so-called ancillas to perform the error detection and correction on the encoded state directly. This is the subject of fault-tolerant quantum error correction and, more generally, fault-tolerant quantum computation. It takes into account the fact that the elementary operations which are part of the encoding and decoding network may themselves be imperfect and subject to errors. Thus one needs to make sure that the correction operations do not introduce more errors into the system than they extract. A description of the theory of fault-tolerant quantum computation is beyond the scope of this introduction. A central result states that it is possible, by using concatenated encoding strategies, to maintain the coherence of the logical state of a quantum computer over an arbitrarily long time, given that the error probability (noise level) of the elementary operations (quantum gate, measurement) is below a certain *threshold*. The price one has to pay for this is a certain *overhead* in the number of auxiliary (physical) qubits that scales polynomially [or under certain circumstances even polylogarithmically] with the time of computation. The threshold is very small and of the order of $10^{-4} - 10^{-5}$. The theory of fault-tolerant computation is considered as the general solution of the problem of decoherence and imperfect apparatus for quantum computation. An introduction into the subject is given by Preskill [24], for example.

What is an Error? Let us return to the question whether the model of an error as a random unitary rotation is reasonable. The interaction of the qubits with the environment can be described as a unitary evolution in the Hilbert space of the total system consisting of both the qubits and the environment. Where and in what sense do "errors" happen in this picture? This question is certainly justified. One can show, however, that any interaction of the qubits and the environment can be written in the (integrated) form

$$|\phi\rangle_S|u\rangle_{\text{env}} \to \sum_k (F_k|\phi\rangle_S)|u_k\rangle_{\text{env}} \tag{6.5}$$

where the operators F_k are tensor products of Pauli operators and $|u_k\rangle_{\text{env}}$ states of the environment which in general are neither orthogonal nor normalized [21]. This result remains true if $|\phi\rangle_S$ is replaced by an arbitrary multi-qubit state [21]. In case of a quantum error correcting code, such as the Shor code, one has the additional property that ${}_S\langle\phi|F_k^{\text{1bit}}|\phi\rangle_S = 0$, for all "1-bit operations" F_k^{1bit} which contain a non-trivial Pauli operator only at a single position (that is, for $F_k^{\text{1bit}} \sim \mathbb{1}\otimes\cdots\otimes\mathbb{1}\otimes\sigma_\mu\otimes\mathbb{1}\otimes\cdots\otimes\mathbb{1}$). For weak interactions and for uncorrelated noise, these are the terms of first order in an expansion with respect to the interaction strength. Since the overall evolution on the space of the qubits plus the environment is unitary, the picture of randomly occurring errors, which are subsequently corrected, is only a helpful way of thinking about the problem. The "digitalization of noise" [21] is in fact only introduced via the measurement of certain observables such as the M_k. By measuring the M_k the state (6.5) of the system is projected "back" into the code space $\mathcal{H}_S$ or any of the error spaces

$F_k^{\text{1bit}}\mathcal{H}_S \perp \mathcal{H}_S$ that are orthogonal to it. It is the *dis*entanglement of the qubits from the environment which is the crucial process. The "error" is introduced by the fact that one does not always project back to the code space but sometimes also into an orthogonal error space, so that subsequently a unitary correction operation has to be applied to rotate the state back into the code space. On the Hilbert space of the qubits alone, the entire process can effectively be described as if the environment would apply random rotations $\sigma_{\mu,j}$ on the code, which we then check and possibly correct. Similar remarks apply to codes that correct several errors at the same time, which take into account terms of higher order in the expansion (6.5).

6.3 Entanglement Purification

In quantum communication, entanglement between distant parties plays a predominant role. In the following, we will concentrate on communication scenarios which involve two parties, Alice and Bob. What does it mean when we say that Alice and Bob have entanglement at their disposal? Usually, this means that they own quantum systems whose state is entangled, or, in technical terms, that the density operator which describes the state of the two quantum systems cannot be written as a convex combination of product states [25]. The two entangled quantum systems are usually called EPR pairs, due to the famous paper by Einstein, Podolsky and Rosen [9]. In the context of quantum information theory, the EPR pairs often consist of two entangled two-level systems (qubits), one owned by Alice, and the other by Bob. Maximally entangled two-qubit states are called *Bell states*; one can find four orthogonal Bell states, which form a basis of the two-qubit Hilbert space, the *Bell basis*.

The importance of entanglement is due to the fact that it is a resource which is equivalent to a quantum channel: If Alice and Bob are connected with a quantum channel, Alice can create an EPR pair locally and send one half through the quantum channel to Bob. On the other hand, if Alice and Bob own EPR pairs, they can use them to teleport qubits [8], even when they are not connected via some "real" quantum channel like an optical fiber.

The questions remains however, how can Alice and Bob obtain perfect EPR pairs if they can only communicate via a noisy channel? Any real quantum channel interacts with the quantum systems which are sent through it: it becomes entangled with them. This fact is important if Alice uses the channel in order to distribute EPR pairs. If the EPR pairs are subsequently used for teleportation, then the teleported qubits become entangled with the quantum channel.

Entanglement purification protocols [5,6,7] can be used to overcome this problem. Simply speaking, these protocols create an ensemble of highly entangled pairs out of a larger ensemble of pairs with low fidelity. The fidelity of a quantum state ρ is defined as its overlap with a given Bell state $|\Phi^+\rangle$, say, i. e. $F = \langle\Phi^+|\,\rho\,|\Phi^+\rangle$.

The purified pairs provide Alice and Bob with a purified quantum teleportation channel. If this channel is used for quantum communication, the qubits

are protected against an unwanted interaction with the channel. In the next sections, we will see that this fact can be exploited for quantum cryptography protocols.

In order to perform an entanglement purification protocol, classical communication between Alice and Bob is necessary. This means, that both Alice and Bob perform measurements on their respective qubits, and tell each other the measurement outcomes. For some protocols only *one-way* communication is required, i. e. only Alice will send classical messages to Bob. It has been shown by Bennett *et al.* [6], that these one-way entanglement purification protocols are equivalent to quantum error correcting codes (see Sect. 6.2). A tutorial introduction to the basic idea of entanglement purification is given in [26]

6.3.1 2-Way Entanglement Purification Protocols

The two-way entanglement purification protocols (2-EPP) which we present here have been developed by Bennett *et al.* [5] and, later, by Deutsch *et al.* [7]. Since these protocols work in a recursive way, they are often referred to as *recurrence protocols*. In order to distinguish between both protocols, we will call them *IBM* and *Oxford* protocol, respectively. The IBM protocol introduces a *twirling* operation after each purification step, which transforms the state of the EPR pairs into the Werner form. Since Werner states [25] are described by only one real parameter, all calculations can be done analytically. A disadvantage of the IBM protocol is that it is less efficient in producing pure states from noisy ones than the Oxford protocol. Qualitatively, there is no difference between both protocols.

To be precise, we want to distinguish between the purification *protocol* and the *distillation process* (see Fig. 6.2).

In each step, the purification protocol acts on two pairs of qubits. For the sake of simplicity, we shall assume that these two pairs are described by the density operator $\rho_{AB} \otimes \rho_{AB}$, which is thus a four-qubit density operator. The Oxford protocol (see Fig. 6.2) consist of the following steps:

1. Alice and Bob perform one-qubit $\pi/4$ rotations about the x-axis on each of their qubits (in opposite directions). If the qubits were stored in atomic/ionic degrees of freedom inside a trap, this could be implemented by (simple) laser pulses.
2. Both Alice and Bob perform a CNOT-operation (controlled NOT) [12], where they use their respective particle of pair one (two) as the source (target). This is the part of the protocol which is most difficult to perform experimentally.
3. Finally, both Alice and Bob measure the qubits which belong to pair two in the σ_z-basis, and tell each other the results (two-way communication). Whenever the results coincide, the keep pair one, otherwise they discard it. In either case, they have to discard the second pair, because it is projected onto a product state by the measurement.

In order to see how this protocol works, it is useful to write the density matrices in the Bell basis, i. e. in the basis of the two qubit Hilbert space which consists

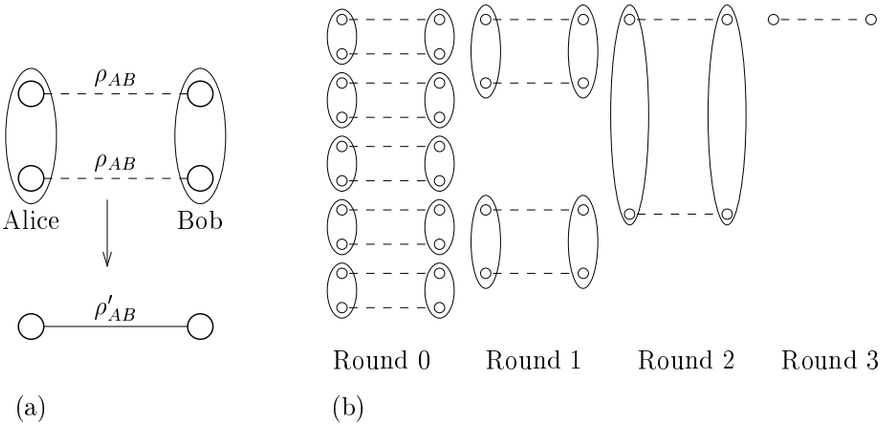


Fig. 6.2. (**a**) The entanglement purification protocol is a (probabilistic) protocol, which creates a higher entangled pair of qubits out of two pairs with lower entanglement. Usually these pairs are called source and target pair, respectively. Through an interaction between the qubits of the source and the target pair, realizing a so-called CNOT operation on each side, the states of all four qubits become correlated. By measuring the qubits of the target pair, the source pair is probabilistically projected into a new state ρ'_{AB}, which is more entangled than the original state ρ_{AB}.(**b**) The distillation process consists of several *rounds*. In each round, the pairs are combined into groups of two at a time, and the purification protocol is applied to them. From round to round, the entanglement of the remaining pairs is increased

of the four Bell states $|\Phi^{\pm}\rangle = 1/\sqrt{2}\,(|00\rangle \pm |11\rangle)$ and $|\Psi^{\pm}\rangle = 1/\sqrt{2}\,(|01\rangle \pm |10\rangle)$:

$$\begin{aligned}\rho_{AB} = \quad & A\left|\Phi^{+}\right\rangle\!\left\langle\Phi^{+}\right| + B\left|\Psi^{-}\right\rangle\!\left\langle\Psi^{-}\right| \\ & + C\left|\Psi^{+}\right\rangle\!\left\langle\Psi^{+}\right| + D\left|\Phi^{-}\right\rangle\!\left\langle\Phi^{-}\right| \quad + \text{off-diag. elements}\end{aligned} \tag{6.6}$$

The coefficients A, B, C, and D are called the *Bell diagonal elements* of the density matrix ρ_{AB}. For any physical state, these coefficients have to fulfill the normalization condition $\operatorname{tr}\rho_{AB} = A + B + C + D = 1$.

As it turns out, the Bell diagonal elements A', B', C' and D' of the remaining pair do not depend on the off-diagonal elements of ρ_{AB}. For this reason, we can find a recurrence relation for the Bell diagonal elements, which describes their evolution during the distillation process (the index n belongs to the state of the pairs at the beginning of round number n in the distillation process:

$$\begin{aligned} A_{n+1} &= \frac{A_n^2 + B_n^2}{N}, \quad B_{n+1} = \frac{2C_nD_n}{N} \\ C_{n+1} &= \frac{C_n^2 + D_n^2}{N}, \quad D_{n+1} = \frac{2A_nB_n}{N} \end{aligned} \tag{6.7}$$

The normalization $N_n = (A_n + B_n)^2 + (C_n + D_n)^2$ is equal to the probability p_{success} that Alice and Bob obtain the same measurement results in step 3 of the

protocol. Even though no analytical solution has been found for this recurrence relation, it has been shown (numerically in [7] and later analytically [27]) that it converges to the fixpoint $A^\infty = 1, B^\infty = C^\infty = D^\infty = 0$, whenever the initial fidelity is greater than $1/2$. In this case, also the off-diagonal elements will vanish, since the density matrix has to be positive. In other words, whenever Alice and Bob are supplied with EPR pairs with a fidelity of more than 50%, they can distill (asymptotically) perfect EPR pairs.

For the IBM protocol, one only needs one recurrence relation, since (one-parametric) Werner states, described by $A = F, B = C = D = (1 - F)/3$ and vanishing off-diagonal elements in (6.6), are mapped onto Werner states. This map is shown in Fig. 6.3a. The map has tree fixpoints. Two of these fixpoints are attractive (at $F = 1/4$ and $F = 1$), and the remaining one (at $F = 1/2$) is repulsive. Thus, if one starts the distillation process with a fidelity greater than $1/2$, one will finally reach EPR pairs in a pure state. If the initial fidelity is smaller than $1/2$, one will finally be left with completely depolarized pairs, which correspond to a Werner state with a fidelity of $1/4$.

6.3.2 Purification with Imperfect Apparatus

Up to now, we have assumed that the only source of decoherence is the quantum channel which connects Alice and Bob. For practical implementations, however, this is an over-simplification. Indeed, there are many operations involved in the distillation process: Qubits have to be stored for a certain time, one- and two-qubit unitary operations will act on them, and there are measurements. Each of these operations is a source of noise by itself. It would be inconsistent to ignore this source of noise. So the following question arises: What are the conditions which we have to impose on the apparatus so that entanglement distillation works at all?

As we have mentioned in the context of fault tolerant quantum computation, there exists a certain noise threshold for the elementary operations, below which fault tolerant quantum computation is possible. In the case of 2-EPP we will find a threshold which is much more favorable than the threshold for fault tolerant quantum computation.

In order to get a qualitative understanding of the influence of noisy operation on the entanglement distillation process, we look again at the purification curve (Fig. 6.3). The curve shows how the fidelity after a purification step depends on the previous fidelity. If noise is introduced in the purification process itself, it is intuitively clear that only a smaller increase in fidelity can be achieved: the purification curve is "pulled down". In Fig. 6.3b this is shown schematically. We thus expect that in the case of noisy operations, one has to start with a greater initial fidelity in order to purify at all, and that the maximum fidelity which can be reached will be smaller than unity.

If the noise level is increased, one reaches the situation that two of the fixpoints will merge. At even higher noise levels, the purification curve has only the trivial fixpoint which corresponds to completely depolarized pairs: the distillation process breaks down and does not work any longer.

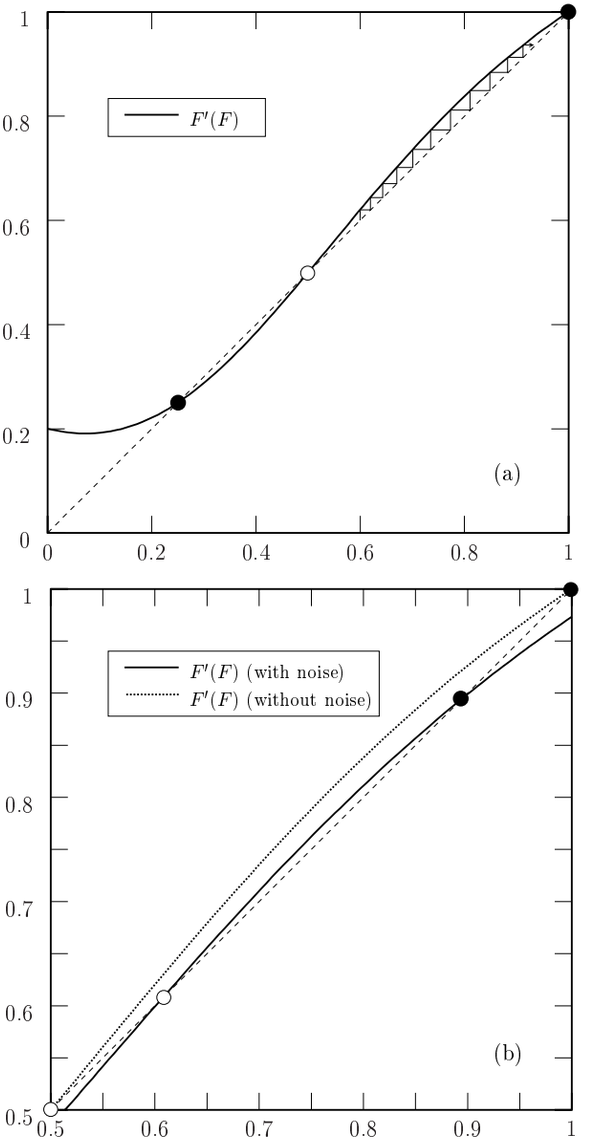


Fig. 6.3. The purification curve for the IBM protocol [5,6] for perfect (i. e. noiseless) apparatus (**a**). The staircase denotes how the fidelity increases from round to round in the distillation process of Fig. 6.3(**b**). If the apparatus is imperfect, the purification curve is "pulled down" (**b**) and the fixpoints move towards each other. The upper fixpoint of the curves indicates the maximum achievable fidelity $F_{\max}$, which can be reached asymptotically by the respective purification protocols; $F_{\max}$ decreases with an increasing noise level. Attractive fixpoints are denoted by black circles, repulsive fixpoints by white circles

The quantitative investigation of entanglement purification with noisy apparatus [28,29] shows that the above considerations are qualitatively correct. For the calculation, the following noise model has been assumed [29]:

- The unitary evolution of the qubits is accompanied by a depolarizing channel. It is well-known that this can be written in a time-integrated form

$$\rho_{AB} \to p\, U_A\, \rho_{AB}\, U_A^{-1} + \frac{1-p}{d}\, \mathbb{1}_A \otimes \mathrm{tr}_A\, \rho_{AB}. \tag{6.8}$$

  Here, ρ_{AB} is the density operator which describes the state of a bipartite quantum system, U_A is the desired unitary operation (which is assumed to act only on the quantum system at party A), d is the dimension of the Hilbert space of A's system, and p is the *reliability* of the quantum operation. For $p = 1$, there is no noise at all, and for $p = 0$, the quantum system at A becomes completely depolarized.
- Measurements give the correct results only with a certain probability η. This can be conveniently described in terms of a POVM (positive operator valued measure [30]),

$$\begin{aligned} M_0 &= \eta\, |0\rangle\langle 0| + (1-\eta)\, |1\rangle\langle 1| \\ M_1 &= \eta\, |1\rangle\langle 1| + (1-\eta)\, |0\rangle\langle 0|\,, \end{aligned} \tag{6.9}$$

  for one-qubit measurements in the σ_z basis. Here, $\mathrm{tr}(M_j \rho)$ describes the probability with which the detector indicates the result "j" for the measured qubit.

As one can see from (6.8), we have to distinguish between one- and two-qubit operations, if they are accompanied by noise: a two-qubit depolarizing channel is different from two one-qubit depolarizing channels. The first is an example of a *correlated* noise channel, the latter of an *uncorrelated* noise channel. The reliability of one- and two-qubit operations is referred to as p_1 and p_2, respectively. Whether or not entanglement purification is possible with a certain protocol, depends on the three parameters p_1, p_2, and η. For all these parameters, one gets a noise threshold in the percent regime, which is about two orders of magnitude better than the noise threshold for fault tolerant quantum computation.

6.4 Quantum Cryptography

One of the practically most advanced fields in quantum communication is quantum cryptography. In this section, we will describe the two basic protocols of quantum cryptography. We show that decoherence in the (untrusted) quantum channel as well as in the (trusted) apparatus plays an important role in the security analysis of quantum cryptography protocols.

The communication scenario in the cryptographic context looks as follows: Alice wants to send a confidential message (*clear-text*) to Bob, while a third communication party, Eve, wants to listen in and learn as much as possible about the message. In order to achieve her goal, Alice encrypts the message using some

cryptographic method. The encrypted message is called *ciphertext*. A cryptographic protocol is considered *good*, if it is possible to restrict the information which Eve can obtain to any desired level.

There exist several categories of *classical* cryptographic protocols; these include symmetric key ciphers, asymmetric key ciphers and one-time pads. All these protocols have advantages and disadvantages, but the most eminent advantage of the one-time pad is that it has been proved to be secure in the information theoretical sense: one can show that an eavesdropper can gain no information (zero bits of information) about the message, even if he or she knows every single bit of the encrypted message. To this end, it is however necessary that Alice and Bob share a secret and random key, which must at least be as long as the message which Alice wants to transmit, and that this key will only be used once (thus the name *one-time pad*).

The one-time pad works as follows: As a key, Alice and Bob share a secret string of zeros and ones $s = (s_1, s_2, \ldots, s_N)$. Similarly, Alice can write the clear-text (like any piece of information) as a string of zeros and ones, using some encoding which Alice and Bob agree on publicly. The clear-text is thus given in a *binary representation* $t = (t_1, t_2, \ldots, t_N)$. For the ciphertext, Alice adds the key and the clear-text bitwise modulo 2: $c = (s_1 \oplus t_1, s_2 \oplus t_2, \ldots, s_N \oplus t_N)$. In order to decrypt the message, Bob simply adds the key bitwise (modulo 2) to the ciphertext, and gets back the binary representation of the clear-text.

The key used in the one-time pad protocol is a valuable resource, to both the legitimate communication parties and to an eavesdropper: Alice and Bob use up the key during the communication. In order to supply themselves with a new key, they have to meet each other physically. On the other hand, if Eve knows the key, the communication between Alice and Bob is no longer a secret for her; for this reason, the cryptographic key might be a valuable target for theft or bribery. The aim of quantum cryptography is to solve this shortcoming of classical cryptography. In most quantum cryptography protocols, the quantum part of the protocol is related to the distribution of a key (quantum key distribution, QKD), which can afterwards, as soon as it is established, be used for a classical one-time pad protocol.

6.4.1 The BB84 Protocol

The first protocol for quantum key distribution was given by Bennett and Brassard in 1984 [10]. This so-called BB84 protocol is widely used in quantum cryptography, since all security considerations are well analyzed, and it is easy to understand.

The protocol works as follows: Alice prepares two random binary strings, the key string $(k_1, k_2, \ldots, k_N)$ and the basis string $(b_1, b_2, \ldots, b_N)$. The randomness of the bits is crucial for the security of the protocol; they may thus not be chosen by a pseudo random number generator.

There are 4 different quantum states which Alice can prepare: $|s_{00}\rangle = |0\rangle$, $|s_{01}\rangle = |1\rangle$, $|s_{10}\rangle = |+\rangle \equiv 1/\sqrt{2}(|0\rangle + |1\rangle)$, $|s_{11}\rangle = |-\rangle \equiv 1/\sqrt{2}(|0\rangle - |1\rangle)$. For simplicity, we will now consider the case of qubits which are represented in the po-

larization degree of freedom of a photon. In this case, the four states which Alice can prepare are horizontally, vertically, or $\pm 45°$ polarized photons.

Alice sends N photons through the quantum channel to Bob. The state in which the qubits are prepared depends on the key- and and the basis string: the ith qubit is prepared in the state $|s_{b_i k_i}\rangle$.

Bob can measure each photon that arrives in his laboratory either in the $|0\rangle\,/\,|1\rangle$-basis (i. e. in the horizontal/vertical basis), or in the $|+\rangle\,/\,|-\rangle$-basis (i. e. in the $\pm 45°$ polarized basis). For each individual photon, he selects the measurement basis randomly, and he writes down the chosen basis and the measurement result. When Bob has received and measured the N photons, he is left with two strings of N bits: the "basis" string and the "result" string.

Alice and Bob exchange their respective basis strings through a classical channel, which may be public; for example, they might announce the basis strings in a newspaper. It is no security breach if Eve knows both basis strings. However, Alice and Bob must make sure that Eve cannot alter these messages. One possibility to achieve this goal is that Alice and Bob posses an initial shared secret, which can be used to check the authenticity and integrity of the basis strings. During the key distribution task, this initial shared secret can be recreated, so that it is not used up; rather, it plays the role of a catalyst. By comparing their basis strings, Alice and Bob can see which photons have been measured in the same basis in which they have been prepared. Whenever the preparation basis and the measurement basis are different, Bob's measurement result is completely random and cannot be used. On the other hand, if the two bases are the same, Bob's measurement result will be strictly correlated with Alice's key bit for the respective photon: Alice's key bits and Bob's measurement results for these photons can be used as a secret key.

Before the key can be used, Alice and Bob have to make sure that the quantum channel has not been eavesdropped. One way to do this is the following: Alice chooses a certain number of the key bits randomly and sends them to Bob through the classical public channel. Bob compares Alice's key bits with his result bits, and if they are equal, they can be sure that there was no eavesdropper who tapped the quantum channel. This is due to the fact that the only quantum operation which does not disturb non-orthogonal quantum states is the identity. In other words: if Eve does not want to disturb the non-orthogonal quantum states which Alice sends, she has to leave them alone.

6.4.2 The Ekert Protocol

The main difference between the BB84 protocol and the so-called E91 protocol found by Ekert in 1991 [11] is that it does not use single photons which one communication party sends to the other, but pairs of entangled photons. While its experimental realization is more difficult than the BB84 protocol, it has a theoretical advantage: the security of the E91 protocol is related to the fact that there exists no local realistic theory which explains the outcomes of Bell-type

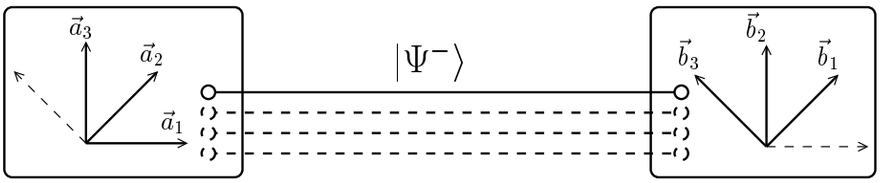


Fig. 6.4. The measurement directions in the Ekert protocol. For each EPR pair, Alice and Bob choose independently and randomly one of the three measurement directions $\vec{a}_1, \vec{a}_2, \vec{a}_3$ and $\vec{b}_1, \vec{b}_2, \vec{b}_3$, respectively

experiments.[4] While in the BB84 protocol one has to believe that the quantum mechanical description of photons is complete (i. e. that there exist no (local) variables – "hidden" or not – which could be used to predict Bob's measurement outcomes[5]), the E91 protocol performs a Bell experiment at the same time, which assures that there cannot exist (local) hidden variables.

In the E91 protocol, pairs of entangled photons are prepared, for example in the state $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. It does not matter whether these pairs are produced in Alice's or Bob's laboratory, or by a (potentially untrusted) source in between. One photon of each pair is sent to Alice, the other to Bob. For each photon, Alice and Bob choose one out of a set of three measurement directions at random, and measure the polarization of the photon in this direction (see Fig. 6.4). As in the BB84 protocol, Eve must not be able to predict the choice of the measurement directions. As soon as all pairs are sent to Alice and Bob and they acknowledge that they have performed the measurements, the information about the measurement directions is exchanged (through a public classical channel). Alice and Bob check for which pairs their respective measurement directions were the same; for all pairs where they have chosen different measurement directions, also the measurement outcomes are exchanged through the public classical channel. With these results, Alice and Bob check that the Bell inequalities [32,33] are violated. The measurement results for the pairs where they have chosen the same measurement direction are strictly anti-correlated, and can be used as a key.

[4] For a recent review of experiments testing Bell's inequalities, see e. g. [31]

[5] In experiments, classical information about the state which has been prepared might leak out of Alice's laboratory through different degrees of freedom, like the frequency of the photon, or the polarization of a second photon in a multi-photon pulse. This information could *in principle* be exploited by Eve without introducing noise. For the E91 protocol, this leakage problem does not exist, since such information does not exist until Alice and Bob perform their measurements

6.4.3 Security Proofs

As we have seen above, the quantum key distribution protocols allow for secure communication, as long as Alice and Bob are connected by a noiseless quantum channel. This is a remarkable result – however, it would be useless for all practical purposes, since all quantum channels are a source of noise. Since Alice and Bob trust only the equipment in their laboratories, they cannot be sure that the noise which they measure can be attributed to the channel. It is *in principle* impossible to distinguish between noise introduced by the quantum channel or by an eavesdropper. For this reason, the communication parties have to deal with the worst case scenario of an eavesdropper, who is present all the time and everywhere, except for the laboratories, which are secure by assumption. The eavesdropper might be hidden behind the noise of the quantum channel, and she might gain partial knowledge of the cryptographic key and, later, of the secret message.

The simplest way to deal with this situation would be to use a better quantum channel. In a practical setting, however, when Alice and Bob are connected by a given quantum channel (e. g. an optical fiber), this possibility is ruled out. In this situation, Alice and Bob can use *privacy amplification* methods, where a shorter and perfectly secure key is distilled out of a longer key, about which Eve might have had considerable knowledge. So-called "ultimate" or "unconditional" security proofs of quantum cryptography show that such protocols do exist.

The first of these proofs has been given by Mayers in 1996 [34] for the BB84 protocol. Shor and Preskill gave a physical interpretation of this proof, as they showed that it could *a posteriori* be understood as a restricted, albeit sufficient, form of quantum error correction and one-way entanglement purification.

A different approach has been taken by Deutsch *et al.* in 1996 [7]. They employ a two-way entanglement purification protocol (2-EPP, see Sect. 6.3) in order to distill almost pure EPR pairs out of many imperfect pairs. If the purified pairs are used for teleportation, the resulting quantum channel is perfectly secure: Since the EPR pairs are in a pure state, they cannot be entangled with any other quantum system. The eavesdropper is thus "factored out" in the total Hilbert space, which we write symbolically as

$$\rho_{\text{Alice,Bob,Eve}} \overset{\text{2-EPP}}{\longrightarrow} \left|\Psi^+\right\rangle_{AB}\!\left\langle\Psi^+\right| \otimes \rho_{\text{Eve}}.$$

As we have already seen in Sect. 6.3.2, in a realistic setting the purification protocol does not converge to perfect EPR pairs, but to some more or less mixed state in the Hilbert space of Alice's and Bob's qubits. But that means that the argument given above does no longer guarantee that Eve is factored out: *a priori*, there could exist residual entanglement with Eve.

6.5 Private Entanglement

In the last section we have seen that entanglement purification (using noisy apparatus) does not *per se* guarantee a provably private communication channel.

Nevertheless, in this section we will show that it suffices for the creation of "private entanglement", i. e. imperfect EPR pairs which are not entangled with an eavesdropper. Private entanglement can thereby serve as noisy but secure quantum channel.

The general idea is the following. Since Alice and Bob use noisy apparatuses for the entanglement distillation process, it is clear that the pairs become entangled with some degree of freedom of the laboratory. However, we will see that the total state of the laboratory and of the (distilled) pairs converges to a pure state, and then the same argument holds as in the case of noiseless entanglement purification: a quantum system in a pure state cannot be entangled with any other quantum system. In particular, Eve cannot be entangled with the distilled pairs. These pairs can then be used for secure, albeit noisy quantum teleportation.

In our analysis it is necessary to keep track of the state of the laboratory. This seems to be a difficult task, since the the details of the structure of the laboratory are unknown and complicated. For this reason one does usually not take care of these details, and describes the system of qubits on which the noisy apparatus acts as an open quantum system, with a master equation that describes their time evolution [35]. As an alternative, in the framework of quantum information theory, we use the concept of completely positive maps [36].

6.5.1 The Lab Demon

In this section, we give a simple model of a noisy laboratory, which allows us to keep track of its state in terms of classical variables.

As long as one cannot "look into" the device that introduced the noise, there is no way to distinguish it from a different device whose action is described by the same positive map. For this reason, our simple noise model is sufficient for the proof, and we need not delve into the complicated details of noisy quantum devices.

In order to keep the argument as transparent as possible, we will restrict our attention to the case that only Alice's laboratory is a source of noise; it would be easy to extend the argument to two noisy laboratories.

Let us assume that in Alice's lab there is a little demon. The lab demon kicks and shakes the qubits from time to time, and is thus a source of noise. However, there are no other sources of noise, and even the lab demon acts on the qubits in a very controlled way: let us assume that the demon has a random number generator that generates in each time step pairs of numbers $(\mu, \nu) \in \{0, 1, 2, 3\}^{\times 2}$, according to a given probability distribution $f_{\mu\nu}$ (which obeys the normalization condition $\sum_{\mu,\nu} f_{\mu\nu} = 1$). The lab demon then applies the (unitary) error operation $\sigma_\mu^{(a_1)} \sigma_\nu^{(a_2)}$ to the two qubits a_1 and a_2, on which Alice acts in the entanglement purification protocol (see Sect. 6.3.1). For $\mu \in \{1, 2, 3\}$, the operators $\sigma_\mu^{(a_i)}$ denote the Pauli matrices acting on qubit a_i, and $\sigma_0^{(a_i)} = \mathbb{1}^{(a_i)}$. In addition, the lab demon writes down which error operations he had applied to which qubits, since he will need this informtion later.

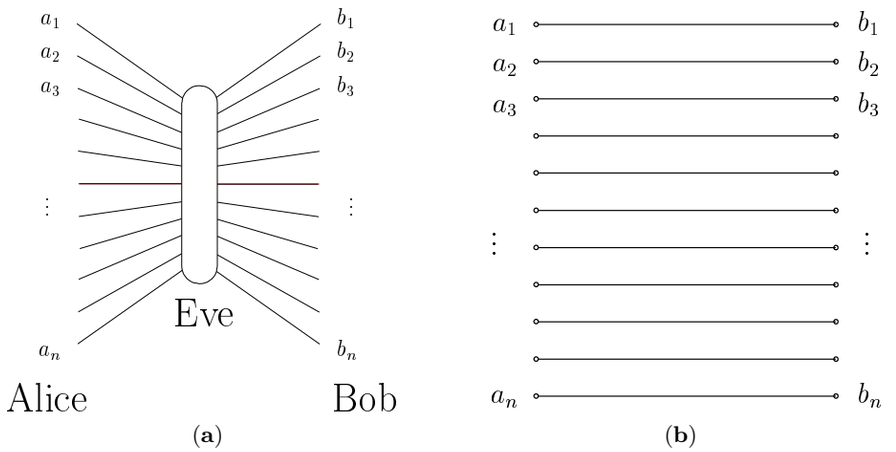


Fig. 6.5. (**a**) The state of N pairs which are distributed through the quantum channel is in the worst case scenario a general $2N$-qubit state, which might moreover be entangled with degrees of freedom under Eve's control. (**b**) After step 1, the state of the N pairs is a classically correlated ensemble of pure Bell states

Alice does not know which of the error operations have been applied to the qubits, and she describes the action of the demon by the average map

$$\rho_{a_1 a_2 \ldots} \to \sum_{\mu,\nu=0}^{3} f_{\mu\nu} \sigma_\mu^{(a_1)} \sigma_\nu^{(a_2)} \rho_{a_1 a_2 \ldots} \sigma_\mu^{(a_1)} \sigma_\nu^{(a_2)} \,. \tag{6.10}$$

The ellipsis $(\ldots)$ denotes other degrees of freedom, on which Alice's lab demon does not act (like Bob's qubits, or some quantum system in Eve's hands). We call the noise channel given by this equation the *correlated two qubit Pauli channel*. It includes, for special choices of the probability distribution $f_{\mu\nu}$, the one- and two-qubit depolarizing channel, and combinations thereof, which have been studied in the context of entanglement purification using imperfect apparatus in [29].

As mentioned above, we introduced the lab demon as a simplified noise model in order to keep track of the internal state of the lab. For that reason, we assume that the lab demon attaches an *error flag* λ to each qubit. The error flag will represent four different values, and it is convenient to divide it into two classical bits.

6.5.2 The State of the Qubits Distributed Through the Channel

In the worst case scenario, all pairs which are distributed between Alice and Bob have been prepared by Eve (see Fig. 6.5a). For that reason, the total state of all pairs is given by a general $2N$-qubit density operator, which can be written in

the form

$$\rho_{AB} = \sum_{\substack{\mu_1 \dots \mu_N \\ \mu_1' \dots \mu_N'}} \alpha_{\substack{\mu_1 \dots \mu_N \\ \mu_1' \dots \mu_N'}} |\mathcal{B}_{\mu_1}^{(a_1 b_1)} \dots \mathcal{B}_{\mu_N}^{(a_N b_N)}\rangle\langle\mathcal{B}_{\mu_1'}^{(a_1 b_1)} \dots \mathcal{B}_{\mu_N'}^{(a_N b_N)}| \,. \tag{6.11}$$

Here, $|\mathcal{B}_{\mu_j}^{(a_j b_j)}\rangle$, $\mu_j = 00, 01, 10, 11$ denote the 4 Bell states associated with the two qubits a_j and b_j and $j = 1, \dots, N$. Specifically, $|\mathcal{B}_{00}\rangle \equiv |\Phi^+\rangle$, $|\mathcal{B}_{01}\rangle \equiv |\Psi^+\rangle$, $|\mathcal{B}_{10}\rangle \equiv |\Phi^-\rangle$, $|\mathcal{B}_{11}\rangle \equiv |\Psi^-\rangle$.

In general, (6.11) will be an entangled $2N$-qubit state, which might moreover be entangled with additional quantum systems in Eve's hands. For the security analysis of the entanglement purification protocol, this state is too complicated and cannot be handled. It would be helpful if there was no entanglement between the different pairs. Fortunately, Alice and Bob can apply the following protocol to the pairs, in order to handle this situation:

Step 1: On each pair of particles (a_j, b_j), they apply randomly one of the four bi-lateral Pauli rotations $\sigma_k^{(a_j)} \otimes \sigma_k^{(b_j)}$, where k = 0,1,2,3.

Step 2: Alice and Bob randomly renumber the pairs, $(a_j, b_j) \to (a_{\pi(j)}, b_{\pi(j)})$ where $\pi \in S(N)$ is a permutation which has been chosen at random.

It is important to note that Alice and Bob deliberately discard the knowledge about which permutation and which of the Pauli rotations have been applied to the pairs. Obviously, they cannot force Eve to do the same thing. So Eve might have a better description of the state of the pairs than Alice and Bob. Thus the question remains whether this additional knowledge might help Eve. It is easy to see that this is not the case: Eve's description of the qubits has to be *statistically consistent* with the state which Alice and Bob or the lab demon assign to the two qubits. As we are going to show, at the end of the distillation process, the lab demon knows that the pairs are pure EPR pairs. Eve can thus not have more information about the pairs than the lab demon.[6]

After step 1, Alice's and Bob's knowledge about the state is summarized by the density operator

$$\tilde{\rho}_{AB} = \sum_{\mu_1 \dots \mu_N} p_{\mu_1 \dots \mu_N} |\mathcal{B}_{\mu_1}^{(a_1 b_1)} \dots \mathcal{B}_{\mu_N}^{(a_N b_N)}\rangle\langle\mathcal{B}_{\mu_1}^{(a_1 b_1)} \dots \mathcal{B}_{\mu_N}^{(a_N b_N)}| \tag{6.12}$$

which corresponds to a *classically correlated ensemble* of pure Bell states (see Fig. 6.5b). The fact that the pairs are classically correlated means that the order in which they appear in the numbered ensemble may have some pattern, which may have been imposed by Eve or by the channel itself. By applying step 2, the order of the pairs is "randomized"; this will prevent Eve from making use of any possibly pre-arranged order of the pairs, which Alice and Bob are meant to follow in the course of the distillation process: they simply ignore this order.

[6] In fact, Eve has *less* information than the lab demon, because she does not know the results of his random number generator.

The only correlation which remains is in the number of pairs which are in a specific Bell state. In the limit of large N, it is consistent for all relevant statistical predictions to describe the ensemble with the density operator

$$\tilde{\tilde{\rho}}_{AB} = \left(\sum_{\mu} p_{\mu} |\mathcal{B}_{\mu}\rangle\langle\mathcal{B}_{\mu}| \right)^{\otimes N} \equiv (\rho_{ab})^{\otimes N}. \tag{6.13}$$

For finite N, the form of the state after step 2 is more complicated; however, the subsequent arguments are also valid in that case.

At this stage, Alice and Bob have to check whether the pairs are "good enough" for the distillation process, i. e. they have to make sure that the fidelity F_0 of the pairs is above the purification/security threshold (which coincide for all practical purposes [37]). They can do this by local measurements on a fraction of the pairs and classical communication.

In order to separate conceptual from technical considerations and to obtain analytical results, we will first concentrate on a toy model where all the pairs are either in the state $|\Phi^+\rangle$ or $|\Psi^+\rangle$. In this case, we talk about *binary pairs*.

6.5.3 Binary Pairs

Let us assume that Alice and Bob initially share pairs in the state

$$\rho_{AB} = A \left|\Phi^+\right\rangle_{AB}\!\left\langle\Phi^+\right| + B \left|\Psi^+\right\rangle_{AB}\!\left\langle\Psi^+\right| \tag{6.14}$$

(binary pairs) with $A = 1 - B > 1/2$, and that the noise is of the form (6.10) with the restriction that the error operators consist only of the identity and spin flip operators:

$$\rho_{a_1 a_2 \ldots} \to \sum_{\mu,\nu=0}^{1} f_{\mu\nu} \sigma_{\mu}^{(a_1)} \sigma_{\nu}^{(a_2)} \rho_{a_1 a_2 \ldots} \sigma_{\mu}^{(a_1)} \sigma_{\nu}^{(a_2)} \,. \tag{6.15}$$

Equation (6.15) describes a *two-bit correlated spin-flip channel*. The indices 1 and 2 indicate the source and target bit of the bilateral CNOT operation, respectively. It is straightforward to show that, using this error model in the 2–EPP, binary pairs will be mapped onto binary pairs.

At the beginning of the distillation process, Alice and Bob share an ensemble of pairs described by (6.14). In this special case, one bit suffices for the error flag. At this stage, all of these bits are set to zero. This reflects the fact that the lab demon has the same *a priori* knowledge about the state of the ensemble as Alice and Bob.

In each purification step, two of the pairs are combined. The lab demon first simulates the noise channel (6.15) on each pair of pairs by using the random number generator as described. Whenever he applies a σ_x operation to a qubit, he inverts the error flag of the corresponding pair. Alice and Bob then apply the 2–EPP to each pair of pairs; if the measurement results in the last step of the protocol coincide, the source pair will be kept. Obviously, the error flag of that

remaining pair will also depend on the error flag of the the target pair, i. e. the error flag of the remaining pair is a function of the error flags of both "parent" pairs, which we call the *flag update function*. In the case of binary pairs, the flag update function maps two bits (the error flags of *both* parents) onto one bit. In total, there exist 16 different functions $g : \{0,1\}^2 \to \{0,1\}$. From these, the lab demon chooses the logical AND function as the flag update function, i. e. the error flag of the remaining pair is set to "1" if and only if both parent's error flags had the value "1".

After each purification step, the lab demon divides all pairs into two subensembles, according to the value of their error flags. By a straightforward calculation, we obtain for the coefficients A_i and B_i, which completely describe the state of the pairs in the subensemble with error flag i, the following recurrence relations:

$$\begin{aligned} A_0' =& \frac{1}{N}(f_{00}(A_0^2 + 2A_0A_1) + f_{11}(B_1^2 + 2B_0B_1) \\ &+ f_s(A_0B_1 + A_1B_1 + A_0B_0)) \,, \\ A_1' =& \frac{1}{N}\left(f_{00}A_1^2 + f_{11}B_0^2 + f_sA_1B_0\right) \,, \\ B_0' =& \frac{1}{N}(f_{00}(B_0^2 + 2B_0B_1) + f_{11}(A_1^2 + 2A_0A_1) \\ &+ f_s(B_0A_1 + B_1A_1 + B_0A_0)) \,, \\ B_1' =& \frac{1}{N}\left(f_{00}B_1^2 + f_{11}A_0^2 + f_sB_1A_0\right) \end{aligned} \tag{6.16}$$

with $N = (f_{00} + f_{11})((A_0 + A_1)^2 + (B_0 + B_1)^2) + 2f_s(A_0 + A_1)(B_0 + B_1)$ and $f_s = f_{01} + f_{10}$.

Since Alice and Bob do not know the values of the error flags, they describe the pairs in terms of $A = A_0 + A_1$ and $B = B_0 + B_1 = 1 - A$ as in (6.14). The fidelity F is thus given by $F = A$.

For the case of uncorrelated noise, the error operations are applied *independently* and with probability $f_\mu (\mu = 0, 1)$ to both qubits. This means that the probability distribution $f_{\mu\nu}$ factorizes into $f_{\mu\nu} = f_\mu f_\nu$. In this special case we obtain the following expression for fixpoints of this map:

$$\begin{aligned} &A_0^\infty = \frac{1}{2} \pm \frac{\sqrt{f_0 - 3/4}}{f_0 - 1} \quad \text{or} \quad A_0^\infty = \frac{1}{2}, \\ &A_1^\infty = 0, \quad B_0^\infty = 0, \quad B_1^\infty = 1 - A_0^\infty. \end{aligned} \tag{6.17}$$

The fixpoint of this map that is "relevant" for our discussion is defined by the plus sign in the expression for A_0^∞ above. It is not *per se* clear that a fixpoint is also an attractor. In fact, we find that (6.17) gives a non-trivial fixpoint of (6.16) for $f_0 \geq 3/4$, but this fixpoint is an attractor only for $f_0 > f_0^{\text{crit}} = 0.77184451$ [37].

To summarize, we can identify three regimes for values of the noise parameter (see Fig. 6.6): for a *high* noise level, when $f_0 < 3/4$, the protocol is not in the purification regime. From Alice's and Bob's point of view, the protocol converges

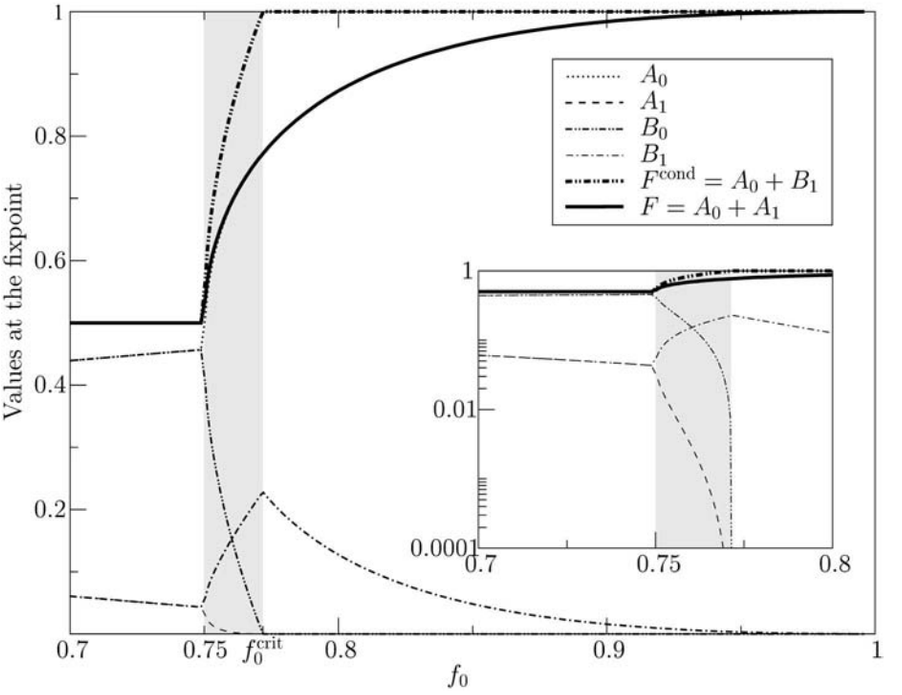


Fig. 6.6. The values of $A_0, A_1, B_0, B_1, F = A_0 + A_1, F^{\mathrm{cond}} = A_0 + B_1$ at the fixpoint as a function of the noise parameter f_0 [37]. For $f_0 < 0.75$, the values of A_1 and B_1, as well as the values of A_0 and B_0 are equal, and the respective lines lie on top of each other. One can clearly see that for $f_0 < 0.75$, the fidelity becomes $1/2$, and the pairs converge to the completely mixed state $1/2\left(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+|\right)$: the protocol is not in the purification regime. For $f_0 > 0.75$, the maximum achievable fidelity increases, and approaches unity for $f_0 \to 1$. This corresponds to the fact that the protocol is in the purification regime, and that it works better if the apparatus is more reliable. However, the fidelity is strictly smaller than unity for $f_0 < 1$. For the conditional fidelity $F^{\mathrm{cond}} = A_0 + B_1$, however, the situation is different: above the critical value f_0^{crit}, it becomes strictly equal to unity. Since F^{cond} is the fidelity or the pairs from the lab demon's point of view, any eavesdropper is factored out, and we call this regime the *security regime*. The regime, where the protocol purifies but does not provide secure EPR pairs is called *intermediate regime* (highlighted in grey). The inset shows the same graphs on a logarithmic scale. In this graph, one can see that the parameters A_1 and B_0 do not vanish only asymptotically, but become zero at $f_0 = f_0^{\mathrm{crit}}$

to the completely mixed binary state $1/2\,(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+|)$. For a *low* noise level, when $f_0 > f_0^{\mathrm{crit}}$, the protocol converges to a state where $A_0 + B_1 = 1$ and $A_1 = B_0 = 0$. This means that all pairs in the subensemble 0 are in the state $|\Phi^+\rangle$, and all pairs in the subensemble 1 are in the state $|\Psi^+\rangle$: From the lab demon's point of view, all pairs are in a pure state! For that reason, we will call this regime the *security regime* of the entanglement purification protocol.

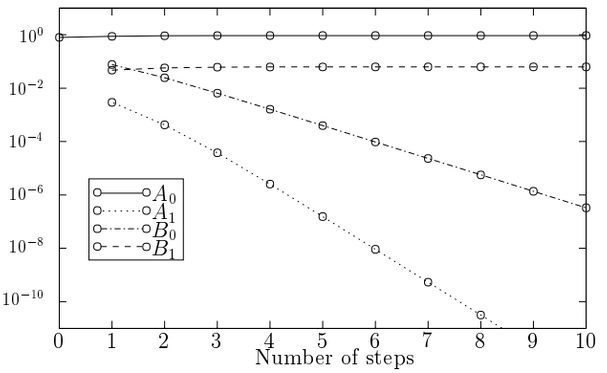


Fig. 6.7. The evolution of the four parameters A_0, A_1, B_0, and B_1 in the security regime. Note that both A_1 and B_0 decrease exponentially fast in the number of rounds. The initial fidelity was 80%, and the values of the noise parameters were $f_{00} = 0.8575$, $f_{01} = f_{10} = f_{11} = 0.0475$

For $3/4 \leq f_0 \leq f_0^{\mathrm{crit}}$, the protocol is in the *intermediate regime*. This regime is of no practical interest, since in this regime, the protocol converges very slowly. However, the mere existence of the intermediate regime is interesting, as it shows that purification and security are not trivially related to each other.

As we have already seen, the sum $A_0 + B_1$ is a measure for the purity of these pairs from the lab demon's point of view. We call this sum the *conditional fidelity* F^{cond}, since this is the fidelity which Alice and Bob would assign to the pairs if they knew the values of the error flags.

We have also evaluated (6.16) numerically in order to investigate correlated noise (see Fig. 6.7). Like in the case of uncorrelated noise, we found that the coefficients A_0 and B_1 reach, during the distillation process, some finite value, while the coefficients A_1 and B_0 decrease exponentially fast, whenever the noise level is moderate.

The distillation process which was described in Fig. 6.2b now looks as in Fig. 6.8, where the ensemble of pairs is now supplemented with an error flag for each pair. One can see that in the course of the distillation process, *strict* correlations are built up between the state of the pairs and the error flags λ_i. In the asymptotic limit, each flag identifies the state of the corresponding pair unambigously.

In other words, whenever the noise level is moderate, the conditional fidelity converges to unity: entanglement purification can be used to create private entanglement.

6.5.4 Bell-Diagonal Initial States

In the previous section we have considered the special case of binary pairs. For arbitrary Bell diagonal states (6.13) and for noise of the form (6.10), the results

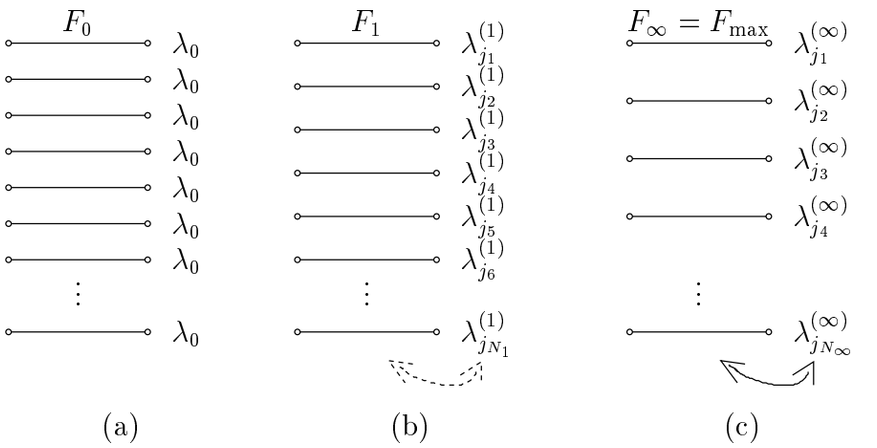


Fig. 6.8. In the course of the distillation process, *strict* correlations are built up between the state of the pairs and the error flags λ_i: (**a**) Initially, all error flags are set to zero, and there exist no correlations between the states of the pairs and the error flags; (**b**) after the first distillation round, there exist weak correlations. (**c**) Finally, in the asymptotic limit, the error flags are strictly correlated with the states of the pairs, and each flag identifies the state of the corresponding pair unambigously

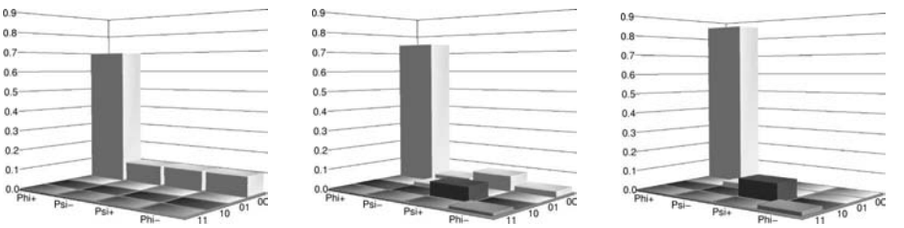


Fig. 6.9. Typical evolution of the 16 parameters $A^{ij}, B^{ij}, C^{ij}, D^{ij}$ with $i, j \in \{0, 1\}$ under the purification protocol. As in (6.6), the coefficients A, B, C, and D correspond to the four Bell states which are indicated by "Phi+", "Psi-", "Psi+", and "Phi-" on one axis. The other axis shows the error flag $\lambda \in \{00, 01, 10, 11\}$. As one can see, only the diagoal elements survive, which means that the error flag identifies the states of all pair unambigously. The noise parameters in this plot are $f_{00} = 0.83981, f_{0j} = f_{i0} = 0.021131$ and $f_{ij} = 0.003712$ for $i, j \in \{1, 2, 3\}$

are quite similar. However, the most important difference is that in the general case the intermediate regime is much smaller than in the case of binary pairs.

As already mentioned, in general the error flag consists of two classical bits. This means that the the lab demon has to use a more complicated flag update function than in the case of binary pairs. In this case, the flag update function has been found by looking at how errors are propagated during the course of the distillation process. The details of this calculation can be found in [37].

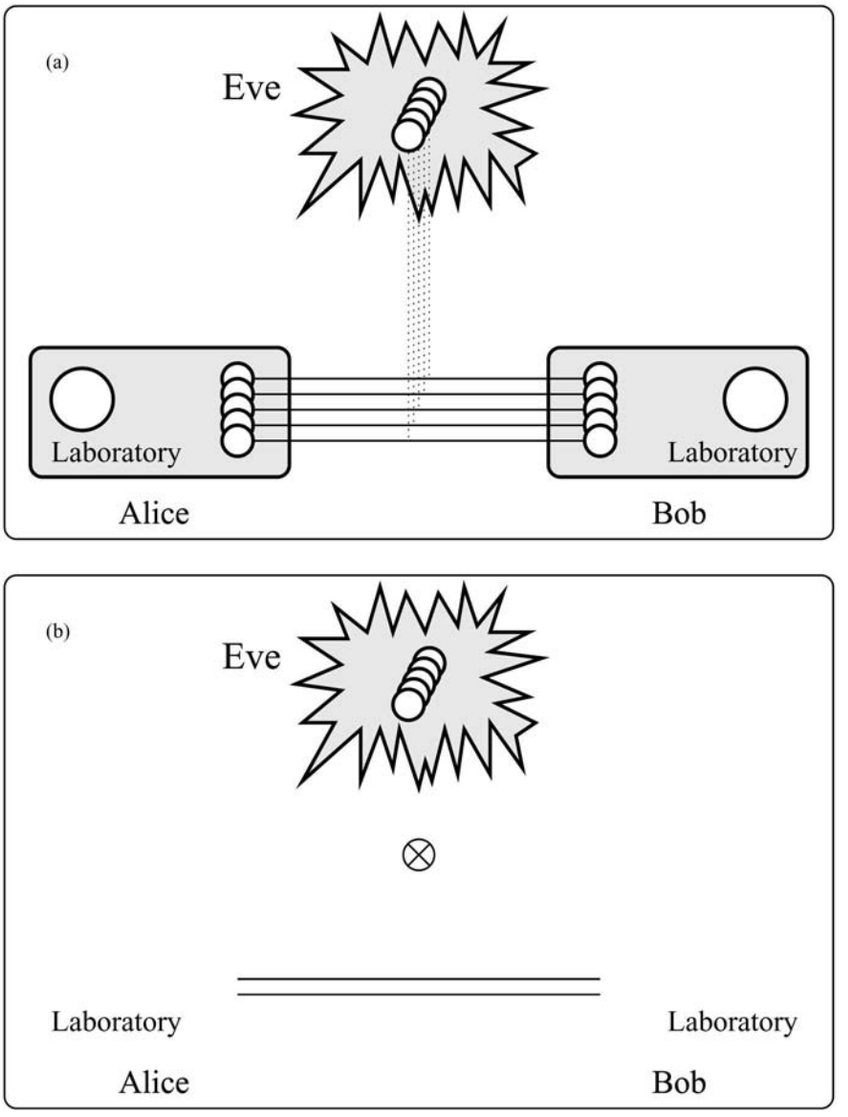


Fig. 6.10. The entanglement distribution process *redistributes* entanglement: at the beginning, the EPR pairs are entangled with the communication channel, or maybe even with an eavesdropper. In the end, however, the remaining EPR pairs are only entangled with the laboratories, and the eavesdropper is *factored out*

Since the error flag represents four different values, the lab demon divides all pairs into four subensembles, according to the value of their error flag λ. In each of the subensembles the pairs are described by a Bell diagonal density operator,

like in (6.6), which now depends on the subensemble. That means, in order to completely specify the state of all four subensembles, we need 16 real numbers $A^{ij}, B^{ij}, C^{ij}, D^{ij}$ with $i, j \in \{0, 1\}$.

Figure 6.9 shows how these 16 parameters evolve under the action of the distillation process: If the protocol is in the security regime, only the "diagonal" elements survive and are identified by unambigously by the corresponding error flag. Again, this means that from the lab demons point of view, all pairs are in a pure state.

To summarize, we have found that in the entanglement distillation process, entanglement is redistributed in the following sense (see Fig. 6.10): in the beginning, there exists (unwanted) entanglement between the EPR pairs and the quantum channel (Eve). The entanglement distillation process is not capable of creating perfect EPR pairs, since the pairs become entangled with the laboratories, due to uncontrolled interactions. Despite this fact, Eve is factored out, and all entanglement between her and the EPR pairs is lost: Alice and Bob succeeded in creating private entanglement and have thus a private, albeit noisy, quantum channel.

Acknowledgement

We would like to thank the organizers of the CohEvol workshop and seminar, who have managed to bring together students and researchers from different fields of physics in an inspiring atmosphere. This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) within the Schwerpunktsprogramm QIV.

References

1. P. W. Shor: SIAM J. Sci. Statist. Comput. **26**,1484 (1997)
2. P. W. Shor: Phys. Rev. A **52**, R2493 (1995)
3. A. R. Calderbank, P. W. Shor: Phys. Rev. A **54** 1098 (1996)
4. A. Steane: Phys. Rev. Lett. **77**, 793 (1996)
5. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, W. K. Wootters: Phys. Rev. Lett. **76**, 722 (1996)
6. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters: Phys. Rev. A **54**, 3824 (1996)
7. D. Deutsch, A. Ekert, R. Josza, C. Macchiavello, S. Popescu, A. Sanpera: Phys. Rev. Lett. **77**, 2818 (1996)
8. C. H. Bennett, G. Brassard, G. Crepeau, R. Josza, A. Peres, W. Wootters: Phys. Rev. Lett. **70**, 1895 (1993)
9. A. Einstein, B. Podolsky, N. Rosen: Phys. Rev. **47**, 777 (1935)
10. C. H. Bennett G. Brassard: 'Quantum Cryptography: Public-key distribution by coin tossing'. In: *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1985)
11. A. K. Ekert: Phys. Rev. Lett. **67**, 661 (1991)
12. D. Bouwmeester, A. Ekert, A. Zeilinger (eds.): *The Physics of Quantum Information* (Springer, Berlin, 2000)

13. H. Aschauer, H.-J. Briegel: Phys. Rev. Lett. **88**, 047902 (2002)
14. D. M. Greenberger, M. Horne, A. Zeilinger: 'Going beyond Bell's theorem'. In: *Bell's theorem, quantum theory, and conceptions of the universe*, ed. by M. Kafatos (Kluwer, Dortrecht, 1989)
15. N. D. Mermin: Am. J. Phys. **58**, 731 (1990)
16. D. P. DiVincenzo, A. Peres: Phys. Rev. A **55**, 4089 (1997)
17. R. Laflamme, C. Miquel, J. P. Paz, W. H. Zurek: Phys. Rev. Lett. **77**, 198 (1996)
18. E. Knill, R. Laflamme: Phys. Rev. A **55**, 900 (1997)
19. A. Steane: Proc. R. Soc. Lond. A **452**, 2551 (1996)
20. A. Steane: 'Quantum error correction'. In: *Introduction to quantum computation and information*, ed. by H. K. Lo, S. Popescu, T. Spiller (World Scientific, Singapore, 1998)
21. A. Steane: 'General theory of quantum error correction and fault tolerance'. In: *The Physics of Quantum Information*, ed. by D. Bouwmeester, A. Ekert, A. Zeilinger (Springer, Berlin, 2000)
22. D. Gottesman: *An introduction to quantum error correction*. Preprint: arXiv:quant-ph/0004072 (2000)
23. D. Gottesman: Phys. Rev. A **54**, 1862 (1996)
24. J. Preskill: 'Fault-tolerant quantum computation'. In: *Introduction to quantum computation and information*, ed. by H. K. Lo, S. Popescu, T. Spiller (World Scientific, Singapore, 1998)
25. R. F. Werner: Phys. Rev. A **40**, 4277 (1989)
26. H. J. Briegel: 'Principles of entanglement purification'. In: *The Physics of Quantum Information*, ed. by D. Bouwmeester, A. Ekert, A. Zeilinger (Springer, Berlin, 2000)
27. C. Macchiavello: Phys. Lett. A **246**, 385–388 (1998)
28. G. Giedke, H.-J. Briegel, J. I. Cirac, P. Zoller: Phys. Rev. A **59**, 2641 (1999)
29. H.-J. Briegel, W. Dür, J. I. Cirac, P. Zoller: Phys. Rev. Lett. **81**, 5932 (1998)
30. A. Peres: *Quantum Theory: Concepts and Methods* (Kluwer, New York, 1993)
31. A. Zeilinger: Rev. Mod. Phys. **71**, 288 (1999)
32. J. S. Bell: Physics **1**, 195 (1964). Reprinted in: *Speakable and unspeakable in quantum mechanics* (Cambidge University Press, 1987)
33. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt: Phys. Rev. Lett. **23**, 880 (1969)
34. D. Mayers: In *Advances in Cryptology – Proceedings of Crypto '96*, pp. 343–357 (Springer-Verlag, New York, 1996). See also preprint arXiv:quant-ph/9802025
35. C. W. Gardiner: *Quantum Noise* (Springer Verlag, Berlin, 1991)
36. K. Kraus: *States, Effects, and Operations* (Lecture Notes in Physics, vol. 190) (Springer Verlag, Berlin, 1983)
37. H. Aschauer, H. J. Briegel: to be published in Phys. Rev. A (2002). Preprint arXiv:quant-ph/0111066.

7 How to Correct Small Quantum Errors

Michael Keyl and Reinhard F. Werner

Institut für Mathematische Physik, TU Braunschweig,
Mendelssohnstr. 3, D-38106 Braunschweig, Germany

7.1 Introduction

Controling decoherence is one of the key problems for making quantum information processing and quantum computation work. From the outset, when Peter Shor announced his algorithm [20,21], many physicists felt that somewhere there would be a price to pay for the miraculous exponential speedup. For example, if the algorithm would require exponentially good adherence to specifications for the quantum circuitry and exponentially low noise levels, it would have been totally useless. Indeed it is far from easy to show that it does not make such requirements.

In this article we look at the simpler, but equally fundamental problem of quantum information transmission or storage. Is it possible to encode the quantum data in such a way that even after some degradation they can largely be restored by a suitable decoding operation? Assuming that the degrading decoherence effects are small to begin with, can restoration be made nearly perfect?

For classical information it is very simple to do this, namely by redundant coding. If we want to send one bit through a noisy channel, we can reduce errors by sending it three times and deciding by majority vote which value we take at the output. Clearly, if errors have a small probability ε for a single channel, they will have order ε^2 for the triple channel, because we go wrong only when two independent errors occur. Unfortunately, such a scheme cannot work in the quantum case because it involves a copying operation, which is forbidden by the No-Cloning Theorem [25]. So we have to look for subtler ways of distributing quantum information among several systems and thereby reducing the probability of errors. Indeed such schemes exist [4,22] and are the subject of the exciting new field of quantum error correcting codes (see also the discussion in Sect. 6.2).

The efficiency of such a scheme is measured by two parameters, namely how many uses of the noisy channel are required, and the error level after correction. The above simple classical scheme can be iterated to get the errors for a single bit down to ε^{2^n} with 3^n parallel uses of the channel. This is a large overhead to correct a single bit. Better procedures work classically by coding several bits at a time, and one can manage to make errors as small as desired with only a finite overhead per bit. The minimal required overhead (or rather its inverse) is, in fact, the central quantity of the coding theory [19] for noisy channels: one defines the *capacity* of a channel as the number of bit transmissions per use of

A. Buchleitner and K. Hornberger (Eds.): LNP 611, pp. 263–286, 2002.

the channel, in an optimal coding scheme for messages of length $L \to \infty$ with the property that the error probability goes to zero in this limit.

It is not a priori clear that the notion of channel capacity makes sense for quantum information, i.e. that the capacity of a channel which produces only small errors is nonzero and close to that of the ideal (errorless) channel. This is indeed not even evident from most existing presentations of the theory of quantum error correcting codes. Papers which address this problem, at least for special cases like depolarizing channels are [5,7] and [16, Sec 7.16.2], while the general case is treated more recently in [8,13]. The purpose of this paper is less the presentation of new results but to show in an elementary and self-contained way that small quantum errors can be corrected with an asymptotically small effort.

The paper is organized as follows. We first review the basic notions concerning quantum channels (Sect. 7.2), and give an abstract definition of the capacity together with some elementary properties (Sect. 7.3). Then we discuss the theory of error correcting codes (Sect. 7.4) and a particular scheme to construct such codes which is based on graph theory (Sect. 7.5). In Sect. 7.6 and 7.7 we apply this scheme to channel capacities and finally we draw our conclusions in Sect. 7.8.

7.2 Quantum Channels

According to the rules of quantum mechanics, every kind of quantum systems is associated with a Hilbert space $\mathcal{H}$, which for the purpose of this article we can take as finite dimensional. Since even elementary particles require infinite dimensional Hilbert spaces, this restriction means that we are usually only trying to coherently manipulate a small part of the system. The simplest quantum system has a two dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$, and is called a *qubit*, for 'quantum bit'. The observables and states of the system are given by hermitian linear operators on $\mathcal{H}$. We write this as $A \in \mathcal{B}(\mathcal{H})$ for observables and as $\rho \in \mathcal{B}_*(\mathcal{H})$ for states (=density operators). This notation reflects distinctions[1], which become relevant only for infinite dimensional $\mathcal{H}$, but helps even in the finite dimensional case to distinguish Heisenberg and Schrödinger picture.

A *quantum channel*, which transforms input systems described by a Hilbert space $\mathcal{H}_1$ into output systems described by a (possibly different) Hilbert space $\mathcal{H}_2$ is represented mathematically by a *completely positive [24,15], unit preserving map* $T : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$. Equivalently [12,23], each T can be written in the form

$$T(A) = \sum_{j=1}^{n} F_j^\dagger A F_j, \tag{7.1}$$

[1] "$\mathcal{B}$" stands for *bounded* operators, i.e., operators such that $\|A\psi\| \leq c\|\psi\|$ for a suitable constant $c \equiv \|A\|$. On the other hand, $\mathcal{B}_*(\mathcal{H})$ is the trace class, consisting of all operators for which $\|\rho\|_1 = \operatorname{tr}\sqrt{\rho^\dagger \rho}$ is finite. For extensive background material see [17].

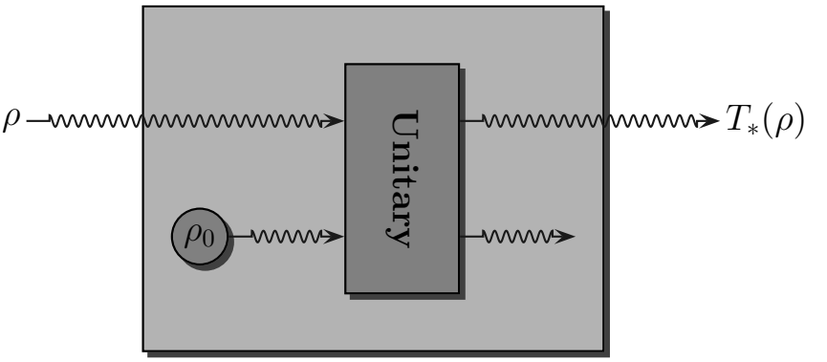


Fig. 7.1. Noisy channel T, represented by coupling to the environment (cf. (7.2))

where the F_j are operators $\mathcal{H}_2 \to \mathcal{H}_1$, which are called *Kraus operators*, and satisfy the normalization condition $\sum_{j=1}^{n} F_j^\dagger F_j = \mathbb{1}$.

The physical interpretation of T is the following. The expectation value of an A measurement ($A \in \mathcal{B}(\mathcal{H}_2)$) at the output side of the channel, on a system which is initially prepared in the state $\rho \in \mathcal{B}_*(\mathcal{H}_1)$ is given in terms of T by $\mathrm{tr}[\rho T(A)]$. Alternatively we can introduce the map $T_* : \mathcal{B}_*(\mathcal{H}_1) \to \mathcal{B}_*(\mathcal{H}_2)$ which is *dual* to T, i.e. $\mathrm{tr}[T_*(\rho)A] = \mathrm{tr}[\rho T(A)]$. It is uniquely determined by T (and vice versa) and we can say that T_* represents the channel in the *Schrödinger picture*, while T provides the *Heisenberg picture* representation.

Let us consider now the special case that $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. For example, T describes the transmission of photons through an optical fiber or the storage in some sort of quantum memory. Here we would most like to have channels which do not affect the information at all, i.e. $T = \mathrm{Id}$, the identity map on $\mathcal{B}(\mathcal{H})$. We will call this case the *ideal channel*. In real situations, however, interaction with the environment, i.e. additional, unobserved degrees of freedom, can not be avoided. The general structure of such a *noisy channel* is given by

$$\rho \mapsto T_*(\rho) = \mathrm{tr}_{\mathcal{K}}\big(U(\rho \otimes \rho_0)U^\dagger\big). \tag{7.2}$$

where $U : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$ is a unitary operator describing the common evolution of the system (Hilbert space $\mathcal{H}$) and the environment (Hilbert space $\mathcal{K}$) and $\rho_0 \in \mathcal{S}(\mathcal{K})$ is the initial state of the environment (cf. Fig. 7.1). Note that each T can be represented in this way (this is an easy consequence of the Stinespring theorem[23]), however there are in general many possible choices for such an "ancilla representation".

7.3 Channel Capacities

As we have already pointed out in the introduction, the *capacity* of a quantum channel is, roughly speaking, the number of qubits transmitted per channel usage. In this section we will come to a more precise description.

7.3.1 The cb-Norm

As a first step we need a measure for the difference between a noisy channel $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ and its ideal counterpart. There are several mathematical ways of expressing this, which turn out to be equivalent for our purpose. We find it most convenient to take a norm difference, i.e., to consider $\|T - \mathrm{Id}\|_{\mathrm{cb}}$ as a quantitative description of the noise level in T, where $\|\cdot\|_{\mathrm{cb}}$ denotes a certain norm, called the *norm of complete boundedness* ("cb-norm" for short). Its physical meaning is that of the largest difference between probabilities measured in two experimental setups, differing only by the substitution of T by Id. Since this setup may involve further subsystems, and the measurement and preparation may be entangled with the systems under consideration, we have to take into account such additional systems in the definition of the norm. For a general linear operator $T : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$ we set

$$\|T\|_{\mathrm{cb}} = \sup\left\{ \|(T \otimes \mathrm{Id}_n)(A)\| \;\middle|\; n \in \mathbb{N}; A \in \mathcal{B}(\mathcal{H}_2 \otimes \mathbb{C}^n); \|A\| \leq 1 \right\}, \tag{7.3}$$

where Id_n denotes the identity on the $n \times n$-matrices $\mathcal{B}(\mathbb{C}^n)$. The cb-norm improves the sometimes annoying property of the usual operator norm that quantities like $\|T \otimes \mathrm{Id}_n\|$ may increase with the dimension n. A particular example for a map with such a behavior is the transposition, e.g., the matrix transposition Θ_d on the $d \times d$-matrices $\mathcal{B}(\mathbb{C}^d)$. Then

$$\|\Theta_d \otimes \mathrm{Id}_n\| = \sup_{\|A\| \leq 1} \|(\Theta_d \otimes \mathrm{Id}_n)(A)\| \geq \min\{d, n\} . \tag{7.4}$$

In fact, equality holds here [15], but we need only the inequality as stated. To show it, we take $A = \sum_{i,j=1}^{m} |ij\rangle\langle ji|$ with $m = \min\{d, n\}$. Then $\|A\| = 1$, but $(\Theta_d \otimes \mathrm{Id}_n)A = \sum_{i,j=1}^{m} |jj\rangle\langle ii|$, which is m times the one-dimensional projection onto the normalized vector $m^{-1/2} \sum_i |ii\rangle$. Hence the supremum in (7.4) is at least m. In particular, we get $\|\Theta_d\|_{\mathrm{cb}} \geq d$. This also covers the case $d = \infty$, so we see that the sequence of suprema (7.3) with n fixed is unbounded. Conversely, operators for which it stays bounded, i.e., those with $\|T\|_{\mathrm{cb}} < \infty$ are called completely bounded, from where the cb-norm has its name. Of course, in a finite dimensional setup each linear map is automatically completely bounded. The cb-norm has some nice features which we will use frequently. This includes its multiplicativity $\|T_1 \otimes T_2\|_{\mathrm{cb}} = \|T_1\|_{\mathrm{cb}} \, \|T_2\|_{\mathrm{cb}}$ and the fact that $\|T\|_{\mathrm{cb}} = 1$ for every channel. For more properties of the cb-norm we refer to [15].

7.3.2 Achievable Rates and Capacity

How can we reduce the error level $\|T - \mathrm{Id}\|_{\mathrm{cb}}$? As an example, consider a small unitary rotation, i.e., $T(X) = U^\dagger X U$, with $\|T - \mathrm{Id}\|_{\mathrm{cb}} \leq 2\|U - \mathbb{1}\|$ small. Then if we know U, it is easy to correct T by the inverse rotation, either before T, as an "encoding" (E), or afterwards, as a "decoding" (D) operation. More generally, we may use both, i.e., we are trying to make the combination $ETD \approx$

Id, by careful choice of the channels E and D. Note that in this way we may look at channels T which have different input and output spaces, and hence cannot be compared directly with the ideal channel on any system. For such channels there is no intrinsic way of defining "errors" as deviations from a desired standard. Moreover, we are free to choose the Hilbert space $\mathcal{H}_0$ such that $ETD : \mathcal{B}(\mathcal{H}_0) \to \mathcal{B}(\mathcal{H}_0)$. For the product ETD to be defined, it is then necessary that $D : \mathcal{B}(\mathcal{H}_0) \to \mathcal{B}(\mathcal{H}_2)$ and $E : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_0)$. The best error level we can achieve deserves its own notation. We define

$$\Delta(T, M) = \inf_{E,D} \|ETD - \mathrm{Id}_M\|_{\mathrm{cb}} , \tag{7.5}$$

where the infimum is taken over all encodings E and decodings D and M is the dimension of the space $\mathcal{H}_0$. Now for longer messages, e.g., a message of m qubits (so that $M = 2^m$) we need to use the channel more often. In the language of classical information theory, we are using longer code words, say of length n. The error for coding m qubits through n uses of the channel T is then $\Delta(T^{\otimes n}, 2^m)$. Can we make this small while retaining a good rate m/n of bits per channel? Clearly there will be a trade-off between rate and errors, which is the basis of the following Definition. The notation $\lfloor x \rfloor$, read "floor x", denotes the largest integer $\leq x$.

Definition 1 *$c \geq 0$ is called* achievable rate *for T, if*

$$\lim_{n\to\infty} \Delta(T^{\otimes n}, \lfloor 2^{cn} \rfloor) = 0 . \tag{7.6}$$

The supremum of all achievable rates is called the quantum-capacity *of T and is denoted by $Q(T)$.*

Because $c = 0$ is always an achievable rate we have $Q(T) \geq 0$. On the other hand, if every $c > 0$ is achievable we write $Q(T) = \infty$.

Often a coding scheme construction does not work for arbitrary integers, but only for specific values of n, or the dimension of the coding space. However, this is no serious restriction, as the following Lemma shows.

Lemma 2 *Let $(n_\alpha)_{\alpha\in\mathbb{N}}$ be a strictly increasing sequence of integers such that $\lim_\alpha n_{\alpha+1}/n_\alpha = 1$. Suppose M_α are integers such that $\lim_\alpha \Delta(T^{\otimes n_\alpha}, M_\alpha) = 0$. Then any*

$$c < \liminf_\alpha \frac{\log_2 M_\alpha}{n_\alpha} \tag{7.7}$$

is an admissible rate. Moreover, if the errors decrease exponentially, in the sense that $\Delta(T^{\otimes n_\alpha}, M_\alpha) \leq \mu e^{-\lambda n_\alpha}$ $(\mu, \lambda \geq 0)$, then they decrease exponentially for all n with rate

$$\liminf_{n\to\infty} \frac{-1}{n} \ln \Delta(T^{\otimes n}, \lfloor 2^{cn} \rfloor) \geq \lambda. \tag{7.8}$$

Proof. Let us introduce the notation $c_+ = \liminf_\alpha (\log_2 M_\alpha)/n_\alpha$, so $c < c_+$. We pick $\eta > 0$ such that $(1 + \eta)c < c_+$. Then for sufficiently large $\alpha \geq \alpha_0$ we

have $(n_{\alpha+1}/n_\alpha) \leq (1+\eta)$, and $(\log_2 M_\alpha/n_\alpha) \geq (1+\eta)c$. Now let $n \geq n_{\alpha_0}$, and consider the unique index α such that $n_\alpha \leq n \leq n_{\alpha+1}$. Then $n \leq (1+\eta)n_\alpha$ and

$$\lfloor 2^{cn} \rfloor \leq 2^{cn} \leq 2^{c(1+\eta)n_\alpha} \leq M_\alpha. \tag{7.9}$$

Clearly, $\Delta(T^{\otimes n}, M)$ decreases as n increases, because good coding becomes easier if we have more parallel channels, and increases with M, because if a coding scheme works for an input Hilbert space $\mathcal{H}_0$, it also works at least as well for states supported on a lower dimensional subspace. Hence $\Delta(T^{\otimes n}, \lfloor 2^{cn} \rfloor) \leq \Delta(T^{\otimes n_\alpha}, M_\alpha) \to 0$. It follows that c is an admissible rate.

With the exponential bound on Δ we find similarly that

$$\Delta(T^{\otimes n}, \lfloor 2^{cn} \rfloor) \leq \mu\, e^{-\lambda n_\alpha} \leq \mu\, e^{-\lambda/(1+\eta)n}\,, \tag{7.10}$$

so that the $\liminf$ in (7.8) is $\geq \lambda/(1+\eta)$. Since η was arbitrary, we get the desired result. □

7.3.3 Elementary Properties

To determine $Q(T)$ in terms of Definition 1 is fairly difficult, because optimization problems in spaces of exponentially fast growing dimensions are involved. In particular, this renders each direct numerical approach practically impossible. In the classical situation, i.e. if we transfer classical information through a classical channel Φ, we can define a capacity quantity $C(\Phi)$ in the same way as above. An explicit calculation of $C(\Phi)$, however, can be reduced, according to Shannons "noisy channel coding theorem" [19], to an optimization problem over a low dimensional space, which does not involve the limit of inifinitely many parallel channels. A similar coding theorem for the quantum case is not yet known – this is the biggest open problem concerning channel capacities.

Nevertheless, there are some special cases in which the capacity can be computed explicity. The most relevant example is the ideal channel Id_d on $\mathcal{B}(\mathbb{C}^d)$. If $d^n \geq M$ we can embed $\mathbb{C}^M$ into $(\mathbb{C}^d)^{\otimes n}$, hence $\Delta(\mathrm{Id}_d^{\otimes n}, M) = 0$ and we see that the rate $\log_2(d)$ can be achieved. Intuitively we expect that this is the best what can be done, because it is impossible to embed a high- into a low-dimensional space. This intuition is in fact correct, i.e. we have $Q(\mathrm{Id}_d) = \log_2(d)$ for the ideal channel. A precise proof of this statement is, however, not as easy as it looks and we skip the details here. Maybe the easiest approach is to use the quantity $\log_2(\|\Theta T\|_{\mathrm{cb}})$ (where Θ denotes the transposition), which is an upper bound on $Q(T)$ (cf. [10] or [24]). The same idea can be used to show that the quantum capacity of a classical channel, or more generally a channel T which uses classical information at an intermediate step, is zero. This is a reformulation of the "no classical teleportation theorem" (cf. again [24]).

Another useful relation concerns the concatenation of two general channels T_1 and T_2: We transmit quantum information first through T_1 and then through T_2. It is reasonable to assume that the capacity of the composition T_2T_1 can not

be bigger than the capacity of the channel with the smallest bandwidth. This conjecture is indeed true and known as the "*Bottleneck inequality*":

$$Q(T_2T_1) \leq \min\{Q(T_1), Q(T_2)\} \; . \tag{7.11}$$

Alternatively we can use the two channels in parallel, i.e. we consider the tensor product $T_1 \otimes T_2$. In this case the capacity of the resulting channel is at least as big as the sum of $Q(T_1)$ and $Q(T_2)$, i.e. Q is *superadditive*:

$$Q(T_1 \otimes T_2) \geq Q(T_1) + Q(T_2) \tag{7.12}$$

(cf. [10] for a proof of both statements). To decide whether Q is even additive, i.e. whether equality holds in (7.12), is another big open question about channel capacities.

7.4 Quantum Error Correction

The definition of capacity requires that we correct errors in a collection of n parallel channels $T^{\otimes n}$. Here the tensor product means that successive uses of the channel are independent. For example, the physical system used as a carrier is freshly prepared every time we use the channel. This independence is important for error correcting schemes, because it prevents errors happening on different channels to "conspire".

Suggestive as it may be, quantum mechanics cautions us to be very careful with this sort of language: just as we cannot assign trajectories to quantum systems, it is problematic to speak about errors 'happening' in one channel, in a situation where we must expect different classical pictures to 'occur' in quantum mechanical superposition. This is to be kept in mind, when we now describe the theory of quantum error correcting codes in the sense of Knill and Laflamme [11], which is very much based on a classification of errors according to the place where they occur. For example, the coding/decoding pair E, D will typically have the property that $E(T_1 \otimes T_2 \otimes \cdots \otimes T_n)D = \mathrm{Id}$, whenever the number of positions at which $T_i \neq \mathrm{Id}$, i.e., the number of errors, is small (cf. Fig. 7.2).

In our presentation of the Knill–Laflamme Theory, we start from the error corrector's dream, namely the situation in which *all the errors happen in another part of the system*, where we do not keep any of the precious quantum information. This will help us to characterize the structure of the kind of errors which such a scheme may tolerate, or 'correct'. Of course, the dream is just a dream for the situation we are interested in: several parallel channels, each of which may be affected by errors. But the splitting of the system into subsystems, mathematically the decomposition of the Hilbert space of the total system into a tensor product is something we may change by a suitable unitary transformation. This is then precisely the role of the encoding and decoding operations. The Knill–Laflamme theory is precisely the description of the situation where such a unitary, and hence a coding/decoding scheme exists. Constructing such schemes, however, is another matter, to which we will turn in the next section.

7.4.1 An Error Corrector's Dream

So consider a system split into $\mathcal{H} = \mathcal{H}_g \otimes \mathcal{H}_b$, where the indices g and b stand for 'good' and 'bad'. We prepare the system in a state $\rho \otimes |\Omega\rangle\langle\Omega|$, where ρ is the quantum state we want to protect. Now come the errors in the form of a completely positive map $T(A) = \sum_i F_i^\dagger A F_i$. Then according to the error corrector's dream, we would just have to discard the bad system, and get the same state ρ as before.

The hardest demands for realizing this come from pure states $\rho = |\phi\rangle\langle\phi|$, because the only way that the restriction to the good system can again be $|\phi\rangle\langle\phi|$ is that the state after errors factorizes, i.e.

$$T_*(|\phi \otimes \Omega\rangle\langle\phi \otimes \Omega|) = \sum_i |F_i(\phi \otimes \Omega)\rangle\langle F_i(\phi \otimes \Omega)| = |\phi\rangle\langle\phi| \otimes \sigma \ . \tag{7.13}$$

This requires that

$$F_i(\phi \otimes \Omega) = \phi \otimes \Phi_i \ , \tag{7.14}$$

where $\Phi_i \in \mathcal{H}_b$ is some vector, which must be independent of ϕ if such an equation is to hold for all $\phi \in \mathcal{H}_g$. Conversely, condition (7.14) implies (7.13) for every pure state $|\phi\rangle\langle\phi|$ and, by convex combination, for every state ρ.

Two remarks are in order. Firstly, we have *not* required that $F_i = \mathbb{1} \otimes F_i'$. This would be equivalent to demanding that this scheme works with every Ω, or indeed with every (possibly mixed) initial state of the bad system. This would be much too strong for a useful theory of codes. So later on we must insist on a proper initialization of the bad subsystem by a suitable encoding. Secondly, if we have the condition (7.14) for the Kraus operators of some channel T, then it also holds for all channels whose Kraus operators can be written as linear combinations of the F_i. In other words, the "set of correctible errors" is naturally identified with the vector space of operators F such that there is a vector $\Phi \in \mathcal{H}_b$ with $F(\phi \otimes \Omega) = \phi \otimes \Phi$ for all $\phi \in \mathcal{H}_g$. This space will be called the *maximal*

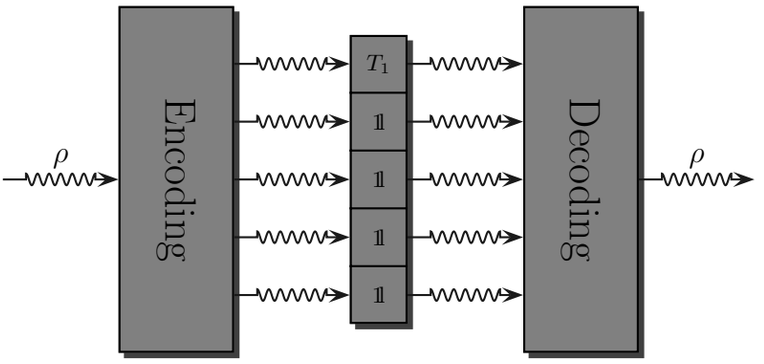


Fig. 7.2. A "five bit quantum code", encoding one qubit into five and correcting one error. An arbitrary noisy channel inserted for T_1 or, similarly, at any one of the other four locations has no effect on the final output

error space of the coding scheme, and will be denoted by $\mathcal{E}_{\max}$. Usually, a code is designed for a given error space $\mathcal{E}$. Then the statement that these given errors are corrected simply becomes $\mathcal{E} \subset \mathcal{E}_{\max}$. The key observation, however, is that the space of errors is a vector space in a natural way, i.e., if we can correct two types of errors, then we can also correct their *superposition*.

7.4.2 Realizing the Dream by Unitary Transformation

Let us now consider the situation in which we want to send states of a small system with Hilbert space $\mathcal{H}_1$ through a channel $T : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_2)$. The Kraus operators of T lie in an error space $\mathcal{E} \subset \mathcal{B}(\mathcal{H}_2)$, which we assume to be given. No more assumptions will be made about T. Our task is now to devise coding E and decoding D so that ETD is the identity on $\mathcal{B}(\mathcal{H}_1)$.

The idea is to realize the error corrector's dream by suitable encoding. The 'good' space in that scenario is, of course, the space $\mathcal{H}_1$. We are looking for a way to rewrite $\mathcal{H}_2$ as $\mathcal{H}_1 \otimes \mathcal{H}_b$. Actually, a unitary isomorphism between these spaces may be asking too much, and indeed it suffices to split a subspace of $\mathcal{H}_2$ into a product of good and bad spaces. This requires an isometry $U : \mathcal{H}_1 \otimes \mathcal{H}_b \to \mathcal{H}_2$, which satisfies $U^*U = \mathbb{1}$, but need not be onto. Then the encoding, written best in the Schrödinger picture, is tensoring with an initial state Ω as before, but now with an additional twist by U:

$$E_*(\rho) = U(\rho \otimes |\Omega\rangle\langle\Omega|)U^\dagger \; . \tag{7.15}$$

The decoding operation D is again taking the partial trace over the bad space $\mathcal{H}_b$, after reversing of U. Since U is only an isometry and not necessarily unitary we need an additional term to make D unit preserving. The whole operation is best written in the Heisenberg picture:

$$D(X) = U(X \otimes \mathbb{1})U^\dagger + \mathrm{tr}(\rho_0 X)(\mathbb{1} - UU^\dagger) \; , \tag{7.16}$$

where ρ_0 is an arbitrary density operator. These transformations are successful, if the error space (transformed by U) behaves as before, i.e., if for all $F \in \mathcal{E}$ there are vectors $\Phi(F) \in \mathcal{H}_b$ such that, for all $\phi \in \mathcal{H}_1$

$$FU(\phi \otimes \Omega) = U(\phi \otimes \Phi(F)) \tag{7.17}$$

holds. This equation describes precisely the elements $F \in \mathcal{E}_{\max}$ of the maximal error space.

To check that we really have $ETD = \mathrm{Id}$ for any channel $T(A) = \sum_i F_i^\dagger A F_i$ with $F_i \in \mathcal{E}_{\max}$, it suffices to consider pure input states $|\phi\rangle\langle\phi|$, and the measurement of an arbitrary observable X at the output:

$$\begin{aligned} \mathrm{tr}\big[|\phi\rangle\langle\phi| ETD(X)\big] &= \sum_i \mathrm{tr}\big[U|\phi \otimes \Omega\rangle\langle\phi \otimes \Omega|U^\dagger F_i U(X \otimes \mathbb{1})U^\dagger F_i\big] \\ &= \sum_i \mathrm{tr}\big[|\phi \otimes \Phi(F_i)\rangle\langle\phi \otimes \Phi(F_i)|X \otimes \mathbb{1}\big] \\ &= \langle\phi, X\phi\rangle \sum_i \|\Phi(F_i)\|^2 = \langle\phi, X\phi\rangle. \end{aligned} \tag{7.18}$$

In the last equation we have used that $\sum_i \|\Phi(F_i)\|^2 = 1$, since E, T, and D each map $\mathbb{1}$ to $\mathbb{1}$.

7.4.3 The Knill–Laflamme Condition

The encoding E defined in (7.15) is of the form $E_*(\rho) = V\rho V^\dagger$ with the *encoding isometry* $V : \mathcal{H}_1 \to \mathcal{H}_2$ given by

$$V\phi = U(\phi \otimes \Omega) \; . \tag{7.19}$$

If we just know this isometry and the error space we can reconstruct the whole structure, including the decomposition $\mathcal{H}_2 = \mathcal{H}_1 \otimes \mathcal{H}_b \oplus (\mathbb{1} - UU^\dagger)\mathcal{H}_2$, and hence the decoding operation D. A necessary condition for this, first established by Knill and Laflamme [11], is that, for arbitrary $\phi_1, \phi_2 \in \mathcal{H}_1$ and error operators $F_1, F_2 \in \mathcal{E}$:

$$\langle V\phi_1, F_1^\dagger F_2 V\phi_2\rangle = \langle \phi_1, \phi_2\rangle \omega(F_1^\dagger F_2) \tag{7.20}$$

holds with some numbers $\omega(F_1^\dagger F_2)$ independent of ϕ_1, ϕ_2. Indeed, from (7.17) we immediately get this equation with $\omega(F_1^\dagger F_2) = \langle \Phi(F_1), \Phi(F_2)\rangle$. Conversely, if the Knill–Laflamme condition (7.20) holds, the numbers $\omega(F_1^\dagger F_2)$ serve as a (possibly degenerate) scalar product on $\mathcal{E}$, which upon completion becomes the 'bad space' $\mathcal{H}_b$, such that $F \in \mathcal{E}$ is identified with a Hilbert space vector $\Phi(F)$. The operator $U : \phi \otimes \Phi(F) = FV\phi$ is then an isometry, as used at the beginning of this section. To conclude, the Knill–Laflamme condition is necessary and sufficient for the existence of a decoding operation. Its main virtue is that we can use it without having to construct the decoding explicitly.

7.4.4 Example: Localized Errors

Let us come back to the problem we are addressing in this paper. In that case the space $\mathcal{H}_2$ is the n-fold tensor product of the system $\mathcal{H}$ on which the noisy channels under consideration act. We say that a coding isometry $V : \mathcal{H}_1 \to \mathcal{H}^{\otimes n}$ *corrects f errors*, if it satisfies the Knill–Laflamme condition (7.20) for the error space $\mathcal{E}_f$ spanned linearly by all operators of the kind $X_1 \otimes X_2 \otimes \cdots \otimes X_n$, where at most f places we have a tensor factor $X_i \neq \mathbb{1}$.

When F_1 and F_2 are both supported on at most f sites, the product $F_1^\dagger F_2$, which appears in the Knill–Laflamme condition involves $2f$ sites. Therefore we can paraphrase the condition by saying that

$$\langle V\phi_1, XV\phi_2\rangle = \langle \phi_1, \phi_2\rangle \omega(X) \tag{7.21}$$

for $X \in \mathcal{E}_{2f}$. From Kraus operators in $\mathcal{E}_f$ we can build arbitrary channels of the kind $T = T_1 \otimes T_2 \otimes \cdots \otimes T_n$, where at most f of the tensor factors T_i are channels different from Id. We will use this in the form that $E(R_1 \otimes R_2 \otimes \cdots \otimes R_n)D = 0$, whenever at most f tensor factors are $R_i \neq \mathrm{Id}$, and at least one of them is a difference of two channels.

There are several ways to construct error correcting codes of this type (see e.g. [6,3,1]). Most appropriate for our purposes is the scheme proposed in [18], which is quite easy to describe and admits a simple way to check the error correction condition. This will be the subject of the next section.

7.5 Graph Codes

The general scheme of graph codes works not just for qubits, but for any dimension d of one site spaces. The code will have some number m of input systems, which we label by a set X, and, similarly n output systems, labeled by a set Y. The Hilbert space of the system with label $x \in X \cup Y$ will be denoted by $\mathcal{H}_x$ although all these are isomorphic to $\mathbb{C}^d$, and are equipped with a special basis $|j_x\rangle$, where $j_x \in \mathbb{Z}_d$ is an integer taken modulo d. As a convenient shorthand, we write j_X for a tuple of $j_x \in \mathbb{Z}_d$, specified for every $x \in X$. Thus the $|j_X\rangle$ form a basis of the input space $\mathcal{H}_X = \bigotimes_{x\in X} \mathcal{H}_x$ of the code. An operator F, say, on the output space will be called *localized* on a subset $Z \subset Y$ of systems, if it is some operator on $\bigotimes_{y\in Z} \mathcal{H}_y$, tensored with the identity operators of the remaining sites.

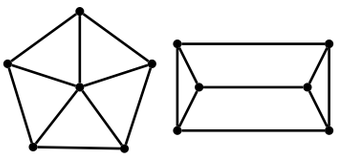


Fig. 7.3. Two graph codes

The main ingredient of the code construction is now an undirected graph with vertices $X \cup Y$. The links of the graph are given by the adjacency matrix, which we will denote by Γ. When we have $|X| = m$ input vertices and $|Y| = n$ output vertices, this is an $(n+m) \times (n+m)$ matrix with $\Gamma_{xy} = 1$ if node x and y are linked and $\Gamma_{xy} = 0$ otherwise. We do allow multiple edges, so the entries of Γ will in general be integers, which can also be taken modulo d. It is convenient to exclude self-linked vertices, so we always take $\Gamma_{xx} = 0$.

The graph determines an operator $V = V_\Gamma : \mathcal{H}_X \to \mathcal{H}_Y$ by the formula

$$\langle j_Y | V_\Gamma | j_X \rangle = d^{-n/2} \exp\left(\frac{\mathrm{i}\pi}{d}\; j_{X\cup Y} \cdot \Gamma \cdot j_{X\cup Y}\right), \tag{7.22}$$

where the exponent contains the matrix element of Γ

$$j_{X\cup Y} \cdot \Gamma \cdot j_{X\cup Y} = \sum_{x,y\in X\cup Y} j_x \Gamma_{xy} j_y \; . \tag{7.23}$$

Because Γ is symmetric, every term in this sum appears twice, hence adding a multiple of d to any j_x or Γ_{xy} will change the exponent in (7.22) by a multiple of 2π, and thus will not change V_Γ.

The error correcting properties of V_Γ are summarized in the following result [18]. It is just the Knill–Laflamme condition (7.20) with a special expression for ω, valid for all error operators such that $F_1^\dagger F_2$ is localized on a set Z.

Proposition 3 *Let Γ be a graph, i.e., a symmetric matrix with entries $\Gamma_{xy} \in \mathbb{Z}_d$, for $x, y \in (X \cup Y)$. Consider a subset $Z \subset Y$, and suppose that the $(Y \setminus Z) \times (X \cup Z)$-submatrix of Γ is non-singular, i.e.,*

$$\forall_{y \in Y \setminus Z} \sum_{x \in X \cup Z} \Gamma_{yx} h_x \equiv 0 \quad \textit{implies } \forall_{x \in X \cup Z}\ h_x \equiv 0 \tag{7.24}$$

where congruences are $\bmod d$. *Then, for every operator $F \in \mathcal{B}(\mathcal{H}_Y)$ localized on Z, we have*

$$V_\Gamma^\dagger F V_\Gamma = d^{-n} \operatorname{tr}(F) \mathbb{1}_X \ . \tag{7.25}$$

Proof. It will be helpful to use the notation for collections of variables, already present in (7.23) more systematically: for any subset $W \subset X \cup Y$ we write j_W for the collection of variables j_y with $y \in W$. The Kronecker-Delta $\delta(j_W)$ is defined to be zero if for any $y \in W$ $j_y \neq 0$, and one otherwise. By $j_W \cdot \Gamma_{WW'} \cdot k_{W'}$ we mean the suitably restricted sum, i.e., $\sum_{x \in W, y \in W'} j_x \Gamma_{xy} k_y$. The important sets to which we apply this notation are $X' = (X \cup Z)$ and $Y' = Y \setminus Z$. In particular, the condition on Γ can be written as $\Gamma_{Y'X'} j_{X'} = 0 \implies j_{X'} = 0$.

Consider now the matrix element

$$\begin{aligned} \langle j_X | V_\Gamma^\dagger F V_\Gamma | k_X \rangle &= \sum_{j_Y, k_Y} \langle j_X | V_\Gamma^\dagger | j_Y \rangle \langle j_Y | F | k_Y \rangle \langle k_Y | V_\Gamma | k_X \rangle \\ &= d^{-n} \sum_{j_Y, k_Y} \exp\left[\frac{\mathrm{i}\pi}{d} \left(k_{X \cup Y} \cdot \Gamma \cdot k_{X \cup Y} - j_{X \cup Y} \cdot \Gamma \cdot j_{X \cup Y} \right) \right] \\ &\qquad \times \langle j_Y | F | k_Y \rangle \ . \end{aligned} \tag{7.26}$$

Since F is localized on Z, the matrix element contains a factor δ_{j_y, k_y} for every $y \in Y \setminus Z = Y'$, so we can write $\langle j_Y | F | k_Y \rangle = \langle j_Z | F | k_Z \rangle \delta(j_{Y'} - k_{Y'})$. Therefore we can compute the sum (7.26) in stages:

$$\langle j_X | V_\Gamma^\dagger F V_\Gamma | k_X \rangle = \sum_{j_Z, k_Z} \langle j_Z | F | k_Z \rangle S(j_{X'}, k_{X'}) \ , \tag{7.27}$$

where $S(j_{X'}, k_{X'})$ is the sum over the Y'-variables, which, of course, still depends on the input variables j_X, k_X and the variables j_Z, k_Z at the error positions:

$$\begin{aligned} S(j_{X'}, k_{X'}) = d^{-n} \sum_{j_{Y'}, k_{Y'}} & \delta(j_{Y'} - k_{Y'}) \\ & \times \exp\left[\frac{\mathrm{i}\pi}{d} \left(k_{X \cup Y} \cdot \Gamma \cdot k_{X \cup Y} - j_{X \cup Y} \cdot \Gamma \cdot j_{X \cup Y} \right) \right] \ . \end{aligned} \tag{7.28}$$

The sums in the exponent can each be split into four parts according to the decomposition X' vs. Y'. The terms involving $\Gamma_{Y'Y'}$ cancel because $k_{Y'} = j_{Y'}$. The terms involving $\Gamma_{X'Y'}$ and $\Gamma_{Y'X'}$ are equal because Γ is symmetric, and together give $2 j_{Y'} \cdot \Gamma_{Y'X'} \cdot (k_{X'} - j_{X'})$. The $\Gamma_{X'X'}$ remain unchanged, but only

give a phase factor independent of the summation variables. Hence

$$\begin{aligned} S(j_{X'}, k_{X'}) &= d^{-n} \exp\left[\frac{\mathrm{i}\pi}{d}\,(k_{X'}\cdot\Gamma\cdot k_{X'} - j_{X'}\cdot\Gamma\cdot j_{X'})\right] \\ &\quad\times\sum_{j_{Y'}} \exp\left[\frac{2\pi\mathrm{i}}{d} j_{Y'}\cdot\Gamma_{Y'X'}\cdot(k_{X'}-j_{X'})\right] \\ &= d^{-n} \exp\left[\frac{\mathrm{i}\pi}{d}\,(k_{X'}\cdot\Gamma\cdot k_{X'} - j_{X'}\cdot\Gamma\cdot j_{X'})\right] \\ &\quad\times d^{|Y'|}\ \delta(\Gamma_{Y'X'}\cdot(k_{X'}-j_{X'})) \\ &= d^{-n+|Y'|} \exp\left[\frac{\mathrm{i}\pi}{d}\,(k_{X'}\cdot\Gamma\cdot k_{X'} - j_{X'}\cdot\Gamma\cdot j_{X'})\right] \\ &\quad\times\delta(k_{X'}-j_{X'}) \\ &= d^{-n+|Y'|}\delta(k_{X'}-j_{X'})\ . \end{aligned} \tag{7.29}$$

Here we used at the first equation that the sum is a product of geometric series as they appear in discrete Fourier transforms. At the second equality the main condition of the Proposition enters: if $\sum_{x\in X'}\Gamma_{yx}\cdot(k_x-j_x)$ vanishes for all $y\in Y'$ as required by the delta-function then (and only then) the vector $k_{X'}-j_{X'}$ must vanish. But then the two terms in the exponent of the phase factor also cancel.

Inserting this result into (7.27), and using that $\delta(h_{X'}) = \delta(h_X)\delta(h_Z)$, we find

$$\begin{aligned} \langle j_X|V_\Gamma^\dagger F V_\Gamma|k_X\rangle &= \delta(j_X-k_X)\ d^{-n+|Y'|}\sum_{j_Z}\langle j_Z|F|j_Z\rangle \\ &= \delta(j_X-k_X)\ d^{-n}\sum_{j_Y}\langle j_Y|F|j_Y\rangle\ . \end{aligned}$$

Here the error operator is considered in the first line as an operator on $\mathcal{H}_Z$, and as an operator on $\mathcal{H}_Y$ in the second line, by tensoring it with $\mathbb{1}_{Y'}$. This cancels the dimension factor $d^{|Y'|}$. □

All that is left to get an error correcting code is to ensure that the conditions of this Proposition are satisfied sufficiently often. This is evident from combining the above Proposition with the example at the end of Sect. 7.4.3.

Corollary 4 *Let Γ be a graph as in the previous Proposition, and suppose that the $(Y\setminus Z)\times(X\cup Z)$-submatrix of Γ is non-singular for* all *$Z\subset Y$ with up to $2f$ elements. Then the code associated to Γ corrects f errors.*

Two particular examples (which actually turn out to be equivalent!) are given in Fig. 7.3. In both cases we have $N=1$, $M=5$ and $K=1$, i.e., there is one input node – which can be chosen arbitrarily – and five output nodes, and the corresponding codes correct one error.

7.6 Discrete to Continuous Error Model

The discrete error correction scheme described in the last section is not really designed to correct *small* errors: it corrects *rare* errors in multiple applications of the channel. A typical example of a small (but not rare) error is a small unitary rotation, $T(X) = U^\dagger XU$. Then $\|T - \mathrm{Id}\|_{\mathrm{cb}}$ can be small, but since the same small error happens to each of the parallel channels in $T^{\otimes n}$, the error patterns of discrete error correction, as encoded in the space $\mathcal{E}$ of Sect. 7.4, at first sight do not seem to be appropriate at all. Nevertheless, the discrete theory can be applied, and this is the content of the following Proposition. It is the appropriate formulation of "reducing the order of errors from ε to ε^{f+1}".

Proposition 5 *Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a channel, and let E, D be encoding and decoding channels for coding m systems into n systems. Suppose that this coding scheme corrects f errors, and that*

$$\|T - \mathrm{Id}\|_{\mathrm{cb}} \leq (f+1)/(n-f-1). \tag{7.30}$$

Then

$$\|ET^{\otimes n}D - \mathrm{Id}\|_{\mathrm{cb}} \leq \|T - \mathrm{Id}\|_{\mathrm{cb}}^{f+1}\, 2^{nH_2((f+1)/n)}\,, \tag{7.31}$$

where

$$H_2(r) = -r\log_2 r - (1-r)\log_2(1-r) \tag{7.32}$$

denotes the Shannon entropy of the probability distribution $(r, 1-r)$.

Proof. Into $ET^{\otimes n}D$, we insert the decomposition $T = \mathrm{Id} + (T - \mathrm{Id})$ and expand the product. This gives 2^n terms, containing tensor products with some number, say k, of tensor factors $(T - \mathrm{Id})$ and tensor factors Id on the remaining $(n-k)$ sites. Now when $k \leq f$, the error correction property makes the term zero. Terms with $k > f$ we estimate by $\|T - \mathrm{Id}\|_{\mathrm{cb}}^k$. Collecting terms we get

$$\|ET^{\otimes n}D - \mathrm{Id}\|_{\mathrm{cb}} \leq \sum_{k=f+1}^{n} \binom{n}{k} \|T - \mathrm{Id}\|_{\mathrm{cb}}^k\,. \tag{7.33}$$

The rest then follows from the next Lemma (with $r = (f+1)/n$). It treats the exponential growth in n for truncated binomial sums. □

Lemma 6 *Let $0 \leq r \leq 1$ and $a > 0$ such that $a \leq r/(1-r)$. Then, for all integers n:*

$$\frac{1}{n}\log\left(\sum_{k=rn}^{n}\binom{n}{k}a^k\right) \leq \log\bigl(a^r\bigr) + H_2(r)\,. \tag{7.34}$$

Proof. For $\lambda > 0$ we can estimate the step function by an exponential, and get

$$\begin{aligned}\sum_{k=rn}^{n}\binom{n}{k}a^k &\leq \sum_{k=0}^{n}\binom{n}{k}a^k e^{\lambda(k-rn)} \\ &= e^{-\lambda rn}\bigl(1+ae^{\lambda}\bigr)^n = M(\lambda)^n\end{aligned} \tag{7.35}$$

with $M(\lambda) = e^{-\lambda r}\big(1+ae^{\lambda}\big)$. The minimum over all real λ is attained at $ae^{\lambda_{\min}} = r/(1-r)$. We get $\lambda_{\min} \geq 0$ precisely when the conditions of the Lemma are satisfied, in which case the bound is computed by evaluating $M(\lambda)$. □

Suppose now that we find a family of coding schemes with $n, m \to \infty$ with fixed rate $r \approx (m/n)$ of inputs per output, and a certain fraction $f/n \approx \varepsilon$ of errors being corrected. Then we can apply the Proposition and find that the errors can be estimated from above by

$$\Delta\left(T^{\otimes n}, d^m\right) \leq \left(2^{H_2(\varepsilon)} \, \|T - \mathrm{Id}\,\|_{\mathrm{cb}}^{\varepsilon}\right)^n \,, \tag{7.36}$$

where d is the Hilbert space dimension of each input system. This goes to zero, and even exponentially to zero, as soon as the expression in parentheses is < 1. This will be the case whenever $\|T - \mathrm{Id}\,\|_{\mathrm{cb}}$ is small enough, or, more precisely,

$$\|T - \mathrm{Id}\,\|_{\mathrm{cb}} \leq \; 2^{-H_2(\varepsilon)/\varepsilon} \,. \tag{7.37}$$

Note in addition that we have for all $n \in \mathbb{N}$

$$2^{H_2(\varepsilon)/\varepsilon} < \frac{\varepsilon - \frac{1}{n}}{1 - \varepsilon + \frac{1}{n}} \,. \tag{7.38}$$

Hence the bound from (7.30) is implied by (7.37).

The function appearing on the right hand side of (7.37) looks rather complicated, so we will often replace it by a simpler one, namely

$$\frac{\varepsilon}{e} \leq 2^{-H_2(\varepsilon)/\varepsilon} \,, \tag{7.39}$$

where e is the base of natural logarithms; cf. Fig. 7.4. The proof of this inequality is left to the reader as exercise in logarithms. The bound is very good (exact to first order) in the range of small ε, in which we are most interested anyhow. In any case, from $\|T - \mathrm{Id}\,\|_{\mathrm{cb}} \leq \varepsilon/e$ we can draw the same conclusion as from (7.37): exponentially decreasing errors, provided we can actually find code families correcting a fraction ε of errors. This will be the aim of the next section.

7.7 Coding by Random Graphs

Our aim in this section is to apply the theory of graph codes to construct a family of codes with positive rate. It is not so easy to construct such families explicitly. However, if we are only interested in existence, and do not attempt to get the best possible rates, we can use a simple argument, which shows not only the existence of codes correcting a certain fraction of errors, but even that "typical graph codes" for sufficiently large numbers of inputs and outputs have this property. Here "typical" is in the sense of the probability distribution, defined by simply setting the edges of the graph independently, and each according to the uniform distribution of the possible values of the adjacency matrix. For the random

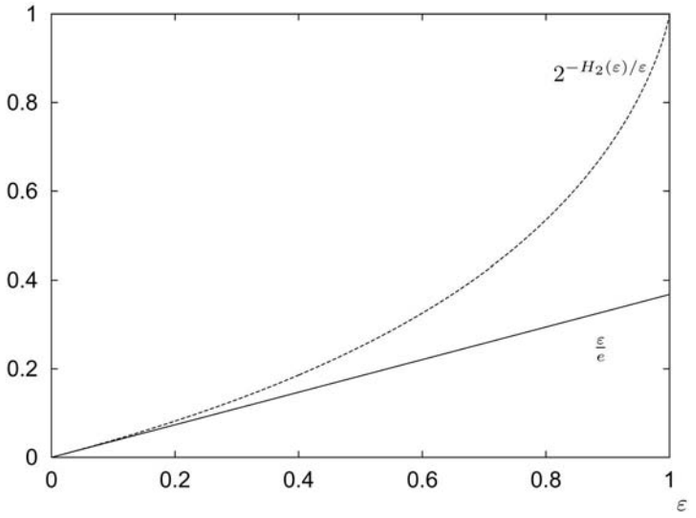


Fig. 7.4. The two bounds from (7.39) plotted as a function of ε

method to work we need the dimension of the underlying one site Hilbert space to be a prime number. This curious condition is most likely an artefact of our method, and will be removed later on.

We have seen that a graph code corrects many errors if certain submatrices of the adjacency matrix have maximal rank. Therefore we need the following Lemma.

Lemma 7 *Let d be a prime, $M < N$ integers and let X be an $N \times M$-matrix with independent and uniformly distributed entries in $\mathbb{Z}_d$. Then X is singular over the field $\mathbb{Z}_d$ with probability at most $d^{-(N-M)}$.*

Proof. The sum of independent uniformly distributed random variables in $\mathbb{Z}_d$ is again uniformly distributed. Moreover, since d is prime, this distribution is invariant under multiplication by non-zero factors. Hence if $x_j \in \mathbb{Z}_d$ $(j = 1, \ldots, N)$ are independent and uniformly distributed, and $\phi_j \in \mathbb{Z}_d$ are non-random constants, not of all of which are zero, $\sum_{j=1}^{N} x_j \phi_j$ is uniformly distributed. Hence, for a fixed vector $\phi \in {\mathbb{Z}_d}^M$, the N components $(X\phi)_k = \sum_{j=1}^{M} X_{kj}\phi_j$ are independent uniformly distributed random variables. Hence the probability for $X\phi = 0$ for some fixed $\phi \neq 0$ is d^{-N}. Since there are $d^M - 1$ vectors ϕ to be tested, the probability for *some* ϕ to yield $X\phi = 0$ is at most d^{M-N}. □

Proposition 8 *Let d be a prime, and let Γ be a symmetric $(n+m) \times (n+m)$-matrix with entries in $\mathbb{Z}_d$, chosen at random such that $\Gamma_{kk} = 0$ and that the Γ_{kj} with $k > j$ are independent and uniformly distributed. Let P be the probability for the corresponding graph code* not *to correct f errors (with $2f < n$). Then*

$$\frac{1}{n}\log P \leq \Big(\frac{m}{n} + \frac{4f}{n} - 1\Big)\log d + H_2\Big(\frac{2f}{n}\Big)\ . \tag{7.40}$$

Proof. Each error configuration is a $2f$-element subset of the n output nodes. According to Proposition 3 we have to decide, whether the corresponding $(n-2f)\times(m+2f)$-submatrix of Γ, connecting input and error positions with the remaining output positions, is singular or not. Since this submatrix contains no pairs Γ_{ij}, Γ_{ji}, its entries are independent and satisfy the conditions of the previous Lemma. Hence the probability that a particular configuration of f errors goes uncorrected is at most $d^{(m+2f)-(n-2f)}$. Since there are $\binom{n}{2f}$ possible error configurations among the outputs, we can estimate the probability of any $2f$ site error configuration to be undetected as less than $\binom{n}{2f} d^{m-n+4f}$. Using Lemma 6 we can estimate the binomial as $\log\binom{n}{2f} \leq nH_2(2f/n)$, which leads to the bound stated. □

In particular, if the right hand side of the inequality in (7.40) is negative, we get $P < 1$, so that there must be at least one matrix Γ correcting f errors. The crucial point is that this observation does not depend on n, but only on the rate-like parameters m/n and f/n. Let us make this behaviour a Definition:

Definition 9 *Let d be an integer. Then we say a pair (μ,ε) consisting of a* coding rate μ *and an* error rate ε *is* achievable*, if for every n we can find an encoding E of at least (μn) d-level systems into n d-level systems correcting $\lfloor \varepsilon n \rfloor$ errors.*

Then we can paraphrase the last Proposition as saying that all pairs (μ,ε) with

$$(1-\mu-4\varepsilon)\log_2 d > H_2(2\varepsilon) \tag{7.41}$$

are achievable. This is all the input we need for the next section, although a better coding scheme, giving larger μ or larger ε would also improve the rate estimates proved there. Such improvements are indeed possible. E.g. for the qubit case ($d = 2$) it is shown in [3] that there is always a code which saturates the *quantum Gilbert–Varshamov bound* $(1-\mu-2\varepsilon\log_2(3)) > H_2(2\varepsilon)$, which is slightly better than our result.

But there are also known limitations, particularly the so-called *Hamming bound*. This is a simple dimension counting argument, based on the error corrector's dream: Assuming that the scalar product $(F,G)\mapsto\omega(F^\dagger G)$ on the error space $\mathcal{E}$ is non-degenerate, the dimension of the "bad space" is the same as the dimension of the error space. Hence with the notations of Sect. 7.4 we expect $\dim\mathcal{H}_0\cdot\dim\mathcal{E}\leq\dim\mathcal{H}_2$. We now take m input systems and n output systems of dimension d each, so that $\dim\mathcal{H}_1 = d^m$ and $\dim\mathcal{H}_2 = d^n$. For the space of errors happening at at most f places we introduce a basis as follows: at each site we choose a basis of $\mathcal{B}(\mathcal{H})$ consisting of d^2-1 operators plus the identity. Then a basis of $\mathcal{E}$ is given by all tensor products with basis elements $\neq \mathbb{1}$ placed at $j\leq f$ sites. Hence $\dim\mathcal{E} = \sum_{j\leq f}\binom{n}{j}(d^2-1)^j$. For large n we estimate this as

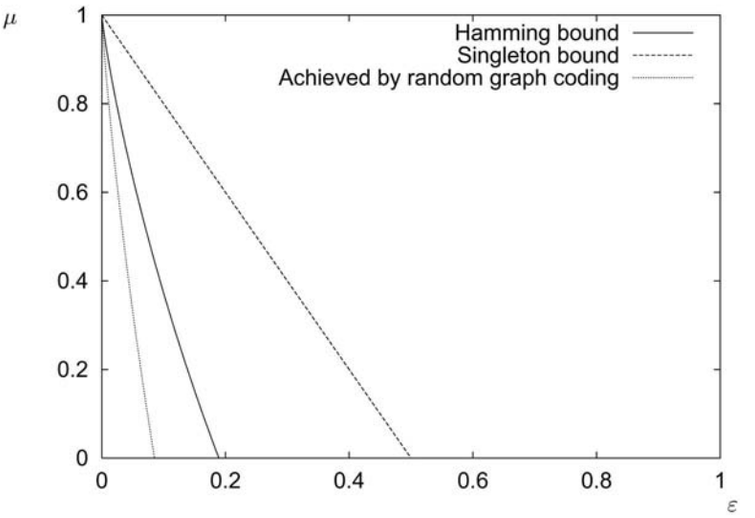


Fig. 7.5. Singleton bound (7.43) and Hamming bound (7.42), together with the rate achieved by random graph coding (for $d = 2$). The allowed regions are below the respective curve

in Lemma 5 as $\log \dim \mathcal{E} \approx (f/n) \log_2(d^2 - 1) + H_2(f/n)$. Hence the Hamming bound becomes

$$\frac{m}{n} \log_2 d + H_2(\varepsilon) + \frac{f}{n} \log_2(d^2 - 1) \leq \log_2 d \tag{7.42}$$

which (with $d^2 \gg 1$) is just (7.41) with a factor $1/2$ on all errors.

If we drop the nondegeneracy condition made above it is possible to find codes which break the Hamming bound [5]. In this case, however, we can consider the weaker *Singleton bound*, which has to be respected by those *degenerate codes* as well. It reads

$$1 - \frac{m}{n} \geq d\frac{f}{n} \,. \tag{7.43}$$

We omit its proof here (see Sect. 12.4 in [14] instead). Both bounds are plotted together with the rate achieved by random graph coding in Fig. 7.5 (for $d = 2$).

7.8 Conclusions

We are now ready to combine our discussion of channel-capacity from Sect. 7.3 with the results about error correction we have derived in the previous sections. Please note that most of the result presented here can be found in [8,13], in some cases with better bounds.

7.8.1 Correcting Small Errors

We first look at the problem which motivated our study, namely estimating the capacity of a channel $T \approx \mathrm{Id}$.

Theorem 10 *Let d be a prime, and let T be a channel on d-level systems. Suppose that for some $0 < \varepsilon < 1/2$,*

$$\| \mathrm{Id} - T \|_{\mathrm{cb}} < 2^{-H_2(\varepsilon)/\varepsilon} \, . \tag{7.44}$$

Then

$$Q(T) \geq (1 - 4\varepsilon) \log_2(d) - H_2(2\varepsilon) \, . \tag{7.45}$$

Proof. For every n set $f = \lfloor \varepsilon n \rfloor$, and $m = \lfloor \mu n \rfloor - 1$, where μ is, up to a $\log_2(d)$ factor, the right hand side of (7.45), i.e. $\mu = 1 - 4\varepsilon - \log_2(d)^{-1} H_2(2\varepsilon)$. This ensures that the right hand side of (7.40) is strictly negative, so there must be a code for d-level systems, with m inputs and n outputs, and correcting f errors. To this code we apply Proposition 5, and insert the bound on $\| \mathrm{Id} - T \|_{\mathrm{cb}}$ into (7.36). Thus $\Delta(T^{\otimes n}, d^{\lfloor \mu n \rfloor - 1}) \to 0$, even exponentially. This means that any number $< \mu \log_2(d)$ is an achievable rate. In other words, $\mu \log_2(d)$ is a lower bound to the capacity. □

If $\varepsilon > 0$ is small enough, the quantity on the right hand side of (7.45) is strictly positive (cf. the dotted graph in Fig. 7.5). Hence each channel which is sufficiently close to the identity allows (asymptotically) perfect error correction. Beyond that we see immediately that $Q(T)$ is continous (in the cb-norm) at $T = \mathrm{Id}$: Since $Q(T)$ is smaller than $\log_2(d)$ and $g(\varepsilon)$ is continuous in ε with $g(0) = \log_2(d)$, we find that for each $\delta > 0$ an $\varepsilon > 0$ exists, such that $\log_2(d) - Q(T) < \varepsilon$ for all T with $\| T - \mathrm{Id} \|_{\mathrm{cb}} < \varepsilon / e$. In other words, if T is arbitrarily close to the identity its capacity is arbitrarily close to $\log_2(d)$. In Corollary 12 below we will show the significantly stronger statement that Q is a lower semicontinuous function on the set of all channels.

7.8.2 Estimating Capacity from Finite Coding Solutions

A crucial consequence of the ability to correct small errors is that we do not actually have to compute the limit defining the capacity: if we have a pretty good coding scheme for a given channel, i.e., one that gives us $ET^{\otimes n}D \approx \mathrm{Id}_d$, then we know the errors can actually be brought to zero, and the capacity is close to the nominal rate of this scheme, namely $\log_2(d)/n$.

Theorem 11 *Let T be a channel, not necessarily between systems of the same dimension. Let $k, p \in \mathbb{N}$ with p a prime number, and suppose there are channels E and D encoding and decoding a p-level system through k parallel uses of T, with error $\Delta = \| \mathrm{Id}_p - ET^{\otimes k}D \|_{\mathrm{cb}} < 1/2\mathrm{e}$. Then*

$$Q(T) \geq \frac{\log_2(p)}{n} (1 - 4e\Delta) - \frac{1}{n} H_2(2e\Delta) \, . \tag{7.46}$$

Moreover, $Q(T)$ is the least upper bound on all expressions of this form.

Proof. We apply Proposition 10 to the channel $\widetilde{T} = ET^{\otimes n}D$. With the random coding method we thus find a family of coding and decoding channels $\widetilde{E}$ and $\widetilde{D}$ from m' into n' systems, of p levels each, such that

$$\| \operatorname{Id} - \widetilde{E}\big(ET^{\otimes k}D\big)^{\otimes n'}\widetilde{D}\|_{\mathrm{cb}} \to 0. \tag{7.47}$$

This can be reinterpreted as an encoding of $p^{m'}$-dimensional systems through kn' uses of the channel T (rather than the channel $\widetilde{T}$), which corresponds to a rate $(kn')^{-1}\log_2(p^{m'}) = (\log_2 p/k)(m'/n')$. We now argue exactly as in the proof of the previous proposition, with $\varepsilon = e\Delta$, so that

$$\| \operatorname{Id}_p - ET^{\otimes k}D\|_{\mathrm{cb}} = \varepsilon/e \le 2^{H_2(\varepsilon)/\varepsilon} \tag{7.48}$$

by (7.39). By random graph coding we can achieve the coding ratio $\mu \approx (m'/n') = 1 - 4\varepsilon - \log_2(p)^{-1}H_2(2\varepsilon)$, and have the errors $\Delta(\widetilde{T}^{\otimes n'}, p^{m'})$ go to zero exponentially. Since

$$\Delta(T^{\otimes kn'}, p^{m'}) \le \Delta(\widetilde{T}^{\otimes n'}, p^{m'}) \le \| \operatorname{Id} - \widetilde{E}\big(ET^{\otimes k}D\big)^{\otimes n'}\|_{\mathrm{cb}}\ , \tag{7.49}$$

we can apply Lemma 2 to the channel T (where the sequence n_α is given by $n_\alpha = n\alpha$) and find that the rate $\mu(\log_2 p/k)$ is achievable. This yields the estimate claimed in (7.46).

To prove the second statement, consider the function $x \to p(x)$ which associates to each real number $x \ge 2$ the biggest prime $p(x)$ with $p(x) \le x$. From known bounds on the length of gaps between two consecutive primes [9][2] it follows that $\lim_{x\to\infty} x/p(x) = 1$ holds, hence we get $2^{kc}/p(2^{kc}) \le 1 + \delta'$ for an arbitrary $\delta' > 0$, provided n is large enough, but this implies

$$c - \frac{\log_2\big[p(2^{kc})\big]}{k} < \frac{\log_2(1+\delta')}{k}\ . \tag{7.50}$$

Since we can choose an achievable rate c arbitrarily close to the capacity $Q(T)$ this shows that there is for each $\delta > 0$ a prime p and a positive integer k such that $|Q(T) - \log_2(p)/k| \le \delta$. In addition we can find a coding scheme E, D for $T^{\otimes k}$ such that (7.48) holds, i.e. the right hand side of (7.46) can be arbitrarily close to $\log_2(p)/k$, and this completes the proof. □

This theorem allows us to derive very easily an important *continuity property* of the quantum capacity, namely the property that the capacity cannot suddenly drop by a finite amount. Technically, this is known as "lower semicontinuity". Equivalent precise statements for a real valued function F on a topological space are [2] (1) $\liminf_{x\to y} F(x) \ge F(y)$ for all y, (2) the upper level sets $F^{-1}\big((x,\infty]\big)$ are open for each $x \in \mathbb{R}$, and (3) that F is the pointwise supremum of continuous functions. The latter characterization immediately applies to capacity: Since the right hand side of (7.46) is continuous in T, and since $Q(T)$ is (according to Proposition 11) the supremum over such quantities, we get:

Corollary 12 *$T \mapsto Q(T)$ is lower semi-continuous in cb-norm.*

[2] If p_n denotes the n^{th} prime and $g(p_n) = p_{n+1} - p_n$ is the length of the gap between p_n and p_{n+1}, it is shown in [9] that $g(p)$ is bounded by $\text{const} \cdot p^{5/8+\varepsilon}$.

7.8.3 Error Exponents

Another consequence of Theorem 11 concerns the rate with which the error $\Delta(T^{\otimes n}, 2^{\lfloor cn \rfloor})$ decays in the limit $n \to \infty$. Theorem 11 says, roughly speaking, that we can achieve *every rate* $c < Q(T)$ by combining a coding scheme E, D with subsequent random-graph coding $\widetilde{E}, \widetilde{D}$. However, the error $\Delta\big[(ET^{\otimes n}D)^{\otimes l}, p^k\big]$ decays according to (7.36) and Proposition 8 exponentially. A more precise analysis of this idea leads to the following (cf. also the work by Hamada [8])

Proposition 13 *If T is a channel with quantum capacity $Q(T)$ and $c < Q(T)$, then, for sufficiently large n, we have*

$$\Delta(T^{\otimes n}, 2^{\lfloor cn \rfloor}) \leq e^{-n\lambda(c)} \;, \tag{7.51}$$

with a positive constant $\lambda(c)$.

Proof. We start as in Theorem 11 with the channel $\widetilde{T} = ET^{\otimes k}D$ and the quantity $\Delta = \| \mathrm{Id}_p - ET^{\otimes k}D\|_{\mathrm{cb}}$. However, instead of assuming that $\Delta = \varepsilon/e$ holds, the full range $e\Delta \leq \varepsilon \leq 1/2$ is allowed for the error rate ε. Using the same arguments as in the proof of Theorem 11, we get an achievable rate

$$c(k, p, \varepsilon) = \frac{\log_2(p)}{k} \left(1 - 4\varepsilon - \frac{H_2(2\varepsilon)}{\log_2(p)}\right) \;, \tag{7.52}$$

and an exponential bound on the coding error:

$$\Delta(T^{\otimes kn'}, p^{m'}) \leq \| \mathrm{Id} - \widetilde{E}\big(ET^{\otimes k}D\big)^{\otimes n'} \|_{\mathrm{cb}} \leq \left(2^{H_2(\varepsilon)}\Delta^{\varepsilon}\right)^{n'} ; \tag{7.53}$$

cf. (7.36) and (7.49).

To calculate the exponential rate $\lambda(c)$ with which the coding error vanishes we have to consider the quantity

$$\lambda(c) = \liminf_{n\to\infty} -\frac{1}{n} \ln \Delta(T^{\otimes n}, \lfloor 2^{nc} \rfloor) \geq \lim_{n'\to\infty} \frac{-1}{kn'} n' \ln \left(2^{H_2(\varepsilon)}\Delta^{\varepsilon}\right) \tag{7.54}$$

$$\geq -\frac{\varepsilon}{k}\left(\ln(\Delta) + \ln 2\frac{H_2(\varepsilon)}{\varepsilon}\right) = -\varepsilon\Lambda(\Delta, \varepsilon)/k \;, \tag{7.55}$$

where we have inserted inequality (7.53). Now we we can apply Lemma 2 (with the sequence $n_\alpha = k\alpha$), which shows that $\lambda(c)$ is positive, if the right hand side of (7.55) is.

What remains to show is that $\lambda(c) > 0$ holds for all $c < Q(T)$. To this end we have to choose k, p, Δ and ε such that $c(k, p, \varepsilon) = c$ and $\Lambda(\Delta, \varepsilon) < 0$. Hence, consider $\delta > 0$ such that $c + \delta < Q(T)$ is an achievable rate. As in the proof of Theorem 11, we can choose $\log_2(p)/k$ such that $\log_2(p)/k > c + \delta$ holds while Δ is arbitrarily small. Hence there is an $\varepsilon_0 > 0$, such that $c(k, p, \varepsilon) = c$ implies $\varepsilon > \varepsilon_0$. The statement therefore follows from the fact that there is a $\Delta_0 > 0$ with $\Lambda(\Delta, \varepsilon) > 0$, for all $0 < \Delta < \Delta_0$ and $\varepsilon > \varepsilon_0$. □

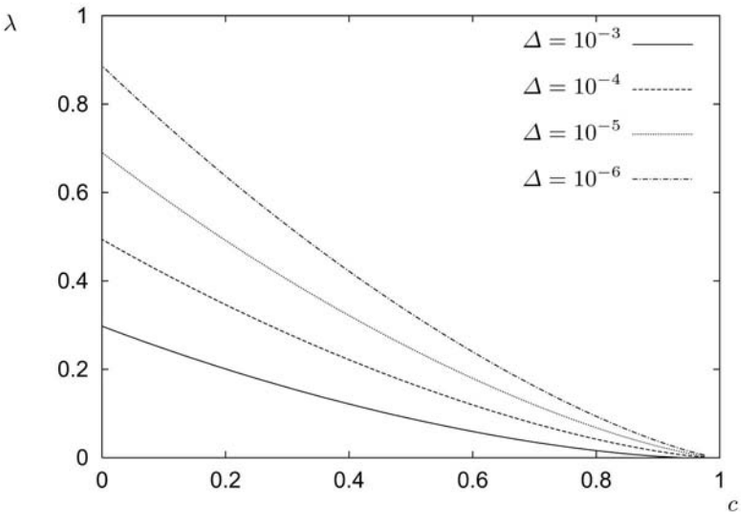


Fig. 7.6. Lower bounds on the error exponent $\lambda(c)$ plotted for $n = 1, p = 2$ and different values of Δ. The curves are generated by combining the estimate (7.52) for capacity c and (7.55) for the exponent λ

In addition to the statement of Proposition 13, we have just derived a lower bound on the error exponent $\lambda(c)$. Since we can not express the error rate ε as a function of k, p and c, we can not specify this bound explicity. However, we can plot it as a parametrized curve (using (7.52) and (7.55) with ε as the parameter) in the (c, λ)-space. In Fig. 7.6 this is done for $k = 1$, $p = 2$, and several values of Δ.

7.8.4 Capacity with Finite Error Allowed

We can also tolerate finite errors in encoding. Let $Q_\varepsilon(T)$ denote the quantity defined exactly like the capacity, but with the weaker requirement that $\Delta(T^{\otimes n}, 2^{\lfloor cn \rfloor}) \leq \varepsilon$ for large n. Obviously we have $Q_\varepsilon(T) \geq Q(T)$, for each $\varepsilon > 0$. Regarded as a function of ε and T, this new quantity admits in addition the following continuity property in ε.

Proposition 14 $\lim_{\varepsilon \to 0} Q_\varepsilon(T) = Q(T)$.

Proof. By definition we can find for each $\varepsilon', \delta > 0$ a tuple n, p, E and D, such that

$$\| \mathrm{Id}_p - ET^{\otimes n} D \|_{\mathrm{cb}} = \frac{\varepsilon' + \varepsilon}{e} \tag{7.56}$$

and $|Q_\varepsilon(T) - \log_2(p)/n| < \delta$ holds. If $\varepsilon + \varepsilon'$ is small enough, however, we find as in Theorem 11 a random graph coding scheme such that

$$Q(T) \geq \frac{\log_2(p)}{n}\big(1 - 4(\varepsilon + \varepsilon')\big) - \frac{1}{n} H_2\big(2(\varepsilon + \varepsilon')\big) = g(\varepsilon + \varepsilon'). \tag{7.57}$$

Hence the statement follows from continuity of g and the fact that $g(0) = \log_2(p)/n$ holds. □

For a classical channel Φ even more is known about the similarly defined quantity $C_\varepsilon(T)$: If $\varepsilon > 0$ is small enough we can not achieve bigger rates by allowing small errors, i.e. $C(T) = C_\varepsilon(T)$. This is called the "strong converse of Shannon's noisy channel coding theorem" [19]. To check whether a similar statement holds in the quantum case is one of the big open problem of the theory.

Acknowledgements

Funding by the European Union project EQUIP (contract IST-1999-11053) and financial support from the DFG (Bonn) is gratefully acknowledged.

References

1. A. Ashikhmin, E. Knill: IEEE T. Inf. Theory **47**, 3065–3072 (2001)
2. N. Bourbaki: *General Topology.* (Addison-Wesley, Reading 1966) Chap. IV§6.2
3. A. R. Calderbank, E. M. Rains, P. W. Shor, N. J. A. Sloane: Phys. Rev. Lett. **78**, 405–408 (1997)
4. A. R. Calderbank, P. W. Shor: Phys. Rev. A **54**, 1098–1105 (1996)
5. D. P. DiVincenzo, P. W. Shor, J. A. Smolin: Phys. Rev. A **57**, 830–839 (1998); Erratum: Phys. Rev. A **59**, 2, 1717 (1999)
6. D. Gottesman: Phys. Rev. A **54**, 1862–1868 (1996)
7. D. Gottesman: Stabilizer codes and quantum error correction. Ph.D. Thesis, California Institute of Technology (1997). Preprint arXiv:quant-ph/9705052
8. M. Hamada: *Exponential lower bound on the highest fidelity achievable by quantum error-correcting codes.* Preprint arXiv:quant-ph/0109114 (2001)
9. A. E. Ingham: Quart. J. Math., Oxford Ser. **8**, 255–266 (1937)
10. M. Keyl: *Fundamentals of quantum information theory.* Preprint arXiv:quant-ph/0202122 (2001)
11. E. Knill, R. Laflamme: Phys. Rev. A **55**, 900–911 (1997)
12. K. Kraus: *States, Effects, and Operations.* (Springer, Berlin 1983)
13. R. Matsumoto, T. Uyematsu: *Lower bound for the quantum capacity of a discrete memoryless quantum channel.* Preprint arXiv:quant-ph/0105151 (2001)
14. M. A. Nielsen, I. L. Chuang: *Quantum computation and quantum information.* (Cambridge University Press, Cambridge 2000)
15. V. I. Paulsen: *Completely bounded maps and dilations.* (Longman Scientific & Technical, Harlow 1986)
16. J. Preskill: *Lecture notes for the course 'information for physics 219/computer science 219, quantum computation'.* Caltech, Pasadena, California (1999). www.theory.caltech.edu/people/preskill/ph229
17. M. Reed, B. Simon: *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis.* (Academic Press, New York, London 1980)
18. D. Schlingemann, R. F. Werner: *Quantum error-correcting codes associated with graphs.* Preprint arXiv:quant-ph/0012111 (2000)
19. C. E. Shannon: Bell. Sys. Tech. J. **27**, 379–423, 623–656 (1948)

20. P. W. Shor: 'Algorithms for quantum computation: Discrete logarithms and factoring', In: *Proc. of the 35th Annual Symposium on the Foundations of Computer Science*, ed. by S. Goldwasser (IEEE Computer Science, Society Press, Los Alamitos, California 1994) pp. 124–134
21. P. W. Shor: Soc. Ind. Appl. Math. J. Comp. **26**, 1484–1509 (1997)
22. A. M. Steane: Proc. Roy. Soc. Lond. A **452**, 2551–2577 (1996)
23. W. F. Stinespring: Proc. Amer. Math. Soc. **6**, 211–216 (1955)
24. R. F. Werner: 'Quantum information theory – an invitation'. In: *Quantum information*, ed. by G. Alber et. al. (Springer, Heidelberg 2001) pp. 14–59
25. W. K. Wootters, W. H. Zurek: Nature **299**, 802–803 (1982)

Index