Erwin Brüning
Francesco Petruccione
Editors

LECTURE NOTES IN PHYSICS 787

Theoretical Foundations of Quantum Information Processing and Communication

Selected Topics

Springer

Lecture Notes in Physics

Founding Editors: W. Beiglböck, J. Ehlers, K. Hepp, H. Weidenmüller

Editorial Board

R. Beig, Vienna, Austria
W. Beiglböck, Heidelberg, Germany
W. Domcke, Garching, Germany
B.-G. Englert, Singapore
U. Frisch, Nice, France
F. Guinea, Madrid, Spain
P. Hänggi, Augsburg, Germany
W. Hillebrandt, Garching, Germany
R. L. Jaffe, Cambridge, MA, USA
W. Janke, Leipzig, Germany
H. v. Löhneysen, Karlsruhe, Germany
M. Mangano, Geneva, Switzerland
J.-M. Raimond, Paris, France
D. Sornette, Zurich, Switzerland
S. Theisen, Potsdam, Germany
D. Vollhardt, Augsburg, Germany
W. Weise, Garching, Germany
J. Zittartz, Köln, Germany

The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching – quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research and to serve three purposes:

- to be a compact and modern up-to-date source of reference on a well-defined topic
- to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas
- to be a source of advanced teaching material for specialized seminars, courses and schools

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive being available at springerlink.com. The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Christian Caron
Springer Heidelberg
Physics Editorial Department I
Tiergartenstrasse 17
69121 Heidelberg / Germany
christian.caron@springer.com

E. Brüning
F. Petruccione (Eds.)

Theoretical Foundations of Quantum Information Processing and Communication

Selected Topics

Springer

Editors
Erwin Brüning
University of KwaZulu-Natal
School of Physics, Westville
Centre for Quantum
Technology
Durban
4000 South Africa
bruninge@ukzn.ac.za

Francesco Petruccione
University of KwaZulu-Natal
School of Physics, Westville
Centre for Quantum
Technology
Durban
4000 South Africa
petruccione@ukn.ac.za

Brüning E., Petruccione F. (Eds.), *Theoretical Foundations of Quantum Information Processing and Communication: Selected Topics*, Lect. Notes Phys. 787 (Springer, Berlin Heidelberg 2010), DOI 10.1007/978-3-642-02871-7

Lecture Notes in Physics ISSN 0075-8450 e-ISSN 1616-6361
ISBN 978-3-642-02870-0 e-ISBN 978-3-642-02871-7
DOI 10.1007/978-3-642-02871-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009935346

© Springer-Verlag Berlin Heidelberg 2010
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: Integra Software Services Pvt. Ltd., Pondicherry

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The present volume contains the text of lectures given at the 18th Chris Engelbrecht Summer School in Theoretical Physics on Theoretical Foundations of Quantum Information Processing and Communication. The Summer School was held in Salt Rock on the Dolphin Coast of KwaZulu-Natal from 14 to 24 January 2007. It was organized under the auspices of the South African National Institute for Theoretical Physics (NITheP) and the National Research Foundation (NRF).

The School was intended to stimulate the interest of an upcoming new generation of South African students to the emerging field of Quantum Information Processing and Communication (QIPC). Topics that are not usually covered in traditional undergraduate courses were presented in nine lectures. All lectures are essentially self-contained and at a level that should be very useful to introduce post-graduate students to the subject. Some are at an introductory level while others describe some recent developments and results in their field of specialization.

Here is a very brief description of the contents of the nine contributions.

The first contribution by Mark Fannes introduces the conceptual and mathematical framework of quantum mechanics, stressing the uniformity in the description of the classical and the quantum world. In order to avoid too many technical complications the authors present the case of finite-dimensional state spaces only which is sufficient for many aspects of QIPC. Thus we learn about the common and the distinguishing features of the classical and the quantum harmonic oscillator, about observables and states of a quantum system and how composite quantum systems are described. Finally the dynamics of a quantum system and its mathematical realization is explained, as well as the description of (ideal) measurements and the importance of completely positive maps for open quantum systems.

The basic idea of the contribution by Bassano Vacchini is that quantum mechanics can naturally be seen as a probability theory (significantly different from classical probability) rather than as an extension of classical mechanics. Accordingly the contribution starts with a brief description of quantum probability. In this way we learn about the statistics of an experiment, about states as preparation procedures, observables as registration procedures, and the description of the statistics of outcomes. Then position and momentum observables are introduced as projection operator-valued measures (POVMs) which are characterized by their

covariance properties with respect to the isochronous Galileo group. Mappings describing state transformations are considered with respect to the general constraints and their covariance properties with respect to symmetry groups. In particular different master equations are analyzed in this way. This is done explicitly for the damped harmonic oscillator, the two-level quantum system, and quantum Brownian motion. Finally, translation-covariant mappings for the description of dissipation and decoherence are discussed, concentrating on the example of the quantum linear Boltzmann equation and the unified description of quite different coherence experiments in terms of Lévy processes.

The purpose of the contribution by Robert Alicki is to present a general formalism based on first principles Hamiltonian models which can be used to describe mathematically and estimate numerically the influence of an environmental noise on quantum devices controlled by time-dependent external forces. Such systems are called controlled quantum open systems (CQOS). The basic idea is to eliminate the degrees of freedom of the environment and use approximative expressions for the reduced dynamics of the quantum system (which might be considered as a device used in QIPC). The main problem is related to the existence of multiple time scales in such models which makes the reduced dynamics very complicated. This necessitates the restriction to subdomains of the parameter domain and to use approximations. Several concrete cases are discussed. At first the author considers the Markov approximation for the case of a time-independent controlling Hamiltonian which leads to the formalism of completely positive quantum dynamical semi-groups with Lindblad–Gorini–Kossakowski–Sudarshan (LGKS) generators. Next the case of slowly varying external forces is considered which also leads to a time-inhomogeneous Markov Master equation of LGKS type. Finally, the non-Markovian Born approximation and the corresponding error formula are presented. This case is illustrated by a generic model of the controlled spin-boson system which in turn can be applied to several implementations of controlled qubits.

The contribution by Sugato Bose investigates how and when complex many-body quantum systems can be used in quantum information applications, in particular in quantum computing, with the motivation that an (ideal) quantum computer itself is the ultimate example of an engineered and controllable many-body quantum system. The first part describes the relevant notions of quantum information, i.e., quantum entanglement and its quantification in terms of concurrence, fidelity, and the operations needed for building a quantum computer, and concludes with the description of many-body spin systems as examples. Then it is examined whether many-body quantum systems can actually serve as a resource for truly "quantum" correlations or entanglement. The next part studies the possibility of using spin chains as channels for quantum communication. Here we learn first about a simple spin chain quantum communication protocol, then about a formula for the fidelity of this protocol and an estimate of its performance. This part ends with a discussion of how entanglement can be transmitted effectively in this context. The next lecture describes the various methods that have been suggested to perfect communication schemes which use spin chains, i.e., engineering the coupling in a spin chain, encoding in many-spin states, coupling qubits weakly to quantum many body systems, dual-rail

encoding and heralded perfect transfer, and controllable coupling of a memory qubit at the end of the transfer. The last lecture discusses two of the many proposals for quantum computation using many-body systems: quantum computation using a one-dimensional Heisenberg chain and realizing logic gates through engineered XY spin chains.

Heinz-Peter Breuer contributed an overview of efficient methods for the description of the non-Markovian dynamics of open quantum systems. The theoretical description of quantum mechanical relaxation and decoherence processes often leads to a non-Markovian dynamics which is determined by pronounced memory effects. This contribution reviews the systematic approach to non-Markovian dynamics which is known as projection operator technique. Accordingly, in the first part the standard operator method is described, recalling in particular Nakajima-Zwanzig and the time-convolution projection operator technique as well as the Markovian approximation of quantum dynamical semi-groups. The second part introduces quite a general class of projection super-operators that project onto correlated system-environment states and are therefore able to describe strong correlations and non-Markovian effects. In the last part, for the case of classically correlated projection super-operators, the non-Markovian generalization of the standard Lindblad equation is derived and some applications of this equation to models with pronounced non-Markovian effects are indicated.

The contribution by Beatrix Hiesmayr aims at showing that particle physics also contributes substantially to answering fundamental questions about quantum mechanics as used in QIPC, the main point being that this testing is done at scales of energies not available for usual quantum systems. The contribution focuses on the massive meson–antimeson system (K-mesons or Kaons) to describe testing of the foundations of quantum mechanics (QM), i.e., local realistic theories versus QM, Bell inequalities, quantum marking and eraser concepts, Bohr's complementary relations, and testing of some fundamental aspects of particle physics, i.e., CP-violation, CPT-violation. In particular, answers to the following three groups of questions are discussed. (1) Bell's inequality: Can one find experimental setups for testing local realistic theories versus QM? What has symmetry violation in particle physics to do with nonlocality? (2) Quantum erasers: Is "erasing the past and impacting the future" possible with K-mesons? Are there new eraser options which can be realized in the (near) future? Decoherence and loss of entanglement: Decoherence models for mesons? How do mesons "loose" certain quantum features, e.g., entanglement and how can this be measured? The constructive answers to all these questions are prepared by providing detailed background information and the important answer to one of these fundamental questions is that experiments confirm that "spooky action at a distance" takes place in the K-meson system.

Pieter Kok introduces the topical subject of optical quantum computing. This contribution thus focuses on an alternative approach to quantum computing as described by Sugato Bose. This is important since at the moment it is not yet known which physical system is best suited for making a quantum computer. This contribution explains some of the many proposals for using light for quantum com-

puting, concentrating on the cases of photonic qubits and leaving out the case of optical field states (continuous variables). Two cases of optical quantum computing are described: linear optical quantum computing using matter qubits and photons and quantum computing using optical nonlinearities (Kerr nonlinearities and Zeno gates). Three introductory sections introduce all the necessary concepts and devices and their theoretical description which are used in optical quantum computing. Thus in particular we learn about photons as qubits, phase shifts, (polarizing) beam splitters, interferometers, gates, the Knill, Laflamme, Milburn (KLM) approach, two-photon interference, circuits, clusters, engineering of clusters (with fusion gates) and complete quantum computer architecture.

Shigeaki Nagamachi describes teleportation of continuous quantum variables in full analogy to teleportation of qubits. In order to achieve this, the scheme of the experimental setup and the decisive steps of the basic experiments are discussed for both teleportation of qubits and teleportation of continuous quantum variables. In order to enable this analogy to come out fully, the author describes teleportation of continuous quantum variables not in terms of Wigner functions as all previous theoretical treatments do, but develops a mathematical description of the experimental devices which implement the teleportation in terms of the Bargmann (or holomorphic) representation of the canonical commutations relations. This step involves naturally some concepts and results of quantum mechanics on infinite-dimensional state spaces which in part is somewhat technical. In order to separate theory and technique, the mathematical proofs respectively calculations are presented in an appendix. Teleportation in a strict sense includes naturally the localization of the objects to be teleported. Accordingly this contribution concludes with a brief discussion of localization/locality aspect in teleportation.

The contribution by Daniel Terno looks beyond present day QIPC in so far as the basic steps of quantum information theory are introduced and discussed in the context of special relativity which faces some additional challenges. At first we learn about the causality constraints on quantum measurements and their description and the implications of these constraints in the realization of standard quantum information protocols. Naturally, in a relativistic context, some basic knowledge of the Poincaré group (semi-direct product of the space-time translation group with the isochronous proper Lorentz group) and its representations are needed which are developed in the second part of the contribution both for the case of massive particles and for photons. The next section discusses the effects of quantum Lorentz transformations on the basic ingredients of quantum information theory, i.e., reduced density matrices, massive qubits and photonic qubits. The effects of the constraints of Lorentz covariance on quantum communication channels is presented next. The following section explains, for the case of two-particle states, some of the difficulties one encounters when entanglement is to be detected in different Lorentz frames; for a proper analysis of these difficult questions it seems necessary to involve field theory to which this section refers. The final part of this contribution gives an overview of some recent developments which look beyond special relativity, in particular at problems involving (relativistic quantum) field theory and taking gravitation into account.

We would like to thank all contributors for the very interesting lectures and for spending much time discussing with the students at the School. Also, we would like to thank the junior members of the Quantum Research Group at the University of KwaZulu-Natal for helping out with the practical organization of the summer school.

Durban

Erwin Brüning
Francesco Petruccione

Contents

An Introduction to Quantum Probability 1
Mark Fannes
1 Motivation 2
2 Observables 7
3 States and Convexity 12
4 Composite Systems 22
5 Dynamics and Measurements 26
References 38

Covariant Mappings for the Description of Measurement, Dissipation and Decoherence in Quantum Mechanics 39
Bassano Vacchini
1 Introduction 39
2 Quantum Mechanics as Quantum Probability 41
3 Open Systems and Covariance 58
References 76

Quantum Open Systems with Time-Dependent Control 79
Robert Alicki
1 Introduction 79
2 The Generic Model of CQOS 80
3 Reduced Dynamics 81
4 Markovian Dynamics 83
5 Weak-Coupling Limit for Constant Hamiltonian 85
6 Adiabatic Control in the WCL 88
7 CQOP in Non-Markovian–Born Regime 90
8 Controlled Spin-Boson System 92
References 95

Five Lectures on Quantum Information Applications of Complex Many-Body Systems 97
Sougato Bose
1 Introduction 97
2 Lecture 1 99
3 Lecture 2 102
4 Lecture 3 107
5 Lecture 4 112
6 Lecture 5 118
References 123

Non-Markovian Quantum Dynamics and the Method of Correlated Projection Super-Operators 125
Heinz-Peter Breuer
1 Introduction 125
2 The Standard Projection Operator Method 126
3 Correlated Projection Super-Operators 132
4 Generalization of the Lindblad Equation 136
5 Conclusions 138
References 138

Testing Quantum Mechanics in High-Energy Physics 141
Beatrix C. Hiesmayr
1 Short Manual to Neutral Kaons 141
2 Bell Inequalities in High-Energy Physics 156
3 Kaonic Quantum Eraser 172
4 Decoherence and Measures of Entanglement 178
References 182

Five Lectures on Optical Quantum Computing 187
Pieter Kok
1 Light and Quantum Information 187
2 Two-Qubit Gates and the KLM Scheme 192
3 Cluster States 198
4 Quantum Computing with Matter Qubits and Photons 204
5 Quantum Computing with Optical Nonlinearities 210
References 218

Quantum Information and Relativity: An Introduction 221
Daniel R. Terno
1 Causality and Distributed Measurements 221
2 Quantum Lorentz Transformations 225
3 Implications of Quantum Lorentz Transformations 231
4 Communication Channels 240

5 Entanglement and Different Lorentz Observers 241
6 Beyond Special Relativity 243
References 246

Index 249

An Introduction to Quantum Probability

Mark Fannes

Using quantum mechanical devices for handling information is not a recent proposal. Recent experimental progress has, however, narrowed the gap with reality, even if there is still a very long way to go in order to achieve the most far-stretching proposals. Modern techniques now really allow to handle very small systems—the paradigms of the Gedankenexperimente in the standard textbooks on quantum mechanics—in an increasingly controlled manner. Because of the counter-intuitive aspects of quantum mechanics, experimental and theoretical aspects have to be developed more or less simultaneously, certainly in order to build complex systems. The renewed interest in theoretical and mathematical aspects of quantum theory is then a natural consequence.

Within the context of information, quantum mechanical devices for secure communication are already commercially available, but implementing the spectacular quantum algorithms seems still a very long-term perspective. To this end one has to exploit the typical quantum assets such as superposing states and efficiently use the unitary dynamics for composite systems with many parties. Randomizing influences of the environment tend to make large systems rather classical and put severe limitations on what is feasible. Another very important problem is to develop techniques for stably controlling the dynamics of such systems.

The aim of these lectures is to sketch a mathematical framework that is sufficiently broad to treat both classical and quantum systems in a uniform way. A uniform description is indeed appropriate; it is after all important to compare classical and quantum aspects of physical systems and to understand to what extent typical quantum features are relevant. These lectures will be on a low mathematical level, essentially finite dimensional. This suffices to introduce the relevant concepts and terminology. The main mathematical tools stem from linear algebra. Most results have, at some technical expense, natural extensions both to the case of standard quantum mechanics on separable Hilbert spaces and to the theory of systems with

M. Fannes (✉)
Instituut voor Theoretische Fysica, K.U.Leuven, Belgium
e-mail: mark.fannes@fys.kuleuven.be

Fannes, M.: *An Introduction to Quantum Probability*. Lect. Notes Phys. **787**, 1–38 (2010)
DOI 10.1007/978-3-642-02871-7_1 © Springer-Verlag Berlin Heidelberg 2010

an infinite number of degrees of freedoms which are studied in statistical mechanics or field theory. Many references at various levels of technicality can be found in, e.g. [1–5].

1 Motivation

Standard courses on classical mechanics and quantum mechanics look quite different. There are hardly common terms in both subjects. Textbooks on classical mechanics often concentrate on the different formulations of the theory: Newton's laws, the Lagrangian formalism, Hamilton's equations of motion, etc. A number of rather exceptional but important examples are analysed, such as the harmonic oscillator, Kepler's problem and the motion of the top. Some more recent texts consider stability issues and chaos. Many textbooks on quantum mechanics on the other hand put a lot of emphasis on wave mechanics solving, e.g. the stationary Schrödinger equation for a number of models. The quantum mechanical formalism is then presented, often in a not so precise mathematical way, in terms of operators, commutators, eigenfunctions and eigenvalues. The connection with classical mechanics is usually rather thin, mainly the correspondence principle and Ehrenfest's theorem. However, R. Finkelstein in his book *Nonrelativistic Mechanics* (Benjamin, Reading, MA, 1973) treats classical and quantum mechanics in parallel.

One of the aims of these lectures is to present a probabilistic model that allows to describe both the classical and the quantum theory. The main ingredients are the observables which define the kinematical structure of the system—what are the components of our system—and the states which should allow us to reproduce or predict the measurements on the system given a certain preparation procedure. Finally we also need the dynamics which describes how a system changes in time either due to its natural evolution or through the actions of an external agent. In order to make the comments above a bit more concrete we consider the basic example of a harmonic oscillator.

1.1 The Classical Harmonic Oscillator

The most basic oscillator is just a one-dimensional mass and spring system: a point particle of unit mass is attached to one end of an ideal spring of unit stiffness while the other end of the spring is attached to the origin. Pulling the mass away from the origin requires some work against the spring and increases therefore the potential energy of the mass and spring system. The mass, after being released, returns towards the origin and acquires some velocity which means that potential energy has been converted into kinetic energy. When it crosses the origin all potential energy has been converted into kinetic energy. The mass now moves away from the origin, gradually transferring its kinetic energy to the spring until it stands still having completed half a period of an oscillation. Without friction this periodic exchange between potential and kinetic energies goes on for ever; this is periodic motion.

The motion is governed by a single ordinary linear differential equation of the second order:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + x = 0\,. \tag{1}$$

Here x is the position of the point particle in Cartesian coordinates; it is a function of a variable t which is proportional to the time. The general solution of (1) is

$$x(t) = A \sin t + B \cos t\,.$$

The two integration constants A and B have to be adjusted by imposing suitable initial conditions. We could, e.g., specify the total energy of the system and the initial position.

The total energy E is the sum of the potential energy

$$V(x) = \frac{1}{2}\, x^2$$

of the spring and the kinetic energy

$$T = \frac{1}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2$$

of the point particle. A straightforward computation yields

$$E = \frac{1}{2} \left(A^2 + B^2\right).$$

Imposing that initially the mass is at the origin yields $B = 0$ and so

$$x(t) = \sqrt{2E}\, \sin t\,.$$

1.2 The Quantum Oscillator

All the information that we can access about an isolated one-dimensional quantum particle in an harmonic well is contained in its wave function ψ which depends in the usual position representation on Cartesian position coordinates and time. The evolution of ψ is governed by *Schrödinger's time-dependent equation*

$$\mathrm{i}\hbar\, \frac{\partial \psi}{\partial t} = H\, \psi = -\frac{\hbar^2}{2}\, \frac{\partial^2 \psi}{\partial x^2} + \frac{1}{2}\, x^2\, \psi\,. \tag{2}$$

To obtain a unique solution we must specify the full knowledge about the particle at $t = 0$, i.e. $\psi(x, 0) = \psi_0(x)$. An idealized initial condition could be a Gaussian

wave packet that minimizes Heisenberg's uncertainty relations. In the course of time such initial wave packets will move and typically spread both in position and in momentum.

The solution to (2) can be computed by using the linearity of the evolution equation. We first determine the standing wave solutions

$$\psi(x,t) = \mathrm{e}^{-\mathrm{i}E/\hbar}\,\varphi(x)$$

by solving *Schrödinger's time-independent equation*

$$H\,\varphi = -\frac{\hbar^2}{2}\,\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} + \frac{1}{2}\,x^2\,\varphi = E\,\varphi\,. \tag{3}$$

It turns out that (3) only admits acceptable solutions for a discrete set of values of E

$$E_n = \hbar(n+\tfrac{1}{2}), \qquad n = 0, 1, 2, \ldots$$

with corresponding functions

$$\varphi_n(x) = \frac{1}{2^{n/2}(n!)^{1/2}(\pi\hbar)^{1/4}}\,\mathrm{e}^{-x^2/2\hbar}\,\mathrm{H}_n\Big(\frac{x}{\sqrt{\hbar}}\Big)\,.$$

The function H_n is the Hermite polynomial of degree n and the normalization constant is chosen in such a way that

$$\int_{-\infty}^{\infty} \mathrm{d}x\;|\varphi_n(x)|^2 = 1\,.$$

In fact, any reasonable square integrable complex function ψ_0 on $\mathbb{R}$ can be written as an infinite linear combination of the φ_n:

$$\psi_0 = \sum_{n=0}^{\infty} \langle\varphi_n\,,\;\psi_0\rangle\,\varphi_n,$$

where

$$\langle\varphi_n\,,\;\psi_0\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\;\overline{\varphi_n(x)}\,\psi_0(x)\,.$$

Here $\overline{\varphi_n(x)}$ denotes the complex conjugate of the complex number $\varphi_n(x)$. This then yields the solution of time-dependent Schrödinger equation with initial value ψ_0:

$$\psi(x,t) = \sum_{n=0}^{\infty} \langle\varphi_n\,,\;\psi_0\rangle\,\mathrm{e}^{-\mathrm{i}E_n/\hbar}\,\varphi_n(x)\,.$$

This computation raises a number of questions about the meaning of the frequencies $E_n/\hbar$ and how to compare the quantum and classical oscillators.

1.3 Comparing Classical and Quantum Oscillators

Let us ask about the relative duration of stay $\rho(x)\,\Delta x$ of the classical oscillating particle of energy E in an interval of length Δx around x. For convenience we chose $E = 1/2$. Because the motion is periodic, it suffices to consider half a period, e.g.

$$x(t) = \sin t\,, \qquad -\frac{\pi}{2} \le t \le \frac{\pi}{2}\,.$$

Let the particle during this half period stay for a time Δt in the Δx vicinity of x, then

$$\begin{aligned}\rho(x)\,\Delta x &= \frac{1}{\pi}\,\Delta t \sim \frac{1}{\pi}\,\frac{\mathrm{d}t}{\mathrm{d}x}\,\Delta x \\ &= \frac{1}{\pi\cos t}\,\Delta x = \frac{1}{\pi\sqrt{1-x^2}}\,\Delta x\,.\end{aligned}$$

Hence ρ is a B-distribution:

$$\rho(x) = \frac{1}{\pi\sqrt{1-x^2}}\,, \qquad -1 < x < 1\,,$$

see Fig. 1.

This classical probability distribution of the position can easily be compared with its quantum counterpart. We first fix the energy of the quantum oscillator to the same value $E = \frac{1}{2}$

$$\hbar\big(n + \tfrac{1}{2}\big) = \tfrac{1}{2} \tag{4}$$

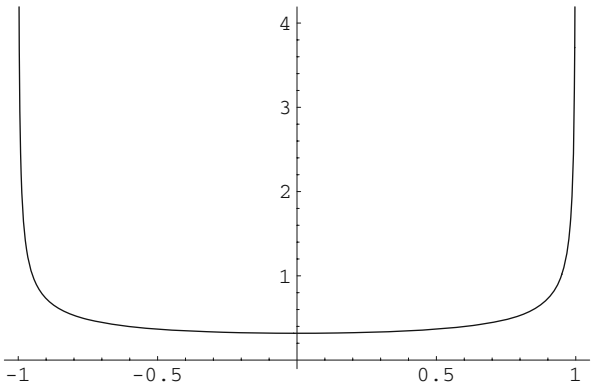


Fig. 1 Probability density of position of classical oscillator

and let then n tend to ∞. This means that we only consider energies that are much larger than the gaps in the spectrum of the quantum Hamiltonian. The position distribution is according to the rules of quantum mechanics

$$\rho_n(x) = |\varphi_n(x)|^2 = \frac{1}{2^n n! \sqrt{\pi\hbar}} \, \mathrm{e}^{-x^2/\hbar} \, \mathrm{H}_n^2\Big(\frac{x}{\sqrt{\hbar}}\Big),$$

with $\hbar$ satisfying (4).

From Figs. 2 and 3 it appears that the density of the quantum mechanical position distribution approaches on the average the classical density when n increases. The typical wave-like fringes remain however and there is certainly no pointwise convergence when $n \to \infty$. It is therefore more sensible to compare the distribution functions, i.e., the probability of finding the particle to the left of a point:

$$F(x) := \int_{-\infty}^{x} \mathrm{d}y \, \rho(y) \, .$$

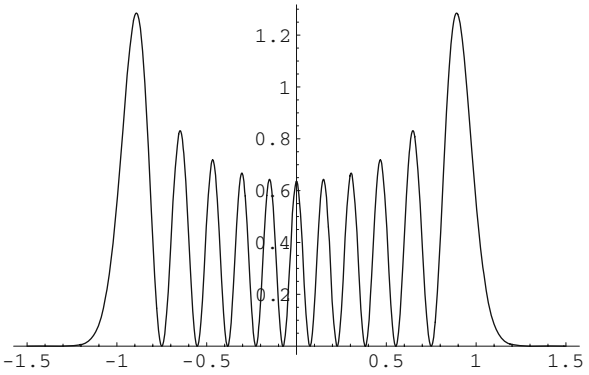


Fig. 2 Probability density of position of quantum oscillator $n = 10$

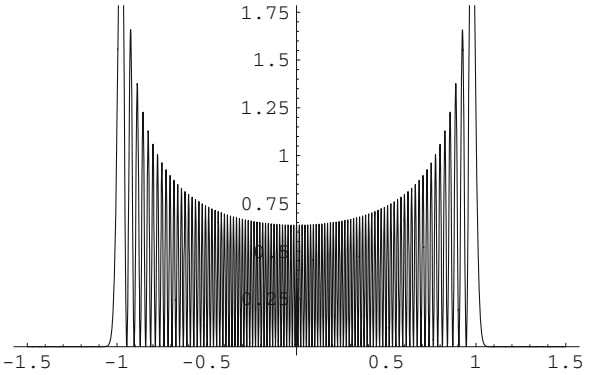


Fig. 3 Probability density of position of quantum oscillator $n = 100$

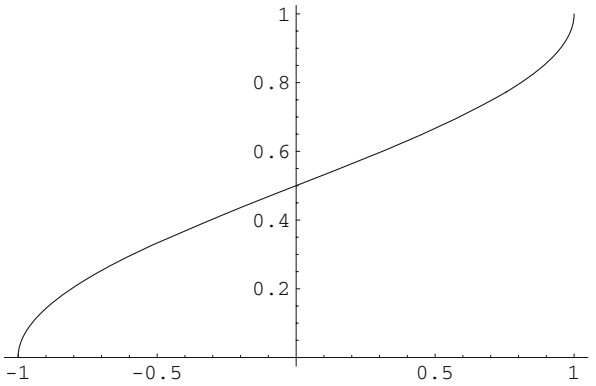


Fig. 4 Distribution function of position of classical oscillator

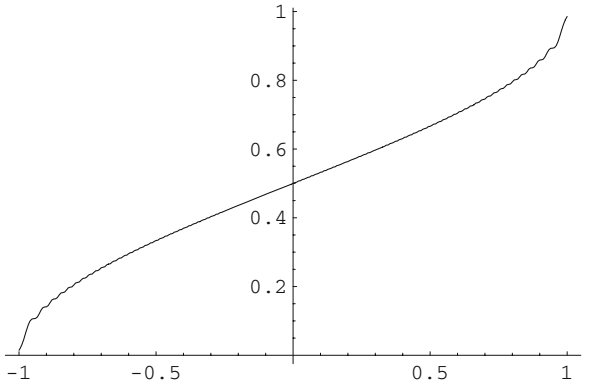


Fig. 5 Distribution function of position of quantum oscillator $n = 100$

In Figs. 4 and 5 the classical and quantum distribution functions are shown. The graphs look very similar except for some tiny wiggles in the quantum case.

We can draw from this example the lesson that we must compare classical and quantum systems at the level of expectations of observables such as position. Of course momentum would be an equally reasonable choice. We formalize this approach in the next section.

2 Observables

A probabilistic description always involves two aspects: the objects that have a random distribution and the actual probability distribution that assigns expectations to these objects. We shall further on rather use the terms *observables and states.*

The set $\mathcal{A}$ of observables of a physical system should specify its kinematical properties: Which particles are involved? What is their nature? What are the geometrical constraints? The following general assumptions lead to a workable model.

1. $\mathcal{A}$ is a *complex algebra*. This means that we can add and multiply observables and rescale them with complex scalars. These operations satisfy usual properties such as

$$A(B + \alpha\, C) = AB + \alpha\, AC\,, \qquad A, B, C \in \mathcal{A},\ \alpha \in \mathbb{C},$$

but the algebra product is generally non-commutative, i.e., we do not necessarily have $AB = BA$. We assume also that $\mathcal{A}$ has an *identity*, denoted by $\mathbb{1}$. This algebraic structure allows us to construct sufficiently many functions of observables. This is important because we want to model more than the average of an observable, also its variance, skewness, etc. The algebra must be complex to handle quantum systems where i always appears somewhere. Think, e.g., about the following common expressions in quantum mechanics:

$$\mathrm{i}\,\frac{\mathrm{d}\psi}{\mathrm{d}t} = H\,\psi, \qquad [P, Q] = -\mathrm{i}\hbar, \qquad [L_x, L_y] = \mathrm{i}\hbar L_z\,.$$

If we only deal with classical systems, then a real algebra suffices.

2. We need an *adjoint operation* $A \to A^*$ on $\mathcal{A}$ to distinguish real observables:

$$\left(A^*\right)^* = A\,, \qquad (A + \alpha\, B)^* = A^* + \overline{\alpha}\, B^*, \qquad (AB)^* = B^* A^*.$$

3. We need a notion of *positivity* to give a meaning to expressions like "the energy is bounded from below". This is tightly connected to a specific kind of norm, called a *C*-norm*. For $A, B \in \mathcal{A}$ and $\alpha \in \mathbb{C}$,

$$\begin{aligned} &\|A\| \geq 0\,, \qquad \|A\| = 0 \text{iff} A = 0\,, \qquad \|A + B\| \leq \|A\| + \|B\| \\ &\|AB\| \leq \|A\|\|B\|\,, \qquad \|\alpha A\| = |\alpha|\,\|A\| \\ &\|A^*\| = \|A\|\,, \qquad \|A^* A\| = \|A\|^2 \quad (C^* - \text{property}). \end{aligned}$$

These abstract properties single out a class of algebras that very naturally connects both with classical probability and with the Hilbert space formulation of quantum mechanics. If we put in the technical requirement that $\mathcal{A}$ is complete w.r.t. the norm, then we have a *C*-algebra* at our disposal.
There are several equivalent definitions of positivity in a C*-algebra, the simplest one being A is positive if it is of the form $B^* B$. The following results link positivity and norm:

$$A = B^* B \qquad \text{iff} \qquad A = A^* \text{and} \left\| \|A\|\mathbb{1} - A \right\| \leq \|A\|\,,$$

and conversely

$$\|A\|^2 = \min\{\lambda \in \mathbb{R} \mid \lambda \mathbb{1} - A^* A \geq 0\}\,.$$

2.1 Examples

Example 1 The continuous complex functions $\mathcal{C}_{\mathbb{C}}\big((0,1)\big)$ on $(0,1)$ equipped with point wise operations is a commutative algebra. The natural adjoint operation and the norm are given by

$$f^*(x) := \overline{f(x)} \qquad \text{and} \qquad \|f\| := \max\{|f(x)| \,:\, x \in (0,1)\}\,.$$

As a uniform limit of continuous functions on $(0,1)$ is still continuous $\mathcal{C}_{\mathbb{C}}\big((0,1)\big)$ is complete.

Example 2 The continuous linear transformations $\mathcal{B}(\mathcal{H})$ of a Hilbert space $\mathcal{H}$ with composition as algebra product is an algebra. It comes with the adjoint and norm

$$A^* := A^\dagger \qquad \text{and} \qquad \|A\| := \sup\{\|A\,\varphi\| \,:\, \varphi \in \mathcal{H},\ \|\varphi\| = 1\}\,.$$

In this expression $A^\dagger$ is the *Hermitian conjugate* of A which is defined through the relation

$$\langle \varphi\,,\ A^\dagger \psi\rangle = \langle A\,\varphi\,,\ \psi\rangle\,.$$

The algebra $\mathcal{B}(\mathcal{H})$ is complete but non-commutative, at least if $\dim(\mathcal{H}) > 1$.

Example 3 Often algebras are defined in terms of generators and their relations, e.g. the ("functions" on the) two-dimensional non-commutative torus. This is generated by $\mathbb{1}$ and U and V with

$$U U^* = U^* U = V V^* = V^* V = \mathbb{1} \qquad \text{and} \qquad U V = \mathrm{e}^{2\pi \mathrm{i} q}\, V U\,. \tag{5}$$

Here $0 \leq q < 1$ is called the deformation parameter. In the absence of deformation $q = 0$ one shows that the resulting algebra is isomorphic to the continuous complex functions on the usual two-dimensional torus. The basic generators U and V can be taken as the basic periodic functions

$$(\theta_1, \theta_2) \mapsto \mathrm{e}^{\mathrm{i}\theta_1} \qquad \text{and} \qquad (\theta_1, \theta_2) \mapsto \mathrm{e}^{\mathrm{i}\theta_2}\,.$$

The deformed relation (5) expresses that the two angles no longer commute.

One could wonder why not to stick with $\mathcal{B}(\mathcal{H})$ or $\mathcal{C}_{\mathbb{C}}(\Omega)$ where Ω is some nice configuration or phase space of a classical system. Some systems have, to begin with, both classical and quantum aspects at least in some approximations. Many

systems with a large number of particles are quite complicated. It is sometimes reasonable and simpler to replace them by systems with an infinite number of particles and which model the asymptotic behaviour of many particle systems. This is known as "taking the thermodynamic limit" in statistical mechanics or "having a property in an asymptotic sense" in quantum information. However, Hilbert spaces as $\otimes^{\mathbb{N}}\mathbb{C}^2$ are really not uniquely defined. So there is no natural Hilbert space around and selecting a particular realization of an infinite system might result in meaningless expressions or computations. A careful description in terms of observables, however, survives the thermodynamical limit and, together with the specification of a state, selects the right Hilbert space for the infinite system.

Two basic results should be mentioned. First of all, every Abelian C*-algebra with unit is isomorphic to the continuous complex functions on a (compact Hausdorff) space Ω. This is *Gelfand's theorem*. Moreover, Ω is unique up to homeomorphisms. Next, every C*-algebra with unit is isomorphic to a (norm closed) self-adjoint sub-algebra with unit of $\mathcal{B}(\mathcal{H})$ for a suitable $\mathcal{H}$.

The classification of general C*-algebras is far from complete but a lot is known about algebras that are almost finite dimensional. This already suffices to treat many interesting infinite systems such as spin lattice systems. We shall only consider finite-dimensional algebras here meaning algebras that are finite dimensional if we consider here only their vector space structure.

2.2 Matrix Algebras

This section is a brief reminder of linear algebra mainly with the purpose of introducing some notation needed further on.

The d-dimensional complex vector space $\mathbb{C}^d$ consists of all column vectors of complex numbers with height d. To save space we write them as transposes of row vectors

$$\varphi^{\mathsf{T}} = (\varphi_1, \varphi_2, \ldots, \varphi_d)^{\mathsf{T}}, \qquad \varphi_j \in \mathbb{C}\,.$$

Addition and scalar multiplication are componentwise and we have a natural complex scalar product

$$\langle \varphi\,,\ \psi\rangle := \sum_{j=1}^{d} \overline{\varphi_j}\,\psi_j\,.$$

The canonical basis in $\mathbb{C}^d$ will be denoted by $\{e_j\ :\ j = 1, 2, \ldots, d\}$ where all entries of e_j are zero except for a one on row j, obviously

$$\varphi = \sum_{j=1}^{d} \langle e_j\,,\ \varphi\rangle\, e_j\,.$$

Linear transformations of $\mathbb{C}^d$ are given by square matrices of dimension d:

$$(A\,\varphi)_i = \sum_{j=1}^{d} A_{ij}\,\varphi_j, \qquad \varphi \in \mathbb{C}^d\,.$$

Here A_{ij} is the entry of A on row i and column j. Adding, composing and rescaling linear maps define the usual addition, product and rescaling of the corresponding matrices, e.g.

$$(AB)_{ij} = \sum_{k=1}^{d} A_{ik}\,B_{kj}\,.$$

We use the notation $\mathcal{M}_d$ for the algebra of d-dimensional complex matrices.

The Hermitian conjugate $A^\dagger$ is easily computed using

$$\langle \varphi\,,\ A^\dagger\psi\rangle = \langle A\varphi\,,\ \psi\rangle \qquad \text{and so} \qquad (A^\dagger)_{ij} = \overline{A_{ji}}\,.$$

There are various ways to compute the norm of A:

$$\begin{aligned}\|A\| &= \max\{\|A\,\varphi\| \,:\, \varphi = 1\}\\ &= \text{largest singular value of} A\\ &= \text{square root of largest eigenvalue of } A^\dagger A\,.\end{aligned}$$

The same holds for checking positivity of A:

- $A = A^\dagger$ and all eigenvalues of A are non-negative,
- there exists a B such that $A = B^\dagger B$,
- $< \varphi,\ A\,\varphi\rangle \geq 0$ for all $\varphi \in \mathbb{C}^d$,
- $A = A^\dagger$ and all determinants of sub-matrices of A are non-negative,
- $A = A^\dagger$ and all elementary symmetric invariants (Schur polynomials) of A are non-negative.

Example 4 The positivity condition for a qubit matrix reads

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \geq 0 \qquad \text{iff} \qquad a_{21} = \overline{a_{12}},\ a_{11} \geq 0 \text{and} |a_{12}|^2 \leq a_{11}a_{22}\,.$$

The *standard matrix units* $\{e_{ij} \,:\, i, j = 1, 2, \ldots, d\}$ are often useful. All entries of the matrix e_{ij} are zero except for a one on row i and column j. The following properties hold:

$$e_{ij}e_{k\ell} = \delta_{jk}e_{i\ell}, \quad e_{ij}^\dagger = e_{ji}, \quad \textstyle\sum_{j=1}^{d} e_{jj} = \mathbb{1},$$

$$\textstyle\sum_{i,j=1}^{d} A_{ij}\,e_{ij} = A, \quad \operatorname{tr} e_{ij} = \delta_{ij}\,.$$

2.3 Finite Dimensional Algebras of Observables

Every finite-dimensional C*-algebra $\mathcal{A}$ is isomorphic to a *direct sum* of matrix algebras

$$\mathcal{A} = \oplus_j \mathcal{M}_{d_j} .$$

The elements of $\mathcal{A}$ can therefore be thought of as *block diagonal matrices*

$$A = \begin{bmatrix} A_1 & 0 & \cdots \\ & \ddots & \\ 0 & \cdots & A_n \end{bmatrix}, \qquad A_j \text{ matrix of dimension } d_j .$$

The adjoint and norm are given by

$$A^* = \begin{bmatrix} A_1^\dagger & 0 & \cdots \\ & \ddots & \\ 0 & \cdots & A_n^\dagger \end{bmatrix} \qquad \text{and} \qquad \|A\| = \max\{\|A_j\| , \ j\}.$$

Moreover, a block diagonal matrix is positive iff all its entries are positive.

Some particular cases are important. If all d_j are one dimensional, then $\mathcal{A}$ is commutative, so it is a function on some space. The space is just the set of indices $\{j\}$ and the values of the function are tabulated on the diagonal. This is the pure (discrete) classical case, e.g. describing a classical register. If there is only a single block with dimension greater than one, then we are in the pure quantum case; the algebra just describes a qubit, qutrit, or qudit.

3 States and Convexity

A *state* is an expectation functional on the observables

$$A \in \mathcal{A} \mapsto \omega(A) \in \mathbb{C}.$$

Here $\omega(A)$ is the average of the observed values of A if the system is perfectly prepared in the state ω. Unlike for classical systems, measuring repeatedly identically prepared quantum systems will typically produce random outcomes. This randomness is intrinsic to quantum mechanics. So $\omega(A)$ gives only the average value of the observable A over many repeated measurements. To describe the distribution of the outcomes we need also to know $\omega(A^2)$, $\omega(A^3)$, etc. The following assumptions are essential to allow for a probabilistic interpretation of the theory:

$$\omega(A + \alpha\, B) = \omega(A) + \alpha\, \omega(B), \quad \omega(\mathbb{1}) = 1, \quad \text{and} \quad \omega(A) \geq 0 \text{if} A \geq 0\,.$$

The set of states on $\mathcal{A}$ or *state space* of $\mathcal{A}$ will be denoted by $\mathcal{S}(\mathcal{A})$ and $(\mathcal{A}, \omega)$ is called a *quantum probability space*, $\mathcal{A}$ being the random variables and ω the probability measure.

We list a few basic inequalities for states; their proofs essentially follow from positivity:

$$\begin{aligned} &\omega(A^*) = \overline{\omega(A)}, \\ &|\omega(A^* B)|^2 \leq \omega(A^* A)\, \omega(B^* B), \\ &|\omega(A^* B\, A)| \leq \|B\|\, \omega(A^* A)\,. \end{aligned} \tag{6}$$

Example 5 Let $\mathcal{A} = \mathcal{C}_{\mathbb{C}}(\Omega)$ and $\omega \in \mathcal{S}(\mathcal{A})$, then there is a unique *probability measure* μ on Ω such that

$$\omega(f) = \int_{\Omega} \mu(\mathrm{d}\omega)\, f(\omega)\,.$$

This is known as *Riesz's representation theorem*; it shows that for Abelian algebras the notion of quantum probability space simply coincides with the usual probability space.

3.1 States on Finite-Dimensional Algebras

Example 6 A *density matrix* of dimension d is a positive (semi-definite) matrix with trace equal to one. We shall now show that every state on $\mathcal{M}_d$ is in one-to-one correspondence with a *density matrix* through

$$\omega(A) = \mathrm{tr}\,(\rho\, A)\,. \tag{7}$$

That (7) defines a state on $\mathcal{M}_d$ is rather obvious; linearity and normalization are immediate and if A is positive, then it can be written as $B^* B$ and

$$\mathrm{tr}\,(\rho\, A) = \mathrm{tr}\,(\rho\, B^\dagger B) = \mathrm{tr}\,(B\, \rho\, B^\dagger) \geq 0\,.$$

The last inequality is true because $B\, \rho\, B^\dagger$ is positive and the trace of a positive matrix is positive.

Conversely, suppose that ω is a state on $\mathcal{M}_d$ and consider the matrix

$$\rho := \sum_{i,j=1}^{d} \omega(e_{ji})\, e_{ij}\,.$$

For an arbitrary matrix unit $e_{k\ell}$

$$\begin{aligned}\mathrm{tr}\,(\rho\, e_{k\ell}) &= \mathrm{tr}\Big(\sum_{i,j=1}^{d} \omega(e_{ij})\, e_{ji}\Big)\, e_{k\ell} = \sum_{i,j=1}^{d} \omega(e_{ij})\mathrm{tr}\,(e_{ji}\, e_{k\ell}) \\ &= \sum_{i,j=1}^{d} \omega(e_{ij})\, \delta_{ik}\, \mathrm{tr}\, e_{j\ell} = \sum_{i,j=1}^{d} \omega(e_{ij})\, \delta_{ik}\, \delta_{j\ell} = \omega(e_{k\ell})\,.\end{aligned}$$

As any matrix A is a linear combination of matrix units,

$$\omega(A) = \mathrm{tr}\,(\rho\, A)\,.$$

We still have to show that ρ is a density matrix. Normalization follows from $\omega(\mathbb{1}) = 1$ and positivity from

$$\langle\varphi\,,\ \rho\,\varphi\rangle = \mathrm{tr}\,(\rho\,|\varphi\rangle\langle\varphi|) = \omega(|\varphi\rangle\langle\varphi|) \geq 0\,.$$

Here $|\varphi\rangle\langle\varphi|$ is the usual notation for the transformation $\psi \mapsto \langle\varphi\,,\ \psi\rangle\,\varphi$ which is positive.

We are now able to describe a general state on a finite-dimensional algebra. Let

$$\mathcal{A} = \oplus_j \mathcal{M}_{d_j}.$$

Then any state ω on $\mathcal{A}$ is given by a set $\{(\mu_j, \rho_j)\}$ where (μ_j) is a *sequence of probabilities* , i.e. $\mu_j \geq 0$ and $\sum_j \mu_j = 1$, and where every ρ_j is a density matrix on $\mathcal{M}_{d_j}$; more precisely

$$\omega\Big(\oplus_j A_j\Big) = \sum_j \mu_j\, \mathrm{tr}\,(\rho_j\, A_j)\,.$$

The set $\{(\mu_j, \rho_j)\}$ is called a *quantum!ensemble*.

3.2 Recovering the Hilbert Space Picture

We shall now show how one can recover the usual quantum mechanical rule for computing expectations, namely sandwiching observables between wave functions. For the sake of simplicity we suppose that our state $\omega \in \mathcal{S}(\mathcal{A})$ is faithful meaning that $\omega(A^*A) = 0$ implies $A = 0$. This restriction can be lifted by taking appropriate quotients. The construction that we sketch is quite general; it goes under the name GNS (*Gelfand–Naimark–Segal representation theorem*) and is unique up to unitary isomorphism.

First $\mathcal{A}$ is turned into an inner product space $\mathcal{H}_\omega$ by introducing the scalar product

$$\langle A\,,\ B\rangle_\omega := \omega(A^*B)\,.$$

We now compute using (6) and the C*-property of the norm defined by the scalar product $\langle\cdot\,,\ \cdot\rangle_\omega$:

$$\|A\,B\|_\omega^2 = \omega(B^*A^*AB) \le \|A\|^2\,\omega(B^*B) = \|A\|^2\,\|B\|_\omega^2\,.$$

This inequality implies that left multiplication

$$\pi_\omega(A) : B \mapsto A\,B$$

is a well-defined linear transformation of the space $\mathcal{H}_\omega$. Moreover, π_ω is a *representation* of the algebra in the linear transformations of $\mathcal{H}_\omega$:

$$\begin{aligned}
&\pi_\omega(A+\alpha\,B) = \pi_\omega(A) + \alpha\,\pi_\omega(B),\\
&\pi_\omega(A\,B) = \pi_\omega(A)\,\pi_\omega(B),\\
&\pi_\omega(A^*) = \pi_\omega(A)^\dagger\,.
\end{aligned}$$

Finally, putting $\Omega_\omega := \mathbb{1}$

$$\begin{aligned}
&\omega(A) = \langle\Omega_\omega\,,\ \pi_\omega(A)\,\Omega_\omega\rangle_\omega,\\
&\pi_\omega(\mathcal{A})\,\Omega_\omega = \mathcal{H}_\omega\,.
\end{aligned} \tag{8}$$

Equation (8) is the usual quantum mechanical rule for computing the expectation of an observable. The main modification is that we have to replace A by $\pi_\omega(A)$. The following explicit GNS construction for a state on a matrix algebra illustrates this remark.

Example 7 Let ρ be the density matrix of dimension d and suppose that all eigenvalues of ρ are strictly positive. This is equivalent to asking that the state corresponding to ρ is faithful. Write now an eigenvalue decomposition of ρ:

$$\rho = \sum_{j=1}^{d} r_i\,|\varphi_i\rangle\langle\varphi_i|\,.$$

We explicitly choose the GNS objects as

$$\begin{aligned}
&\mathcal{H}_\omega := \mathbb{C}^d \otimes \mathbb{C}^d,\\
&\Omega_\omega := \sum_{j=1}^{d} \sqrt{r_j}\,\varphi_j \otimes \varphi_j,\\
&\pi_\omega(A) := A \otimes \mathbb{1}\,.
\end{aligned}$$

It is now a straightforward exercise to verify that these definitions satisfy the requirements of the GNS theorem. This construction shows that we can turn any state on $\mathcal{M}_d$ into a vector state; the price to pay is that we have to enlarge the space by introducing an *ancillary space* (in this case $\mathbb{C}^d$). If some of the eigenvalues of ρ are zero, then the ancillary space can be smaller. In fact the dimension of the ancillary space can be taken equal to the number of non-zero eigenvalues of ρ taking multiplicities into account.

3.3 Convex Subsets of $\mathbb{R}^n$

The state space of an algebra has a natural *convex structure*: we can mix states using a probability vector (μ_j), meaning that $\sum_j \mu_j \omega_j$ is again a state if all the ω_j are. For simplicity we assume that $\mathcal{A}$ is finite dimensional, then also $\mathcal{S}(\mathcal{A})$ is a convex subset of a finite-dimensional real space. Moreover $\mathcal{S}(\mathcal{A})$ is bounded and closed, i.e. compact. We briefly remind some general notions and results about such sets.

A subset $\mathcal{K}$ of $\mathbb{R}^n$ is *convex* if the line segments joining arbitrary pairs of points of $\mathcal{K}$ lie in $\mathcal{K}$. If $\mathcal{X}$ is an arbitrary subset of $\mathbb{R}^n$, then we can form its *convex hull*:

$$\mathrm{Conv}(\mathcal{X}) := \left\{ \sum_j \mu_j x_j \,:\, x_j \in \mathcal{X} \text{ and } (\mu_j) \text{ a probability vector} \right\}.$$

We can consider the *convex boundary* $\partial_c\mathcal{K}$ of a closed and compact subset $\mathcal{K}$ of $\mathbb{R}^n$. It is the set of *extreme points* of $\mathcal{K}$, namely the points $k \in \mathcal{K}$ that admit only trivial decompositions in $\mathcal{K}$. More explicitly, k is extreme if $k = \lambda k_1 + (1-\lambda)k_2$ with $0 < \lambda < 1$ and $k_1, k_2 \in \mathcal{K}$ implies that $k_1 = k_2 = k$. $\partial_c\mathcal{K}$ can be a proper subset of the topological boundary $\partial\mathcal{K}$. In Fig. 6 a compact subset of $\mathbb{R}^2$ is shown; the convex boundary consists of an arc and a single point, both thickly marked.

We can now formulate two basic structural results. The *theorem of Krein and Milman* states that a compact convex set $\mathcal{K}$ (in $\mathbb{R}^n$) is the (closure of the) convex hull of its extreme points:

$$\mathcal{K} = \overline{\mathrm{Conv}\left(\partial_c\mathcal{K}\right)}.$$

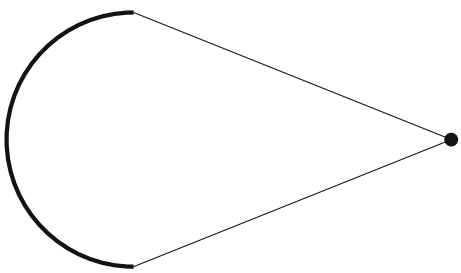


Fig. 6 A compact convex subset of $\mathbb{R}^2$

Carathéodory's theorem states that each point of a convex compact subset of $\mathbb{R}^n$ is a mixture of at most $n+1$ points.

A distinguished class of convex compact subsets of $\mathbb{R}^n$ are the *simplices* characterized by the property that each point of the simplex has a unique convex decomposition in extreme points. In one dimension simplices are closed intervals, in two dimensions triangles (see Fig. 7), in three dimensions tetrahedrons (see Fig. 8), etc. For a simplex one typically needs the maximal number of extreme points as stated in Carathéodory's theorem. We now give some useful examples of convex sets and their extreme boundaries.

Example 8 (Bistochastic matrices) A square matrix $[c_{ij}]$ of dimension d is called *bistochastic* if for all k and ℓ

$$c_{k\ell} \geq 0 \qquad \text{and} \qquad \sum_{i=1}^{d} c_{i\ell} = \sum_{j=1}^{d} c_{kj} = 1\,. \tag{9}$$

These matrices define a particular class of Markov processes; the entry c_{ij} specifies the transition probability of a classical particle from the state i to the state j. Conditions (9) amount to both conservation of particles and invariance of the uniform probability distribution. *Birkhoff's theorem* shows that the extreme boundary of the bistochastic matrices is the set of permutation matrices.

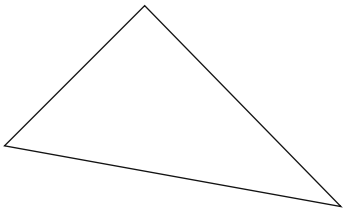


Fig. 7 A simplex in $\mathbb{R}^2$

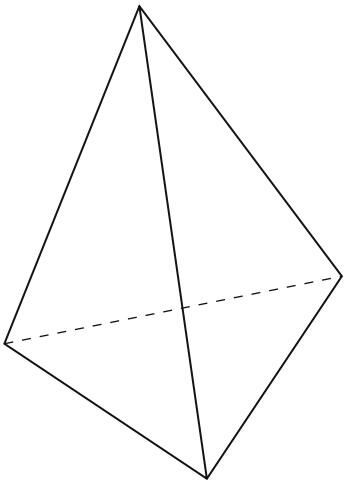


Fig. 8 A simplex in $\mathbb{R}^3$

Example 9 (Stochastic order) Given two real sequences $\mu = (\mu_j)$ and $\nu = (\nu_j)$ of length d such that

$$\sum_{j=1}^{d} \mu_j = \sum_{j=1}^{d} \nu_j \,,$$

μ stochastically dominates or majorizes ν, $\mu \succ \nu$, if for all k

$$\max\{\mu_{j_1} + \cdots + \mu_{j_k} \,:\, j_1 < \cdots < j_k\} \geq \max\{\nu_{j_1} + \cdots + \nu_{j_k} \,:\, j_1 < \cdots < j_k\}\,.$$

Rado's theorem characterizes this order relation: $\mu \succ \nu$ iff ν is a mixture of permutations of μ. This justifies that the terminology ν is *more mixed* than μ.

Stochastic ordering is quite useful to compare quantities like the *Shannon entropies* of measures

$$\mathsf{H}(\mu) := -\sum_{j=1}^{d} \mu_j \ln \mu_j \,.$$

From its definition H is concave:

$$\mathsf{H}(\lambda\mu_1 + (1-\lambda)\mu_2) \geq \lambda \mathsf{H}(\mu_1) + (1-\lambda)(\mu_2), \quad 0 \leq \lambda \leq 1\,.$$

Suppose now that $\mu \succ \nu$; then by Rado's theorem ν is a mixture of permutations of μ. As each permutation of μ has the same Shannon entropy as μ we obtain

$$\mathsf{H}(\nu) \geq \mathsf{H}(\mu)\,.$$

Hence more mixed measures have a higher entropy.

Example 10 (Perturbation of eigenvalues) Let A be an Hermitian matrix with ordered eigenvalues $\lambda_1^A \geq \cdots \geq \lambda_d^A$ repeated according to multiplicity. The *minimax principle* gives a variational method for computing eigenvalues

$$\lambda_k^A = \max\Big\{\min\{\langle\varphi, A\,\varphi\rangle \,:\, \varphi \in \mathcal{K},\ \|\varphi\| = 1\} \,:\, \mathcal{K} \subset \mathbb{C}^d,\ \dim(\mathcal{K}) = k\Big\}\,.$$

Suppose now that $B = B^*$, $C = C^*$ and $A = B + C$; then one shows using the mini–max principle that $\{\lambda_j^A - \lambda_j^B\}$ is majorized by $\{\lambda_j^C\}$ and so there exists a bistochastic $[\gamma_{ij}]$ such that

$$\lambda_j^A = \lambda_j^B + \sum_i \gamma_{ij}\, \lambda_{ij}^C. \tag{10}$$

This is *Lidskii's theorem* which can be used to compare the eigenvalues of $A + B$ with those of A and B. It is not hard to see that this result is optimal in the follow-

ing sense: given two sequences of eigenvalues and a bistochastic matrix, matrix γ, we can always find Hermitian A and B whose eigenvalues coincide with the given sequences such that the eigenvalues of $C := A + B$ satisfy (10).

Example 11 (Unit ball in $\mathcal{M}_d$) The extreme points of the unit ball in $\mathcal{M}_d$, i.e. of $\{A \in \mathcal{M}_d : \|A\| \leq 1\}$, are the unitaries. By using the singular value decomposition of a matrix we can write

$$A = U\, D\, V,$$

where U and V are unitary and D is diagonal and positive. As $\|A\| \leq 1$ the diagonal elements (δ_j) of D all belong to $(0, 1)$. Any vector $(\delta_1, \ldots, \delta_d)$ with all $\delta_j \in (0, 1)$ is a mixture of vectors of the type $(\pm 1, \ldots, \pm 1)$ and each such vector defines a diagonal unitary. Therefore A is also a mixture of unitaries.

Conversely, suppose that a unitary U allows a convex decomposition $U = \lambda A + (1 - \lambda)B$ with A and B in the unit ball. For any vector φ we then write

$$\begin{aligned}\|\varphi\| = \|U\,\varphi\| &= \big\|(\lambda A + (1-\lambda)B)\,\varphi\big\| \\ &\leq \lambda\|A\,\varphi\| + (1-\lambda)\|B\,\varphi\| \leq \lambda\|\varphi\| + (1-\lambda)\|\varphi\| = \|\varphi\|\,.\end{aligned}$$

It follows that

$$\|A\,\varphi\| = \|B\,\varphi\| = \|\varphi\|$$

and hence both A and B have to be unitary. Finally from the strict convexity of the norm on $\mathbb{C}^d$ we conclude that $A = B = U$.

Example 12 (Unit interval in $\mathcal{M}_d$) The unit interval $(0, \mathbb{1})$ in $\mathcal{M}_d$ consists of $\{A \in \mathcal{M}_d : 0 \leq A \leq \mathbb{1}\}$. This convex compact set is important, e.g. for Fermionic systems. We show that the extreme boundary consists of the projectors. Write the eigenvalue decomposition of $0 \leq A \leq \mathbb{1}$

$$A = \sum_{j=1}^{k} \lambda_j\, P_j$$

with $\lambda_1 > \lambda_2 > \cdots > \lambda_k > 0$. Generally k could be less than d as some eigenvalues of A might be degenerate. We now rewrite A as follows:

$$\begin{aligned}A = \lambda_d(P_1 + \cdots + P_d) + (\lambda_{d-1} - \lambda_d)(P_1 + \cdots + P_{d-1}) \\ + \cdots + (\lambda_1 - \lambda_2)P_1 + (1 - \lambda_1)0.\end{aligned}$$

Hence A is a mixture of projectors.

Conversely suppose that $P = \frac{1}{2}\,A + \frac{1}{2}\,B$ with A and B in the unit interval. For a vector φ in the range $\mathrm{Ran}(P)$ of the projector P, i.e. $\varphi \in \mathrm{Ran}(P)$, we have

$$\varphi = P\,\varphi = \tfrac{1}{2}\,A\,\varphi + \tfrac{1}{2}\,B\,\varphi.$$

Hence $A\,\varphi = B\,\varphi = \varphi$ by the strict convexity of the unit ball in $\mathbb{C}^n$. If φ belongs to the nullspace or kernel of P, i.e. $\varphi \in \mathrm{Ker}(P)$, then also $\varphi \in \mathrm{Ker}(A)$ and $\varphi \in \mathrm{Ker}(B)$ by positivity. Hence $A = B = P$.

3.4 Pure and Mixed States

A state on an algebra $\mathcal{A}$ is called *pure* if it is extreme, i.e. if it cannot be decomposed in a non-trivial way in a convex decomposition of other states. A state that is not pure is called *mixed*.

Let $\Omega = \{1, 2, \ldots, d\}$ be a finite set; then the pure states on $\mathcal{C}_{\mathbb{C}}(\Omega)$ are the degenerate or Dirac measures $\{\delta_j : j \in \Omega\}$. Every probability measure μ on Ω can be written in a unique way as

$$\mu = \sum_j \mu_j\,\delta_j.$$

Hence the state space of such a classical system is a simplex and a probability measure can be seen as an ensemble. Conversely, a simplex in $\mathbb{R}^d$ is the pure state space of a classical register with d positions. This example is of course trivial and can be vastly extended.

Let us now consider a discrete purely quantum system and show that *the pure states on $\mathcal{M}_d$ are in one-to-one correspondence with the one-dimensional subspaces of $\mathbb{C}^d$*, i.e. with the one-dimensional projectors on $\mathbb{C}^d$. We have seen in (24) that we may identify the state space on $\mathcal{M}_d$ with the set of d-dimensional density matrices. Writing the eigenvalue decomposition of a density matrix

$$\rho = \sum_i r_i\,|\varphi_i\rangle\langle\varphi_i|,$$

it follows that ρ can only be extreme if all $r_i = 0$ except for one. This means that ρ has to be a one-dimensional projector. Conversely, suppose that there are density matrices ρ_1 and ρ_2 such that

$$|\varphi\rangle\langle\varphi| = \lambda\,\rho_1 + (1-\lambda)\,\rho_2, \quad \text{with} 0 < \lambda < 1\,.$$

Choose $\psi \perp \varphi$; then

$$\begin{aligned} 0 &= \mathrm{tr}\left(|\varphi\rangle\langle\varphi|\,|\psi\rangle\langle\psi|\right) \\ &= \lambda \mathrm{tr}\left(\rho_1\,|\psi\rangle\langle\psi|\right) + (1-\lambda)\,\mathrm{tr}\left(\rho_2\,|\psi\rangle\langle\psi|\right) \\ &= \lambda\,\langle\psi\,,\ \rho_1\,\psi\rangle + (1-\lambda)\,\langle\psi\,,\ \rho_2\,\psi\rangle. \end{aligned}$$

So, $\langle\psi\,,\ \rho_i\,\psi\rangle = 0$ and therefore $\rho_i = |\varphi\rangle\langle\varphi|$.

The topological boundary of $\mathcal{S}(\mathcal{M}_d)$ can easily be characterized as well: ρ *belongs to the boundary of the state space iff* $\mathrm{Det}(\rho) = 0$. Suppose first that $\mathrm{Det}(\rho) > 0$; then all eigenvalues of ρ are strictly positive. Let us now consider a small open neighbourhood of ρ; we may in fact restrict ourselves to

$$\{\rho + \sigma \,:\, \sigma = \sigma^\dagger,\ \mathrm{tr}\,\sigma = 0 \text{and} \|\sigma\| \leq \epsilon\}\,.$$

It follows from perturbation theory, e.g. from Lidskii's theorem, that the eigenvalues of $\rho + \sigma$ are within distance ϵ from the eigenvalues of ρ and hence still positive for ϵ sufficiently small. Therefore ρ belongs to the interior of the density matrices. Suppose conversely that $\mathrm{Det}(\rho) = 0$; then at least one eigenvalue of ρ is equal to 0. This implies that however small $\epsilon > 0$, there always exists a σ with $\sigma = \sigma^\dagger$, $\mathrm{tr}\,\sigma = 0$ and $\|\sigma\| \leq \epsilon$ such that $\rho + \sigma$ has at least one strictly negative eigenvalue. We have therefore shown that ρ lies on the boundary of the state space.

Let us make a parameter count to have a feeling for the size of the extreme and topological boundaries. As a pure state on $\mathcal{M}_d$ corresponds to a one-dimensional subspace of $\mathbb{C}^d$, we have to count how many real parameters we need to label an arbitrary one-dimensional subspace of $\mathbb{C}^d$. Such a subspace is uniquely labelled by a vector

$$\left(\varphi_1, \varphi_2, \ldots, \varphi_{d-1}, \sqrt{1 - |\varphi_1|^2 - \cdots - |\varphi_{d-1}|^2}\right)^{\mathsf{T}}\,.$$

This calls for $2d - 2$ real parameters. For the topological boundary we have to impose two real conditions on an arbitrary Hermitian matrix, namely trace one and determinant zero. This leaves us with $d^2 - 2$ real parameters. Except for dimension two we have much less states in the extreme boundary than in the full boundary; this means that the topological boundary contains many flat parts.

Example 13 (Qubit states) Although a bit misleading it is useful to examine the state space of $\mathcal{M}_2$. Any element in $\mathcal{M}_2$ is a linear combination of the identity matrix and the three Pauli matrices $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$:

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma^y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \qquad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

In particular, any matrix ρ in $\mathcal{M}_2$ can be written as

$$\rho = x_0 \mathbb{1} + \boldsymbol{x} \cdot \boldsymbol{\sigma}\,,$$

where $x_0 \in \mathbb{C}$ and $\boldsymbol{x} \in \mathbb{C}^3$. If ρ is a density matrix then, because $\mathrm{tr}\,\rho = 1$, we have $x_0 = 1/2$. Moreover, ρ is positive definite which implies that ρ is Hermitian, so $\boldsymbol{x} \in \mathbb{R}^3$. Finally, we still have to require that $\mathrm{Det}(\rho) \geq 0$ in order to have both eigenvalues non-negative (we know already that the sum of the eigenvalues is equal to one). This allows us to write that

$$\rho = \tfrac{1}{2}\left(\mathbb{1} + \boldsymbol{r} \cdot \boldsymbol{\sigma}\right) \text{with}\, \boldsymbol{r} \in \mathbb{R}^3 \text{ and } \|\boldsymbol{r}\| \leq 1\,.$$

We have shown that the state space of $\mathcal{M}_2$ is affinely isomorphic to the unit ball in $\mathbb{R}^3$: the *Bloch ball*. This is not at all a simplex; every mixed state has infinitely many decompositions in extreme states. The extreme and topological boundaries coincide in this case but is a rather misleading accident. A pure state on $\mathcal{M}_2$ corresponds to a point with polar angles (θ, φ) on the Bloch sphere, i.e. $\|\mathbf{r}\| = 1$. The corresponding one-dimensional subspace of $\mathbb{C}^2$ is spanned by the (column) vector

$$\left(\cos(\tfrac{1}{2}\,\theta), \mathrm{e}^{\mathrm{i}\varphi} \sin(\tfrac{1}{2}\,\theta)\right)^{\mathrm{T}}.$$

The structure of the state space of a finite-dimensional algebra

$$\mathcal{A} = \oplus_j \mathcal{M}_{d_j}$$

should by now be clear. The pure states live on one of the terms in the direct sum and are given by a one-dimensional projector in that term.

4 Composite Systems

There is in fact nothing particular about observables and states of composite systems in itself. Quantum information is dealing with an additional structure, namely locality. This means that several *parties* in the system have to be distinguishable and that one puts restrictions on the allowed quantum operations on states, so-called *local operations*.

If two parties have observables $\mathcal{A}_1$ and $\mathcal{A}_2$, then the composite system has observables $\mathcal{A}_{12} = \mathcal{A}_1 \otimes \mathcal{A}_2$. The algebra $\mathcal{A}_{12}$ is generated by elementary tensors $A_1 \otimes A_2$ which satisfy

$$\begin{aligned}
\left(A_1 + \alpha B_1\right) \otimes A_2 &= A_1 \otimes A_2 + \alpha B_1 \otimes A_2, \\
A_1 \otimes \left(A_2 + \alpha B_2\right) &= A_1 \otimes A_2 + \alpha A_1 \otimes B_2, \\
\left(A_1 \otimes A_2\right)^* &= A_1^* \otimes A_2^*.
\end{aligned}$$

It follows that $\|A_1 \otimes A_2\| = \|A_1\|\,\|A_2\|$.

For a purely classical system we have

$$\mathcal{C}_{\mathbb{C}}(\Omega_1) \otimes \mathcal{C}_{\mathbb{C}}(\Omega_2) = \mathcal{C}_{\mathbb{C}}(\Omega_1 \times \Omega_2)\,.$$

So the "phase space" of a composite classical system is just the Cartesian product of the spaces of the factors.

For a purely quantum system

$$\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2} = \mathcal{M}_{d_1 d_2}.$$

A matrix in $\mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ can be seen as a d_1-dimensional matrix, the entries of which are themselves d_2-dimensional matrices, or conversely. Furthermore

$$\left(\oplus_j \mathcal{M}_{d_j}\right) \otimes \left(\oplus_k \mathcal{M}_{n_k}\right) = \oplus_{j,k}\left(\mathcal{M}_{d_j} \otimes \mathcal{M}_{n_k}\right) = \oplus_{j,k}\mathcal{M}_{d_j n_k} .$$

4.1 Marginals

In a composite system a subsystem is recognized by the natural embedding

$$\mathcal{A}_1 \hookrightarrow \mathcal{A}_{12} : A_1 \mapsto A_1 \otimes \mathbb{1}_2 .$$

Using this embedding we obtain the *marginals* of states of a composite system

$$\omega_1(A_1) := \omega_{12}(A_1 \otimes \mathbb{1}_2).$$

Here ω_{12} is a state on the composite system and ω_1 is its marginal on the first factor. An analogous expression defines the marginal on the second factor.

For a classical system this amounts to "integrating out" some variables in a multivariate distribution, e.g. if μ_{12} is a probability vector on $\Omega_1 \times \Omega_2$, then its marginal μ_1 is the probability vector

$$\left(\mu_1\right)_j = \sum_k \left(\mu_{12}\right)_{j,k} .$$

For a matrix algebra $\mathcal{A}_{12} = \mathcal{M}_{d_1} \otimes \mathcal{M}_{d_2}$ let ρ_{12} be the density matrix defining the state ω_{12}. Let $\{e_j\}$ and $\{f_k\}$ be the canonical bases of $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$, then $\{e_j \otimes f_k\}$ is the canonical basis of the tensor space and

$$\begin{aligned}
\omega_1(A_1) &= \omega_{12}(A_1 \otimes \mathbb{1}_2) \\
&= \mathrm{tr}_{\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}}\left(\rho_{12}\,(A_1 \otimes \mathbb{1}_2)\right) \\
&= \sum_{j,k} \langle e_j \otimes f_k\,,\ \rho_{12}\,(A_1 \otimes \mathbb{1}_2)\,(e_j \otimes f_k)\rangle \\
&= \sum_{j,k,\ell,m} \langle e_j \otimes f_k\,,\ \rho_{12}\,(e_\ell \otimes f_m)\rangle\, \langle e_\ell \otimes f_m\,,\ (A_1 \otimes \mathbb{1}_2)\,(e_j \otimes f_k)\rangle \\
&= \sum_{j,k,\ell} \langle e_j \otimes f_k\,,\ \rho_{12}\,(e_\ell \otimes f_k)\rangle\, \langle e_\ell\,,\ A_1\, e_j\rangle \\
&= \mathrm{tr}_{\mathbb{C}^{d_1}}(\rho_1\, A_1) .
\end{aligned}$$

Here, ρ_1 is the *partial trace* of ρ_{12} over the second space, explicitly given by

$$\big(\rho_1\big)_{i,j} = \sum_{k=1}^{d_2} \big(\rho_{12}\big)_{ik,jk}\,. \tag{11}$$

Equation (11) is the quantum analog of integrating out a variable in a bi-variate distribution.

Example 14 Let us work out the partial trace for $d_1 = d_2 = 2$ where the canonical basis vectors of $\mathbb{C}^2 \otimes \mathbb{C}^2$ are labelled in the standard lexicographical order:

$$\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\},$$

$$\rho_{12} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \longrightarrow \rho_1 = \begin{bmatrix} r_{11} + r_{22} & r_{13} + r_{24} \\ r_{31} + r_{42} & r_{33} + r_{44} \end{bmatrix}.$$

A similar operation yields ρ_2. It is crucial to remark that partial trace does not preserve purity; the marginal of a pure state can be very mixed. An extreme example is given by

$$\rho_{12} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \rho_1 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The state ρ_{12} is pure on $\mathcal{M}_4$; it is determined by the Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ while ρ_1 is the completely mixed state on $\mathcal{M}_2$

4.2 Product States and Entanglement

A state ω_{12} on $\mathcal{A}_1 \otimes \mathcal{A}_2$ typically generates many *conditional states* on $\mathcal{A}_1$. Take $0 \leq P \in \mathcal{A}_2$, then

$$A_1 \mapsto \frac{\omega_{12}(A_1 \otimes P)}{\omega_{12}(\mathbb{1} \otimes P)}$$

is a state on $\mathcal{A}_1$ and the dependence of this state on P tells us about the correlations between the parties that are encoded in ω_{12}.

Given two states ω_1 and ω_2 on $\mathcal{A}_1$ and $\mathcal{A}_2$, one can always consider the *product state* $\omega_1 \otimes \omega_2$ on the composite system

$$\big(\omega_1 \otimes \omega_2\big)(A_1 \otimes A_2) := \omega_1(A_1)\,\omega_2(A_2).$$

In this state there are no correlations between the subsystems; conditioning on the second does not affect the first:

$$A_1 \mapsto \frac{\big(\omega_1 \otimes \omega_2\big)(A_1 \otimes P)}{\big(\omega_1 \otimes \omega_2\big)(\mathbb{1} \otimes P)} = \omega_1(A_1)\,.$$

This actually means that both subsystems are *independent*.

A product state is said to be *separable or non-entangled*. More generally, the separable states are the convex hull of the product states

$$\begin{aligned}\mathcal{S}_{\text{sep}}(\mathcal{A}_1 \otimes \mathcal{A}_2) &= \text{Conv}\Big(\Big\{\omega_1 \otimes \omega_2 \,:\, \omega_i \in \mathcal{S}(\mathcal{A}_i)\Big\}\Big) \\ &= \text{Conv}\Big(\Big\{\omega_1 \otimes \omega_2 \,:\, \omega_i \in \mathcal{S}(\mathcal{A}_i) \text{ and } \omega_i \text{ pure}\Big\}\Big)\,.\end{aligned}$$

Studying entanglement for two or more parties has become a sub-field of research within quantum information. One tries to find both *entanglement witnesses* and entanglement quantifiers. A state on a two-party system with one of the parties classical is always separable, so entanglement is a pure quantum notion.

4.3 The Schmidt Decomposition

There is a nice and useful way to write a normalized element η of $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$: one can find, generically in a unique way, a probability vector r and orthonormal families $\{\varphi_i\}$ and $\{\psi_k\}$ in $\mathbb{C}^{d_1}$ and $\mathbb{C}^{d_2}$ such that

$$\eta = \sum_j r_j^{1/2}\, \varphi_j \otimes \psi_j\,.$$

Suppose that the statement is true; then compute the first marginal ρ_1 of the pure state determined by η in the basis $\{\varphi_i\}$:

$$\langle \varphi_i\,,\ \rho_1\, \varphi_j\rangle = \sum_k \langle \varphi_i \otimes \psi_k\,,\ \eta\rangle\, \langle \eta\,,\ \varphi_j \otimes \psi_k\rangle = r_i\, \delta_{ij}\,.$$

This means that ρ_1 is diagonal in the $\{\varphi_i\}$ basis with the r_i as eigenvalues. To prove the Schmidt decomposition, one goes around the other way, namely diagonalizing the marginal of the state determined by η.

Let ρ_1 be the marginal of $|\eta\rangle\langle\eta|$ on the first party and diagonalize ρ_1:

$$\rho_1 = \sum_j r_j\, |\varphi_j\rangle\langle\varphi_j|\,. \tag{12}$$

Next write

$$\eta = \sum_j \varphi_j \otimes \psi_j.$$

Then

$$|\eta\rangle\langle\eta| = \sum_{ij} |\varphi_i\rangle\langle\varphi_j| \otimes |\psi_i\rangle\langle\psi_j|,$$

and taking the partial trace over the second space

$$\rho_1 = \sum_{ij} \langle\psi_j\ ,\ \psi_i\rangle\, |\varphi_i\rangle\langle\varphi_j|.$$

Comparing this with the eigenvalue decomposition (12) we see that

$$\langle\psi_j\ ,\ \psi_i\rangle = \delta_{ij}\, r_j\ .$$

Therefore also the $\{\psi_j\}$ form an orthogonal family.

The number of non-zero elements in r is called the *Schmidt number* of the state η. States with Schmidt number 1 are product states and their marginals are pure and vice versa: hence the Schmidt number characterizes entanglement of pure states, however in a very discontinuous way. There are much nicer entanglement quantifiers as will be explained in other lectures.

To conclude this section let us remark that in a bipartite system both marginals of a pure state have the same eigenvalues, except for multiplicities of 0. Typically the marginals of pure states on a $d_1 \times d_2$ dimensional system have $\min(\{d_1, d_2\})$ non-zero eigenvalues. Moreover, if some marginal of a multi-partite state is pure, then the state factorizes. Hence pure states decouple from "the rest of the world" which is somehow logically necessary in order to treat closed quantum system without regard to the outside world.

5 Dynamics and Measurements

For a closed quantum system two different types of dynamics should be distinguished. A system that is left alone will evolve according to its own unitary dynamics which can in principle last for ever. Another possibility is that a measurement is performed on the system. This will modify the system in an irreversible way and is stochastic in nature. For open systems the situation is more complex. The system we are interested in is coupled to an environment and the composite system can both evolve reversibly according to a unitary evolution and be subject to a measurement. This calls for the introduction of general quantum operations.

5.1 Unitary Dynamics

Symmetries and reversible dynamics of a closed system are given in Heisenberg picture in terms of *automorphisms* of the observables, i.e. the observables move.

An automorphism is a map γ that preserves the full algebraic structure and that is invertible, so

$$\gamma(A\,B) = \gamma(A)\,\gamma(B), \quad \gamma(A + \alpha B) = \gamma(A) + \alpha\gamma(B), \quad \gamma(A^*) = \gamma(A)^* . \tag{13}$$

As γ is invertible, also γ^{-1} is an automorphism. We may as well use the Schrödinger picture where the states move:

$$\gamma^*(\omega)(A) = \omega\big(\gamma(A)\big) .$$

The characteristic property of γ^* is now that it is an invertible affine transformation of the state space. Such maps will preserve the purity of states, i.e. no information is lost.

For a purely classical system automorphisms are permutations g of the configuration space

$$\gamma(f)(\epsilon) = f(g(\epsilon)),$$

while for a purely quantum system automorphisms are given by unitaries

$$\gamma(A) = U\,A\,U^\dagger .$$

For mixed systems blocks of the same dimension can be permuted and within blocks unitaries can act. It is obvious that classical systems with continuous configuration spaces allow much more interesting automorphisms.

The evolution in time of an autonomous system is given by a *group of automorphisms* $\{\gamma_t \,:\, t \in \mathbb{R}\}$. Each γ_t is an automorphism of the observables and

$$\gamma_{t_1} \circ \gamma_{t_2} = \gamma_{t_1+t_2} .$$

Moreover, the maps γ_t depend continuously on time. Because of the group property one can describe the γ_t in terms of a *generator* δ,

$$\delta(A) := -\mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}t}\,\gamma_t(A)\Big|_{t=0} = -\mathrm{i}\,\lim_{t\to 0}\frac{\gamma_t(A) - A}{t} . \tag{14}$$

The group may be reconstructed by *exponentiating* the derivation

$$\gamma_t(A) = \exp(\mathrm{i}t\delta)(A) = A + \mathrm{i}t\delta(A) + \frac{(\mathrm{i}t)^2}{2!}\,\delta(\delta(A)) + \cdots .$$

This means in fact that the evolution can be obtained as the solution of the following differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\gamma_t(A) = \mathrm{i}\delta\big(\gamma_t(A)\big) \qquad \text{and} \qquad \gamma_0(A) = A . \tag{15}$$

Equation (15) is closely related to the *Heisenberg equation of motion*. The generator (14) is called a *derivation* because it has properties that remind one of differentiation:

$$\delta(A\,B) = \delta(A)\,B + A\,\delta(B) \qquad \text{and} \qquad \delta(A^*) = -\delta(A)^*\,. \tag{16}$$

These properties are the differential versions of (13).

There is a nice general form of a derivation on a matrix algebra in terms of a commutator with a Hamiltonian. We begin by defining such a Hamiltonian

$$h := -\mathrm{i}\sum_j \delta(e_{j1})\,e_{1j}\,.$$

Using the properties (16) one proves that

$$h = h^* \qquad \text{and} \qquad \delta(A) = [h, A]\,.$$

Indeed,

$$\delta(\mathbb{1}) = \delta(\mathbb{1}\circ\mathbb{1}) = \delta(\mathbb{1}) + \delta(\mathbb{1}) = 2\delta(\mathbb{1});$$

therefore $\delta(\mathbb{1}) = 0$. Let us now check that $h = h^*$.

$$\begin{aligned} h^* &= \mathrm{i}\sum_j e_{1j}^*\delta(e_{j1})^* \\ &= \mathrm{i}\sum_j e_{j1}\,\delta(e_{1j}) \\ &= \mathrm{i}\delta\Big(\sum_j e_{j1}\,e_{1j}\Big) - \mathrm{i}\sum_j \delta(e_{j1})\,e_{1j} \\ &= \mathrm{i}\delta(\mathbb{1}) - \mathrm{i}\sum_j \delta(e_{j1})\,e_{1j} \\ &= -\mathrm{i}\sum_j \delta(e_{j1})\,e_{1j} \\ &= h\,. \end{aligned}$$

A similar computation shows that $\delta(A) = [h, A]$; it suffices to verify the formula for every matrix unit $e_{k\ell}$. At this point one could wonder, why not use just Hamiltonians? There is certainly an issue with infinite systems where Hamiltonians loose their meaning while derivations remain nicely defined objects. A more down-to-earth reason is that it is useful to work on the level of linear transformations of the observables. Such linear maps are generally known as super-operators. Later super-operators will prove useful in modelling open system dynamics. Finally, the simple description of above can be extended to include time-dependent Hamil-

tonians. They correspond to *controlled closed systems* where parameters in the Hamiltonian can be adjusted from outside the system.

Autonomous dynamics of a finite-dimensional system is rather dull. Let us diagonalize the Hamiltonian h

$$h\,\varphi_k = \epsilon_k\,\varphi_k.$$

Then

$$\delta\big(|\varphi_k\rangle\langle\varphi_\ell|\big) = (\epsilon_k - \epsilon_\ell)\,|\varphi_k\rangle\langle\varphi_\ell|\,.$$

The $\epsilon_k - \epsilon_\ell$ are the *Bohr frequencies* of the system and

$$\begin{aligned}\gamma_t(A) &= \gamma_t\Big(\sum_{k\ell}\langle\varphi_k,\, A\,\varphi_\ell\rangle\,|\varphi_k\rangle\langle\varphi_\ell|\Big)\\ &= \sum_{k\ell} \mathrm{e}^{\mathrm{i}t(\epsilon_k-\epsilon_\ell)}\,\langle\varphi_k,\, A\,\varphi_\ell\rangle\,|\varphi_k\rangle\langle\varphi_\ell|\,.\end{aligned}$$

If the Bohr frequencies are commensurate, i.e. if they are multiples of some frequency, then the evolution of observables will be periodic. In the more general case the motion is only quasi-periodic, which means that it will, waiting sufficiently long, repeat itself within any given precision. This behaviour can obviously not lead to a randomizing behaviour where a system tends to some equilibrium situation.

5.2 Ideal von Neumann Measurements

Observing quantum systems is modelled by an action on the system from the outside. It is still an open problem how to include measurements within the theory. The simplest measurements are exclusive yes–no measurements; they are irreversible and correspond to a random intervention on the system. Let $\{\varphi_i \,:\, i \in I\}$ be an orthonormal basis with corresponding projectors $P_i := |\varphi_i\rangle\langle\varphi_i|$. The $\{P_i \,:\, i \in I\}$ are the von Neumann measurement operators; the index set I labels the possible readings of the measuring apparatus. Preparing a system in a pure state φ and performing the measurement on the system we obtain a probability distribution on the outcome:

$$\mathrm{Prob}\{\text{reading } i\} = |\langle\varphi, \varphi_i\rangle|^2 = \|P_i\varphi\|^2\,.$$

If we filter out readings with outcome in $I_0 \subset I$ we obtain the state

$$\frac{\sum_{i\in I_0}\|P_i\varphi\|^2\,|\varphi_i\rangle\langle\varphi_i|}{\sum_{i\in I_0}\|P_i\varphi\|^2}$$

with probability $\sum_{i\in I_0}\|P_i\varphi\|^2$. Without any filtering we still affect the system; after going through the apparatus we no longer have a pure state but rather a mixture:

$$|\varphi\rangle\langle\varphi| \longrightarrow \sum_{i\in I} |\langle\varphi_i,\varphi\rangle|^2\,|\varphi_i\rangle\langle\varphi_i|\,. \tag{17}$$

The transformation (17) is an example of a *quantum operation*.

The broader and more natural class of *positive operator valued measures*, POVMs, arises by coupling a system to an auxiliary one and performing a von Neumann measurement on the second system. Let $\{|i\rangle\}$ be an orthonormal basis of the second system and $[U_{ij}]$ a unitary on the composite system, the U_{ij} being operators of the first system. This unitary corresponds to a global evolution of the composite system during some time interval. Consider now an observable A of the first system, embed it in the composite system, let it evolve for some time and filter out the reading k for the second system. Doing so, we actually perform the operation

$$\begin{aligned} A &\mapsto U^\dagger A\otimes \mathbb{1} U\\ &= \sum_{i,k,\ell}\Big(\big(U_{ki}\big)^\dagger A\,U_{\ell i}\Big)\otimes|k\rangle\langle\ell|\\ &\mapsto \sum_{i}\Big(\big(U_{ki}\big)^\dagger A\,U_{ki}\Big)\otimes|k\rangle\langle k|\,. \end{aligned}$$

Note that, because U is unitary,

$$\sum_i \big(U_{ki}\big)^\dagger U_{ki} = \mathbb{1}\,.$$

In general,

$$\{X_\alpha\,:\,\alpha\in A\}\qquad\text{with}\qquad X_\alpha\geq 0 \text{ and } \sum_\alpha X_\alpha = \mathbb{1}$$

is called a *POVM* and it describes a generalized measurement; $\langle\varphi, X_\alpha\varphi\rangle$ is the probability of the outcome α when the system is in the state φ. An instrument describes how the state is modified by the measurement. We need more than the X_α, namely a collection of V_α such that $X_\alpha = \big(V_\alpha\big)^* V_\alpha$. A set $\{V_\alpha\,:\,\alpha\in A\}$ such that

$$\sum_\alpha \big(V_\alpha\big)^* V_\alpha = \mathbb{1}$$

is called an *operational partition of unity* and $\{\big(V_\alpha\big)^* V_\alpha\,:\,\alpha\in A\}$ its associated POVM. An operational partition describes a quantum operation on the system, in Heisenberg picture

$$A \mapsto \sum_{\alpha} \big(V_\alpha\big)^* A\, V_\alpha\,.$$

Example 15 (Joint measurement of spin $\frac{1}{2}$ components) Let us try to measure jointly the x and z components of a spin $\frac{1}{2}$ system. There are four possible outcomes $(\pm, \pm)$. A von Neumann measurement of σ^x would use the projectors $\{\frac{1}{2}(\mathbb{1} \pm \sigma^x)\}$ and similar expressions for σ_z. Multiplying these measurement operators $\{\frac{1}{4}(\mathbb{1} \pm \sigma^x)(\mathbb{1} \pm \sigma^z)\}$ does not yield a POVM and neither does $\Big\{\big\{\frac{1}{2}(\mathbb{1} \pm \sigma^x),\, \frac{1}{2}(\mathbb{1} \pm \sigma^z)\big\}\Big\}$. The most symmetrical expression that yields a POVM is

$$\Big\{\tfrac{1}{4}\Big(\mathbb{1} \pm \frac{\sigma^x}{\sqrt{2}} \pm \frac{\sigma^z}{\sqrt{2}}\Big)\Big\}\,.$$

How efficient is this for measuring σ^z? Consider the eigenstate spin up of σ^z, i.e. the pure state generated by $|+\rangle$. It yields the probability measure

$$\Big(\Big\langle +\,,\ \tfrac{1}{4}\Big(\mathbb{1} \pm \frac{\sigma^x}{\sqrt{2}} \pm \frac{\sigma^z}{\sqrt{2}}\Big) + \Big\rangle\Big)$$

on $(\pm, \pm)$. We now compute the marginal of this measure on the σ^z observable; we get the probabilities

$$\Big(\tfrac{1}{2}\Big(1 \pm \tfrac{1}{\sqrt{2}}\Big)\Big) \approx (0.85, 0.15)\,.$$

So by this procedure of computing the probabilities our POVM introduces an uncertainty of about 15% on the outcome for σ^z.

5.3 Open Systems and Completely Positive Maps

The general idea is to concentrate on a small system which is in interaction with a large environment but to discard all detailed information on the environment. This can only work well if the system and the environment can easily be distinguished, e.g. if they do not interact too strongly and if there are very different time scales involved in the proper evolution of both separately. So one starts at $t = 0$ with independent system and environment in states ω_0 and τ and one is interested in

$$\omega_{t,\omega_0}(A) := \big(\omega_0 \otimes \tau\big)(\gamma_t(A \otimes \mathbb{1}))\,.$$

The state τ is usually a relevant state of the environment, e.g. a thermal state or the vacuum and γ_t describes the global evolution of system and environment. The map

$$\omega_0 \mapsto \omega_{t,\omega_0}$$

is affine and has therefore also a Heisenberg picture version

$$A \mapsto \Gamma_t(A) \qquad \text{with} \qquad \omega_{t,\omega_0}(A) = \omega_0\big(\Gamma_t(A)\big)\,.$$

Γ_t is called the *reduced dynamics* of the small system; it has in general a quite complicated time dependence. If Γ_t is a semi-group then one says that the reduced evolution is *Markovian* in time meaning that there are no memory effects:

$$\Gamma_{t_1} \circ \Gamma_{t_2} = \Gamma_{t_1+t_2}, \quad t_1,\ t_2 \in \mathbb{R}^+\,.$$

We can now wonder when such a Markovian behaviour can be expected. Under which assumptions can one prove that a reduced dynamics is Markovian? This question is answered by the theories of weak coupling and singular coupling limits. Further questions concern the classification of Markovian semi-groups and their generators. Fortunately, quantum operations which arise as above have an additional property which is deeply connected with quantum mechanics, namely *complete positivity*. This makes the analysis of semi-groups and their generators feasible, at least under some technical conditions. Complete positivity will be the last topic of these lectures.

Positive maps

A linear transformation $\Gamma : \mathcal{A} \to \mathcal{A}$ that maps positive elements in positive elements is called *positive*, P. Sums, positive multiples and compositions of P maps are still P. A linear map $\Gamma : \mathcal{A} \to \mathcal{A}$ is *unity preserving*, UP, if $\Gamma(\mathbb{1}) = \mathbb{1}$. If $\omega \in \mathcal{S}(\mathcal{A})$ and Γ is PUP, then

$$\Gamma^*(\omega)(A) := \omega\big(\Gamma(A)\big)$$

is again a state. Affine maps like Γ^* are called *positive and state preserving*, PTP. The T comes from "trace preserving". Γ is in *Heisenberg picture* while Γ^* is in *Schrödinger picture* and the pictures are dual

$$\Gamma^*(\omega) = \omega \circ \Gamma\,.$$

At first glance, PTP maps exactly satisfy the requirements for general quantum maps compatible with the probabilistic structure of quantum mechanics. They could be used to describe general channels, i.e. black boxes that accept a state on one end and emit a, generally distorted, state at the other:

$$\omega \longrightarrow \blacksquare \longrightarrow \Gamma^*(\omega)\,.$$

There is, however, a fundamental problem with this: P is not robust for composing systems.

Example 16 (The transposition) The standard example of a map that is PUP but not well-behaved with respect to composition is transposition

$$\mathsf{T} : \mathcal{M}_2 \to \mathcal{M}_2 : A \mapsto A^{\mathsf{T}} .$$

Clearly transposition is PUP but it turns out that the trivial extension

$$\mathrm{id} \otimes \mathsf{T} : \mathcal{M}_2 \otimes \mathcal{M}_2 \to \mathcal{M}_2 \otimes \mathcal{M}_2 : A \otimes B \mapsto A \otimes B^{\mathsf{T}}$$

is no longer positive!

Order the standard product basis in $\mathbb{C}^2 \otimes \mathbb{C}^2$ in the usual way

$$\big\{|00\rangle,\ |01\rangle,\ |10\rangle,\ |11\rangle\big\}$$

and let Q be the projector on $\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$

$$Q = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} .$$

The map $\mathrm{id} \otimes \mathsf{T}$ acts on Q by replacing $\langle 00,\, Q\, 11\rangle$ by $\langle 01,\, Q\, 10\rangle, \ldots$

$$\mathrm{id} \otimes \mathsf{T}(Q) = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cong \frac{1}{2}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\},$$

which is clearly not positive.

Completely positive maps

The example of above justifies the following definition: a linear map $\Gamma : \mathcal{A} \to \mathcal{A}$ is *completely positive*, CP, if $\mathrm{id} \otimes \Gamma : \mathcal{M}_n \otimes \mathcal{A} \to \mathcal{M}_n \otimes \mathcal{A}$ is P for $n = 2, 3, \ldots$. Sums, positive multiples and compositions of CP maps are still CP.

Example 17 (*A CP map*) Let $\mathcal{A} = \mathcal{M}_d$, choose an arbitrary $V \in \mathcal{M}_d$ and put

$$\Gamma(A) := V^\dagger A\, V .$$

We claim that Γ is CP. Indeed, take $X \in \mathcal{M}_n \otimes \mathcal{M}_d$ positive, then $X = Y^\dagger Y$ for a certain $Y \in \mathcal{M}_n \otimes \mathcal{M}_d$ and

$$\begin{aligned}\mathrm{id}\otimes\Gamma(X) &= \left(\mathbb{1}_n\otimes V^\dagger\right) X \left(\mathbb{1}_n\otimes V\right)\\ &= \left(\mathbb{1}_n\otimes V^\dagger\right) Y^\dagger Y \left(\mathbb{1}_n\otimes V\right)\\ &= \left(Y\,\mathbb{1}_n\otimes V\right)^\dagger \left(Y\,\mathbb{1}_n\otimes V\right)\\ &\geq 0\,.\end{aligned}$$

A linear map $\Lambda : \mathcal{M}_d \to \mathcal{M}_d$ is often called a *super-operator*. A useful description of such maps is provided by the *Jamiołkowski–Choi encoding*. Let $\{e_{ij}\}$ be the standard matrix units in $\mathcal{M}_d$, then every element of $\mathcal{M}_d$ can be written as

$$A = \sum_{ij} A_{ij}\, e_{ij}\,.$$

This implies that Λ is fully determined by providing the image of every matrix unit under Λ. A global way of doing this consists in forming the large matrix

$$\begin{aligned}\mathsf{C}(\Lambda) &:= \sum_{ij} e_{ij}\otimes\Lambda(e_{ij})\\ &= \begin{bmatrix}\Lambda(e_{11}) & \Lambda(e_{12}) & \cdots & \Lambda(e_{1d})\\ \vdots & & & \\ \Lambda(e_{d1}) & \cdots & & \Lambda(e_{dd})\end{bmatrix} \in \mathcal{M}_d\otimes\mathcal{M}_d\,.\end{aligned}$$

Remark that the map $\Lambda \mapsto \mathsf{C}(\Lambda)$ is linear and one to one so that it truly is an encoding. $\mathsf{C}(\Lambda)$ is called the *Choi matrix* of Λ. Similarly,

$$\mathsf{J}(\Lambda^*) := \frac{1}{d}\sum_{ij} e_{ij}\otimes\Lambda^*(e_{ij}) \in \mathcal{M}_d\otimes\mathcal{M}_d$$

is called the *Jamiołkowski state* of Λ^*. The name state will be justified soon.

Example 18 (*Choi matrix of the identity map*). Take $\Lambda = \mathrm{id}$, then

$$\begin{aligned}\mathsf{C}(\mathrm{id}) &= \sum_{ij} e_{ij}\otimes e_{ij}\\ &= \sum_{ij} |e_i\rangle\langle e_j|\otimes|e_i\rangle\langle e_j|\\ &= \Big|\sum_i e_i\otimes e_i\Big\rangle\Big\langle\sum_j e_j\otimes e_j\Big|\,.\end{aligned}$$

This is, up to a factor d, the projector on a maximally entangled state. In particular $\mathsf{C}(\Lambda)$ is positive.

Choi's theorem characterizes the CP maps in terms of their encoding: *$\Lambda : \mathcal{M}_d \to \mathcal{M}_d$ is CP iff $C(\Lambda)$ is positive. If this is the case, there exist $\{V_i\}$ such that*

$$\Lambda(X) = \sum_i V_i^\dagger \, X \, V_i \,. \tag{18}$$

Equation (18) is called a *Kraus representation* of Λ.

Suppose first that Λ is CP. Applying $\mathrm{id} \otimes \Lambda$ to

$$\sum_{ij} e_{ij} \otimes e_{ij} = \Big|\sum_i e_i \otimes e_i\Big\rangle\Big\langle\sum_j e_j \otimes e_j\Big| \geq 0$$

we obtain

$$\mathsf{C}(\Lambda) = \sum_{ij} e_{ij} \otimes \Lambda(e_{ij}) \geq 0 \,.$$

Suppose conversely that $\mathsf{C}(\Lambda) \geq 0$, then write its eigenvalue decomposition

$$\mathsf{C}(\Lambda) = \sum_\alpha \lambda_\alpha \, |f_\alpha\rangle\langle f_\alpha| = \sum_\alpha |\varphi_\alpha\rangle\langle\varphi_\alpha| \,,$$

where $\varphi_\alpha := \sqrt{\lambda_\alpha}\, f_\alpha$ which is possible because $\lambda_\alpha \geq 0$. Consider now a single term $|\varphi\rangle\langle\varphi|$, $\varphi \in \mathbb{C}^d \otimes \mathbb{C}^d$, and find out of which super-operator it is the encoding

$$\varphi = \sum_i e_i \otimes \varphi_i, \quad \varphi_i \in \mathbb{C}^d,$$
$$|\varphi\rangle\langle\varphi| = \sum_{ij} e_{ij} \otimes |\varphi_i\rangle\langle\varphi_j| \,.$$

But this is the encoding of

$$\Phi : e_{ij} \mapsto |\varphi_i\rangle\langle\varphi_j| = V^\dagger \, e_{ij} \, V$$

with $V^\dagger \, e_i := \varphi_i$. Hence, by Example 4, Λ is a sum of CP maps and is therefore CP.

Example 19 (*Rescaling spin* $\frac{1}{2}$). We begin with the Pauli matrices

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma^y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \qquad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$
$$\sigma^x \sigma^y = \mathrm{i}\,\sigma^z, \qquad \big(\sigma^x\big)^\dagger = \sigma^x, \qquad \big(\sigma^x\big)^2 = \mathbb{1},$$

and consider a rescaling map

$$\Lambda(\mathbb{1}) = \mathbb{1} \qquad \text{and} \qquad \Lambda(\overline{\sigma}) = \lambda\,\overline{\sigma}, \quad \lambda \in \mathbb{R} \,.$$

Λ is positive iff $-1 \leq \lambda \leq 1$. Putting $\lambda = -1$ is a quite bad quantum operation as it destroys the commutation relations. When is Λ CP? To answer that question we

must compute $\mathsf{C}(\Lambda)$ using

$$\Lambda(e_{11}) = \Lambda\Big(\tfrac{1}{2}\,\mathbb{1} + \tfrac{1}{2}\,\sigma^z\Big) = \tfrac{1}{2}\,\mathbb{1} + \tfrac{\lambda}{2}\,\sigma^z,$$
$$\Lambda(e_{12}) = \Lambda\Big(\tfrac{1}{2}\,\sigma^x + \tfrac{\mathrm{i}}{2}\,\sigma^y\Big) = \frac{\lambda}{2}\,\sigma^x + \frac{\mathrm{i}\lambda}{2}\,\sigma^y = \lambda\, e_{12}\,.$$

We now have

$$\mathsf{C}(\Lambda) = \begin{bmatrix} \frac{1+\lambda}{2} & 0 & 0 & \lambda \\ 0 & \frac{1-\lambda}{2} & 0 & 0 \\ 0 & 0 & \frac{1-\lambda}{2} & 0 \\ \lambda & 0 & 0 & \frac{1+\lambda}{2} \end{bmatrix} \cong \begin{bmatrix} \frac{1-\lambda}{2} & 0 \\ 0 & \frac{1-\lambda}{2} \end{bmatrix} \oplus \begin{bmatrix} \frac{1+\lambda}{2} & \lambda \\ \lambda & \frac{1+\lambda}{2} \end{bmatrix}.$$

It follows that $\mathsf{C}(\Lambda) \geq 0$ iff $-\frac{1}{3} \leq \lambda \leq 1$, so we can flip all components of $\overline{\sigma}$ but have to rescale them at least with a factor three in order to guarantee complete positivity.

We end this section with some general properties of CP maps.

1. The notion of complete positivity can easily be extended to maps between different algebras.
2. If $\Lambda : \mathcal{A} \to \mathcal{B}$ is P and $\mathcal{A}$ is $\mathcal{B}$ are Abelian, then Λ is CP. This means that complete positivity, just as entanglement, is a purely quantum mechanical notion with no classical counterpart.
3. If Λ is P, then

$$\Lambda(X^*) = \Lambda(X)^* \qquad \text{and} \qquad \|\Lambda(X)\| \leq \|X\|\,\|\Lambda(\mathbb{1})\|\,.$$

4. If Λ is PUP, then

$$\Lambda(X^*)\,\Lambda(X) \leq \Lambda(X^*X) \quad \text{for normal } X:\ X\,X^* = X^*X\,.$$

   This is sometimes called the Schwarz inequality for positive maps.
5. If Λ is CPUP, then

$$\Lambda(X^*)\,\Lambda(X) \leq \Lambda(X^*X) \quad \text{for all } X\,.$$

   This very useful inequality is called the *two-positivity inequality*. In fact it holds already if the extension of Λ over the two-dimensional matrices is positive, which is a strictly weaker condition than complete positivity.

Semi-groups of UPCP maps on $\mathcal{M}_d$

The last topic of these introductory lectures will be the structure of semi-groups of completely positive unity preserving maps on $\mathcal{M}_d$. As mentioned before, such maps

give the simplest description of a dissipative evolution. The weak coupling theory shows that this sort of approximation arises when one has a weak interaction of the system with the environment and when there is a clear separation between the fast time scale on which the environment relaxes and the natural time scale of the small system. We do not enter into the details of the derivation but rather start with the outcome: a semi-group of UPCP maps.

We deal more precisely with the following structure:

$$\begin{aligned}&\Gamma_t : \mathcal{M}_d \to \mathcal{M}_d \text{ is UPCP for every } t \geq 0,\\ &\Gamma_t \circ \Gamma_s = \Gamma_{t+s}, \quad s,\ t \geq 0,\\ &\Gamma_0 = \mathrm{id} \text{ and } t \mapsto \Gamma_t \text{ is continuous.}\end{aligned}$$

Writing the semi-group in terms of a generator

$$\Gamma_t = \exp(t\mathcal{L}) \qquad \text{with} \qquad \mathcal{L}(X) = \lim_{t \downarrow 0} \frac{\Gamma_t(X) - X}{t},$$

we ask about the general form of $\mathcal{L}$. The result is given by *Lindblad's theorem*: there are $H = H^\dagger \in \mathcal{M}_d$ and $V_i \in \mathcal{M}_d$ such that

$$\mathcal{L}(X) = \mathrm{i}[H, X] + \sum_i \Big(V_i^\dagger X\, V_i - \tfrac{1}{2}\, V_i^\dagger V_i\, X - \tfrac{1}{2}\, X\, V_i^\dagger V_i \Big).$$

The idea of the proof is to expand the maps in Jamiołkowski–Choi encoding around $t = 0$ and to express that for t slightly positive the Choi matrices of the corresponding maps have to be positive:

$$\mathsf{C}\big(\exp(t\mathcal{L})\big) = \mathsf{C}(\mathrm{id}) + t\,\mathsf{C}(\mathcal{L}) + \mathrm{o}(t).$$

We computed $\mathsf{C}(\mathrm{id})$ in Example 18; it is d times the projector on a maximally entangled vector. Write now $\mathbb{C}^d \otimes \mathbb{C}^d = \mathbb{C} \oplus \mathbb{C}^{d^2-1}$ where the first dimension is spanned by the maximally entangled vector. Then

$$\mathsf{C}(\mathcal{L}) = \begin{bmatrix} a & \alpha^{\mathsf{T}} \\ \alpha & A \end{bmatrix}, \tag{19}$$

with $a \in \mathbb{R}$, $\alpha \in \mathbb{C}^{d^2-1}$ and $A \in \mathcal{M}_{d^2-1}$. Requiring complete positivity amounts to requiring that $\mathsf{C}(\mathrm{id}) + t\mathsf{C}(\mathcal{L}) + \mathrm{o}(t)$ be positive for $t \geq 0$ and small. We can use the following result from linear algebra: on $\mathbb{C}^m \oplus \mathbb{C}^n$ consider the block matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

with, e.g., $A_{12} : \mathbb{C}^n \to \mathbb{C}^m$, then $A \geq 0$ is equivalent with

$$\begin{aligned} &A_{21} = \left(A_{12}\right)^{\dagger}, \\ &A_{22} \geq 0, \\ &A_{12}\,\mathrm{Ker}(A_{22}) = 0, \\ &A_{11} \geq A_{12}\left(A_{22}\right)^{-1} A_{21}\,. \end{aligned}$$

In this way we obtain for the generator (19) the conditions

$$A \geq 0 \qquad \text{and} \qquad \alpha \in \mathrm{Ran}(A)\,.$$

To finish the proof, we write the eigenvalue decomposition of A and repeat the computation we made for the Jamiołkowski–Choi–Kraus theorem.

Acknowledgments This work was partially funded by the Bilateral Cooperation Project BIL/04/48 between Flanders and South Africa.

References

1. R. Alicki, M. Fannes: *Quantum Dynamical Systems* (Clarendon Press, Oxford, 2001) 2
2. I. Bengtsson, K. Życzkowski: *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006) 2
3. H.P. Breuer, F. Petruccione: *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002) 2
4. O. Bratteli, D.W. Robinson: *Operator Algebras and Quantum Statistical Mechanics: C*- and W*-Algebras. Symmetry Groups. Decomposition of States*, volume 1 (Springer, Berlin, 1979) and *Operator Algebras and Quantum Statistical Mechanics: Equilibrium States. Models in Quantum Statistical Mechanics*, volume 2, 2nd edn (Springer, Berlin, 1997) 2
5. M.A. Nielsen, I.L. Chuang: *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000) 2

Covariant Mappings for the Description of Measurement, Dissipation and Decoherence in Quantum Mechanics

Bassano Vacchini

The general formalism of quantum mechanics for the description of statistical experiments is briefly reviewed, introducing in particular position and momentum observables as POVM characterized by their covariance properties with respect to the isochronous Galilei group. Mappings describing state transformations both as a consequence of measurement and of dynamical evolution for a closed or open system are considered with respect to the general constraints they have to obey and their covariance properties with respect to symmetry groups. In particular different master equations are analyzed in view of the related symmetry group, recalling the general structure of mappings covariant under the same group. This is done for the damped harmonic oscillator, the two-level system, and quantum Brownian motion. Special attention is devoted to the general structure of translation-covariant master equations. Within this framework a recently obtained quantum counterpart of the classical linear Boltzmann equation is considered, as well as a general theoretical framework for the description of different decoherence experiments, pointing to a connection between different possible behaviors in the description of decoherence and the characteristic functions of classical Lévy processes.

1 Introduction

Since its very beginning quantum mechanics has urged physicists and other scientists, getting interested in or involved with it, to radically change their classical picture of reality, as well as their way to describe and understand experiments. Despite the elapsed time, the matured knowledge about quantum mechanics and the growing number of applications, the process of deeper understanding of quantum mechanics and of its truly basic features is still on its way. In these notes we will stress the standpoint that quantum mechanics actually is a probability theory, enlarging and modifying the horizons of the classical one and allowing to describe quantitatively

B. Vacchini (✉)
Dipartimento di Fisica dell'Università di Milano and INFN, Milan, Italy
e-mail: bassano.vacchini@mi.infn.it

Vacchini, B.: *Covariant Mappings for the Description of Measurement, Dissipation and Decoherence in Quantum Mechanics*. Lect. Notes Phys. **787**, 39–77 (2010)
DOI 10.1007/978-3-642-02871-7_2 © Springer-Verlag Berlin Heidelberg 2010

experiments at the microscopic level. The probabilistic standpoint aiming at an understanding of the statistical structure of quantum theory has in fact turned out to be of great significance in recent achievements in the description of quantum mechanical systems, leading to the introduction of new relevant concepts and tools often in analogy with classical probability theory. In particular we will build on the growing impact that the modern formulation of quantum mechanics in terms of statistical operators, POVM and instruments is having in recent experiments and theoretical studies, as well as the relevance of the characterization of mappings giving the dynamical evolution of both closed and open systems. This applies as well to one-step transformations describing the overall effect of a measurement, as to measurements resolved in their time duration, or more generally irreversible evolutions taking place as a consequence of the interaction of the system of interest with some external system, typically even though not necessarily having many degrees of freedom. The presentation will pay particular attention to the structural features of such mappings and especially to their covariance with respect to the representation of a symmetry group, whenever this applies.

This contribution grew out of the lectures given at the Summer School in Theoretical Physics in Durban and is organized as follows. Section 2 supports the point of view according to which quantum mechanics actually is a probability theory and its general formulation is presented in this spirit, starting from the description of statistical experiments. States are introduced as preparation procedures mathematically represented by statistical operators, observables as registration procedures are described in terms of POVM, and the statistics of experimental outcomes is given in terms of the trace formula. Finally, general state transformations as a consequence of measurement are associated with instruments, which provide the transformed state as well as the statistics of the measurement. Examples are provided focussing on position observables, as well as joint position and momentum observables, understood as specified according to their covariance properties with respect to the isochronous Galilei group. Section 3 explores the relevance of the concept of mapping acting on a space of operators in a given Hilbert space for situations ranging from free evolutions to open system dynamics. Dynamical mappings are characterized in view of possible constraints depending on the physical situation of interest and leading to important informations on their possible structure. In particular we consider the notion of covariance with respect to the representation of a symmetry group, complete positivity and semigroup composition law, corresponding to a Markov approximation. This leads to the Lindblad characterization of generators of quantum-dynamical semigroups, possibly also including covariance requirements. Particular examples of master equation are given stressing their covariance properties with respect to the proper symmetry, also providing the general characterization of master equations covariant under the same group. This is done for the damped harmonic oscillator and shift-covariance, a two-level system and rotation-covariance, quantum Brownian motion and translation-covariance. Finally it is shown how the general expression of a translation-covariant generator, building on a quantum non-commutative version of the Lévy-Khintchine formula, actually encompasses a quantum version of the classical linear Boltzmann

equation for the description of the motion of a quantum test particle in a gas, as well as a unified theoretical framework for the explanation of different decoherence experiments.

2 Quantum Mechanics as Quantum Probability

In the present section we will briefly introduce the basic tools necessary in order to describe in the most general way a quantum mechanical system and the possible measurements that can be performed on it. The basic idea that we would like to convey or at least draw to the reader's attention is that quantum mechanics is indeed naturally to be seen as a probability theory, significantly different from the classical one, rather than an extension of classical mechanics. Experiments at the microscopic level are of statistical nature in an essential way and their quantitative description asks for a probabilistic model which is the quantum one, emerged in the twenties and first thoroughly analyzed by von Neumann [1], actually before the foundations of classical probability theory were laid down by Kolmogorov in the thirties [2]. The fact that quantum mechanics is a probability theory different from the classical one, containing the latter as a special case, brings with itself that quantum experiments and their statistical description exhibit new features, which sometimes appear unnatural or paradoxical when somehow forced to fit in a classical probabilistic picture of reality, which is closer to our intuition. Our presentation is more akin to the introduction to quantum mechanics one finds in textbooks on quantum information and communication theory rather than standard quantum mechanics textbooks, even though at variance with the former we will mainly draw examples from systems described in an infinite-dimensional Hilbert space. The standpoint according to which quantum mechanics actually is a probability theory is by now well understood, and even though it is still not in the spirit of typical textbook presentations, it has been developed and thoroughly investigated in various books and monographs (see, e.g., [3–10]), to which we refer the reader for more rigorous and detailed presentations. A more concise account of similar ideas has also been given in [11].

2.1 Classical Statistical Description

The basic setting of classical probability theory as clarified by Kolmogorov is described within the mathematical framework of measure theory. A classical probability model is fixed by specifying a measure space, which is the space of elementary events, a σ-algebra on this measure space characterizing the meaningful events to which we want to ascribe probabilities, and a probability measure on it. The observable quantities are then given by real measurable functions on this space, i.e., random variables. Take for example the case of the classical description of the dynamics of a point particle in three-dimensional space. Then the measure space is

given by the usual phase-space $\mathbb{R}^3 \times \mathbb{R}^3$ endowed with the Borel σ-algebra, where the points of phase-space can be identified with position and momentum of the particle. The probability measure can be expressed by means of a probability density $f(\mathbf{x}, \mathbf{p})$, i.e., a positive and normalized element of $L^1(\mathbb{R}^3 \times \mathbb{R}^3)$ and observables are described as random variables given by real functions $X(\mathbf{x}, \mathbf{p})$ in $L^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$, so that exploiting the canonical duality relation between L^1 and L^∞ mean values are given by

$$\langle X \rangle_f = \int_{\mathbb{R}^3 \times \mathbb{R}^3} d^3\mathbf{x} d^3\mathbf{p}\, X(\mathbf{x}, \mathbf{p}) f(\mathbf{x}, \mathbf{p}).$$

The very same probability density $f(\mathbf{x}, \mathbf{p})$ allows to calculate the expectation value of any random variable, i.e., of any observable. In particular any observable taking values in $\mathbb{R}$ defines a probability measure on this space according to the formula.

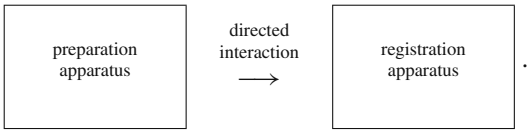
$$\mu^X(M) = \int_{X^{-1}(M)} d^3\mathbf{x} d^3\mathbf{p}\, f(\mathbf{x}, \mathbf{p}),$$

where M is a Borel set in the outcome space $\mathbb{R}$ of the observable and again the same probability density $f(\mathbf{x}, \mathbf{p})$ appears.

As a special case for the position observable $X(\mathbf{x}, \mathbf{p}) = \mathbf{x}$ the measure $\mu^{\mathbf{x}}$ can be expressed by means of the probability density $f^{\mathbf{x}}(\mathbf{x})$ obtained by taking the marginal of $f(\mathbf{x}, \mathbf{p})$ with respect to momentum and similarly for the momentum observable $X(\mathbf{x}, \mathbf{p}) = \mathbf{p}$. In particular one can notice that all observables commute, the points of phase-space can be taken as meaningful elementary events and any probability measure can be uniquely decomposed as a convex mixture of the extreme points of the convex set of probability measures, given by the measures with support concentrated at these elementary events given by single points in phase-space. This probabilistic description is however not mandatory in the classical case, where a deterministic description applies, and it only becomes a very convenient or the only feasible tool for systems with a very high number of degrees of freedom. The situation is quite different in the quantum case.

2.2 Statistics of an Experiment

In the quantum case experiments are by necessity of statistical nature. The most simple setup can typically be described as a suitably devised macroscopic apparatus, possibly made up of lots of smaller components, preparing the microscopical system we would like to study, which in turn triggers another macroscopic device designed to measure the value of a definite quantity. The reproducible quantity to be compared with the theory is the relative frequency according to which the preparation apparatus triggers the registration apparatus in a high enough number of repetitions of the experiment under identical circumstances. A most simple sketch of such a setup can be given by the so-called Ludwig's Kisten [3, 12]



More complicated setups can be traced back to this one by suitably putting together different apparatuses in order to build a new preparation apparatus, and similarly for the registration part. In order to describe such experiments one has to introduce a suitable probability theory, which can actually account for the various experimental evidences to be gained at microscopic level. This is accomplished by introducing mathematical objects describing the preparation and the registration, as well as a statistical formula to extract from these two objects the probability densities to be compared with the experimental outcomes.

States as preparation procedures

In the quantum case a preparation procedure is generally described by a statistical operator. Given the Hilbert space $\mathcal{H}$ in which the system one considers has to be described, e.g., $L^2(\mathbb{R}^3)$, for the center of mass degrees of freedom of a particle in three-dimensional space, statistical operators are positive trace class operators on $\mathcal{H}$ with trace equal to one:

$$\rho \in \mathcal{K}(\mathcal{H}) = \{\rho \in \mathcal{T}(\mathcal{H}) |\, \rho = \rho^\dagger, \rho \geq 0, \operatorname{Tr} \rho = 1\}.$$

The set $\mathcal{K}(\mathcal{H})$ of statistical operators is a convex subset of the space $\mathcal{T}(\mathcal{H})$ of trace class operators on the Hilbert space $\mathcal{H}$, so that any convex mixture of statistical operators is again a statistical operator:

$$\rho_1, \rho_2 \in \mathcal{K}(\mathcal{H}) \Rightarrow w = \mu\rho_1 + (1-\mu)\rho_2 \in \mathcal{K}(\mathcal{H}) \text{ for } 0 \leq \mu \leq 1.$$

In particular the extreme points of such a set are given by one-dimensional projections, that is to say pure states, which cannot be expressed as a proper mixture. We stress the fact that statistical operators are actually to be associated to the considered statistical preparation procedure, rather than to the system itself. More precisely they describe a whole equivalence class of preparation procedures which all prepare the system in the same state, even though by means of quite different macroscopic apparatuses. This correspondence between statistical operators and equivalence classes of preparation procedures is reflected in the fact that a statistical operator generally admits infinitely many different decompositions, as mixtures of pure states or other statistical operators. Relying on the spectral theorem a statistical operator can always be written in the form:

$$\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|, \quad \lambda_j \geq 0, \quad \sum_j \lambda_j = 1,$$

with $\{|\psi_j\rangle\}$ as an orthonormal set. However infinite many others not necessarily orthogonal demixtures generally exist. Think for example of the most simple case of a statistical operator describing the spin state of a fully unpolarized beam of spin $1/2$ particles: $\rho = \frac{1}{2}\mathbb{1}$. Then any orthogonal basis in $\mathbb{C}^2$ (e.g., the eigenvectors of the spin operator along a given arbitrary direction) provides an orthogonal decomposition of the considered statistical operator. Such decompositions correspond to different possible macroscopic procedures leading to such a preparation. In the N runs of the statistical experiment the beam is prepared $\frac{N}{2}$ times with spin $+\frac{\hbar}{2}$ along a fixed direction, and $\frac{N}{2}$ times with spin $-\frac{\hbar}{2}$. All such preparations, differing in the choice of direction, lead to the same state, but they cannot be performed together since no apparatus can measure the spin along two different directions: they are therefore incompatible. The prepared state is however the same and the actual preparation cannot be distinguished on the basis of any other subsequent statistical experiment whatsoever performed on the obtained state. At variance with the classical case, therefore, states are given by operators which generally admit infinitely many convex decompositions and represent equivalence class of preparation procedures.

Observables as registration procedures

On the same footing one has to associate a mathematical object to a macroscopic apparatus assembled in order to measure the value of a certain quantity. Once again utterly different and generally incompatible macroscopic procedures and apparatuses can possibly be used to assign a value to the same physical quantity. The operator describing an observable is therefore to be understood as the mathematical representative of a whole equivalence class of registration procedures. In full generality an observable in the sense clarified above is given by a positive operator-valued measure (POVM), the measure theoretic aspect appearing since one is in fact interested in the probability that the quantity of interest lies within a certain interval.

A POVM is a mapping defined on a suitable measure space and taking values in the set of positive operators within $\mathcal{B}(\mathcal{H})$, that is to say the Banach space of bounded operators on $\mathcal{H}$. Taking for the sake of concreteness an observable assuming values in $\mathbb{R}^3$, such as the position of a particle in three-dimensional space, a POVM is given by a mapping F defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^3)$:

$$\begin{aligned} F : \mathcal{B}(\mathbb{R}^3) &\to \mathcal{B}(\mathcal{H}), \\ M &\to F(M) \end{aligned}$$

associating to each interval $M \in \mathcal{B}(\mathbb{R}^3)$ a positive bounded operator in such a way that

$$\begin{aligned} &0 \leq F(M) \leq \mathbb{1}, \\ &F(\emptyset) = 0, \quad F(\mathbb{R}^3) = \mathbb{1}, \\ &F(\cup_i M_i) = \sum_i F(M_i) \, if \; M_i \cap M_j = \emptyset \, \text{when} \, i \neq j, \end{aligned}$$

where the first condition will turn out to be necessary for the statistical interpretation, the second one expresses normalization, always associating the null operator to the empty set and the identity to the whole space, while the last condition amounts to σ-additivity. For a fixed set $M \in \mathcal{B}(\mathbb{R}^3)$ the operator $F(M)$, positive and between zero and one, is called *effect*. Note that observables are given here by generally noncommuting operators. Moreover we have not requested $F(M)$ to be a projection operator, that is to say a self-adjoint and idempotent operator such that $F^2(M) = F(M)$. If this further condition holds for all $M \in \mathcal{B}(\mathbb{R}^3)$ one has a very special case of POVM, also called projection-valued measure (PVM), since it is a measure taking values in the space of projections on the Hilbert space $\mathcal{H}$. For such measures we shall use the symbol $E(M)$. In the case of PVM there is a one-to-one correspondence between the PVM and a uniquely defined self-adjoint operator, thus explaining the standard definition of observable as self-adjoint operator.

The operator associated to the PVM turns out to be a very convenient tool for the calculation of mean values and higher order moments, such as variances. Let us call $E^{\mathbf{A}}$ the PVM for the description of measurements on the quantity $\mathbf{A}$ taking values in $\mathbb{R}^k$. The first moment of the measure

$$\mathbf{A} = \int_{\sigma(\mathbf{A})} \mathbf{x} dE^{\mathbf{A}}(\mathbf{x})$$

actually identifies k commuting self-adjoint operators, the integral being calculated over the support of the measure or equivalently the spectrum of $\mathbf{A}$, and higher moments of the measure can be identified with powers of these operators, according to the functional calculus:

$$\mathbf{A}^n = \int_{\sigma(\mathbf{A})} \mathbf{x}^n dE^{\mathbf{A}}(\mathbf{x}).$$

In particular a whole collection of commuting self-adjoint operators can be obtained considering the integrals of a measurable function g from $\mathbb{R}^k$ to $\mathbb{R}$:

$$g(\mathbf{A}) = \int_{\sigma(\mathbf{A})} g(\mathbf{x}) dE^{\mathbf{A}}(\mathbf{x}),$$

corresponding to measurements of functions of the quantity $\mathbf{A}$. These facts are no longer true for a generic POVM.

Statistics of outcomes

Having introduced statistical operators as general mathematical representatives of a state, in the sense of characterization of preparation apparatuses, and POVM as mathematical representatives of observables, associated to registration apparatuses, we now have to combine states and observables in order to express the probabilities to be compared with the outcomes of an experiment. This is done by considering the duality relation between the spaces of states and observables. As in the classical

description the space of observables L^∞ is the dual of the space L^1 of states; here the Banach space of bounded operators is the dual of the space $\mathcal{T}(\mathcal{H})$ of trace class operators to which statistical operators do belong. The duality relation is given by the trace evaluation:

$$\begin{aligned} \mathrm{Tr} : \mathcal{B}(\mathcal{H}) \times \mathcal{T}(\mathcal{H}) &\to \mathbb{C}, \\ (X, w) &\to \mathrm{Tr}\, X^\dagger w, \end{aligned}$$

where taking any basis in $\mathcal{H}$, e.g., $\{u_n\}$, the trace can be evaluated as

$$\mathrm{Tr}\, X^\dagger w = \sum_n \langle u_n | X^\dagger w | u_n \rangle,$$

the series being convergent for any bounded operator X and trace class operator w and the result independent of the choice of orthonormal basis. Given a system prepared in the state ρ, the probability that a quantity described by the POVM F takes value in the set M and is given by the statistical formula:

$$\mathrm{Tr}\, \rho F(M). \tag{1}$$

The property of ρ and F ensure that $\mathrm{Tr}\, \rho F(M)$ is indeed a positive number between zero and one, and in particular for every pair ρ and F the mapping

$$\begin{aligned} \mathrm{Tr}\, \rho F(\cdot) : \mathcal{B}(\mathbb{R}^3) &\to [0, 1], \\ M &\to \mathrm{Tr}\, \rho F(M) \end{aligned}$$

is a classical probability measure assigning to each set M the probability $\mathrm{Tr}\, \rho F(M)$ that the outcome of the experiment lies in that set. For a given state to each observable one can therefore associate a classical probability measure; however only commuting observables are described by the same probability measure, to different observables one generally has to associate distinct probability measures. The formula (1) when considered for the particular case of a pure state $|\psi\rangle$ and a PVM E leads to the usual expression

$$\| E(M)\psi \|^2$$

for the evaluation of the statistics of an experiment measuring E once the system has been prepared in the state $|\psi\rangle$. Considering within this framework the usual notion of position and momentum observables one immediately realizes that the related measures can be expressed by means of the two well-distinct probability densities $|\psi(\mathbf{x})|^2$ and $|\tilde{\psi}(\mathbf{p})|^2$, respectively ($\tilde{\psi}$ denoting as usual the Fourier transform). At variance with the classical case there is in general no common probability density allowing to express the probability measure of all observables. There is in fact no sample space of elementary events. Note that for fixed F the mapping $\rho \to \mathrm{Tr}\, \rho F(\cdot)$

is an affine mapping from the convex set $\mathcal{K}(\mathcal{H})$ of statistical operators into the convex set of classical probability measures on $\mathcal{B}(\mathbb{R}^3)$. For a fixed observable this is all we need in order to compare with the experimental outcomes. Since any such affine mapping can be written in this form for a uniquely defined POVM one can come to the conclusion that POVM indeed provide the most general description of the statistics of experimental outcomes compatible with the probabilistic interpretation of quantum mechanics. The statistical formula (1) is the key point where the theory can be compared with experiment and allows us to better understand the meaning of the equivalence classes. Two preparation apparatuses are in the same equivalence class if they produce the very same statistics of outcomes for any observable, and similarly two registration apparatuses are in the same equivalence class if they lead to the same statistics of outcomes for any state.

2.3 Example of POVM for Position and Momentum

Now we want to consider a few examples of POVM and PVM concentrating on position and momentum, showing in particular how symmetry properties can be a very important guiding principle in the determination of meaningful observables, once we leave the correspondence principle focussed on quantum mechanics as a new mechanics with respect to the classical one. As we shall see while for the case of either position or momentum alone POVM are essentially given by a suitable coarse-graining with respect to the usual PVM; if one wants to give statistical predictions for the measurement of both position and momentum together the corresponding observable is given by necessity in terms of a POVM. For a more detailed and mathematically accurate exposition we refer the reader to [7, 13, 14].

Covariant mapping

Let us first start by introducing the notion of a mapping covariant under a given symmetry group G. As we will show this notion is of great interest in many situations, both for the construction of POVM and general dynamical mappings. Consider a measure space $\mathcal{X}$ with the σ-algebra of Borel sets $\mathcal{B}(\mathcal{X})$. Such a space is called a G-space if there exist an action of G on $\mathcal{X}$ defined as a mapping that sends group elements $g \in G$ to transformation mappings μ_g on $\mathcal{X}$ in such a way as to preserve group composition and identity

$$\mu_g \mu_h = \mu_{gh}, \qquad \forall g, h \in G, \qquad \mu_e = 1_{\mathcal{X}},$$

where $1_{\mathcal{X}}$ denotes the identity function on $\mathcal{X}$. If furthermore G acts transitively on $\mathcal{X}$, in the sense that any two point of $\mathcal{X}$ can be mapped one into the other with $\mu_{g'}$ for a suitable $g' \in G$, then $\mathcal{X}$ is called a transitive G-space. Consider for example $\mathcal{X} = \mathbb{R}^3$; then $\mathcal{X}$ is a transitive G-space with respect to the group of translations. The elements of the group are three-dimensional vectors acting in the obvious way on the Borel sets of $\mathbb{R}^3$, i.e., $\mu_{\mathbf{a}} = M + \mathbf{a}$ for all $\mathbf{a} \in \mathbb{R}^3$ and for all $M \in \mathcal{B}(\mathbb{R}^3)$.

Consider as well a unitary representation $U(g)$ of the same group G on a Hilbert space $\mathcal{H}$:

$$|\psi_g\rangle = U(g)|\psi\rangle \text{ for } \psi \in \mathcal{H} \text{ and } g \in G,$$

in terms of which one also has a representation of G on a space $\mathcal{A}(\mathcal{H})$ of operators acting on $\mathcal{H}$:

$$A_g = U^\dagger(g) A U(g) \text{ for } A \in \mathcal{A}(\mathcal{H}).$$

A mapping $\mathcal{M}$ defined on $\mathcal{B}(\mathcal{X})$ and taking values in $\mathcal{A}(\mathcal{H})$ is said to be covariant with respect to the symmetry group G provided it commutes with the action of the group in the sense that

$$U^\dagger(g)\mathcal{M}(X)U(g) = \mathcal{M}(\mu_{g^{-1}}(X)) \qquad \forall X \in \mathcal{B}(\mathcal{X}) \quad \forall g \in G. \tag{2}$$

A symmetry transformation on the domain of the mapping is mapped into the symmetry transformation corresponding to the same group element on the range of the mapping.

Position observable

As an example of an observable in the sense outlined above we now want to introduce the position observable. Rather than relying on the usual correspondence principle with respect to classical mechanics, we want to give an operational definition of the position observable, fixing its behavior with respect to the action of the relevant symmetry group, which in this case is the isochronous Galilei group, containing translations, rotations and boosts, that is to say velocity transformations. The group acts in the natural way on the Borel sets of $\mathbb{R}^3$, and the covariance equations that we require for an observable to be interpreted as position observable are the following:

$$\begin{aligned} U^\dagger(\mathbf{a})F^{\mathbf{x}}(M)U(\mathbf{a}) &= F^{\mathbf{x}}(M-\mathbf{a}) && \forall\, \mathbf{a} \in \mathbb{R}^3, \\ U^\dagger(\mathsf{R})F^{\mathbf{x}}(M)U(\mathsf{R}) &= F^{\mathbf{x}}(\mathsf{R}^{-1}M) && \forall\, \mathsf{R} \in SO(3), \\ U^\dagger(\mathbf{q})F^{\mathbf{x}}(M)U(\mathbf{q}) &= F^{\mathbf{x}}(M) && \forall\, \mathbf{q} \in \mathbb{R}^3. \end{aligned} \tag{3}$$

The mapping $F^{\mathbf{x}}$ to be interpreted as a position observable has to transform covariantly with respect to translations and rotations as in (2), and to be invariant under a velocity transformation. These equations can also be seen as a requirement on the possible macroscopic apparata possibly performing such a measurement. The apparatus used to test whether the considered system is localized in the translated region $M - \mathbf{a}$ should be in the equivalence class to which the translated apparatus used to test localization in the region M belongs, and similarly for rotations. Localization measurements should instead be unaffected by boost transformations. A solution of these covariance equations, that is to say a POVM complying with (3), is now a position observable. If one looks for such a solution asking moreover that the

POVM be in particular a PVM, the solution is uniquely given by the usual spectral decomposition of the position operator:

$$E^{\mathbf{x}}(M) = \chi_M(\hat{\mathbf{x}}) = \int_M d^3\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|, \tag{4}$$

where χ_M denotes the characteristic function of the set M. The first moment of the spectral measure gives the usual triple of commuting position operators:

$$\hat{\mathbf{x}} = \int_{\mathbb{R}^3} d^3\mathbf{x}\,\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|,$$

whose powers coincide with the higher moments of $E^{\mathbf{x}}$:

$$\hat{\mathbf{x}}^n = \int_{\mathbb{R}^3} d^3\mathbf{x}\,\mathbf{x}^n|\mathbf{x}\rangle\langle\mathbf{x}|.$$

In particular for a given state ρ mean values and variances of the classical probability distribution giving the position distribution can be expressed by means of the operator $\hat{\mathbf{x}}$:

$$\begin{aligned}\mathsf{Mean}(E^{\mathbf{x}}) &= \operatorname{Tr}\rho\hat{\mathbf{x}} = \langle\hat{\mathbf{x}}\rangle_\rho,\\ \mathsf{Var}(E^{\mathbf{x}}) &= \operatorname{Tr}\rho\hat{\mathbf{x}}^2 - (\operatorname{Tr}\rho\hat{\mathbf{x}})^2 = \langle\hat{\mathbf{x}}^2\rangle_\rho - \langle\hat{\mathbf{x}}\rangle_\rho^2.\end{aligned}$$

The pair $(U, E^{\mathbf{x}})$, where U is the unitary representation of the symmetry group, here the isochronous Galilei group, and $E^{\mathbf{x}}$ a PVM covariant under the action of U is called a system of imprimitivity. More generally a solution of (3) as a POVM is obtained as follows. Let us introduce a rotationally invariant probability density $h(\mathbf{x})$:

$$h(\mathbf{x}) \geq 0, \qquad \int d^3\mathbf{x}\,h(\mathbf{x}) = 1, \qquad h(\mathsf{R}\mathbf{x}) = h(\mathbf{x}),$$

with variance given by

$$\mathsf{Var}(h) = \int d^3\mathbf{x}\,\mathbf{x}^2 h(\mathbf{x}).$$

One can then indeed check that the expression

$$F^{\mathbf{x}}(M) = (\chi_M * h)(\hat{\mathbf{x}}) = \int_M d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\,h(\mathbf{x}-\mathbf{y})|\mathbf{x}\rangle\langle\mathbf{x}|, \tag{5}$$

where $*$ denotes convolution, actually is a POVM complying with (3), and in fact provides the general solution of (3). The POVM (5) actually is a smeared version of the usual sharp position observable, the probability density $h(\mathbf{x})$ which fixes the

POVM being understood as the actual, finite resolution of the registration apparatus. For any state ρ the first moment of the associated probability density can still be expressed as the mean value of the usual position operator, since $\mathsf{Mean}(h) = 0$:

$$\begin{aligned}\mathsf{Mean}(F^{\mathbf{x}}) &= \int_{\mathbb{R}^3} d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, \mathbf{y} h(\mathbf{x}-\mathbf{y}) \operatorname{Tr}\rho |\mathbf{x}\rangle\langle\mathbf{x}| \\ &= \langle \hat{\mathbf{x}} \rangle_\rho .\end{aligned}$$

The second moment however differs:

$$\begin{aligned}\mathsf{Var}(F^{\mathbf{x}}) &= \int_{\mathbb{R}^3} d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, \mathbf{y}^2 h(\mathbf{x}-\mathbf{y}) \operatorname{Tr}\rho |\mathbf{x}\rangle\langle\mathbf{x}| \\ &\quad - \left(\int_{\mathbb{R}^3} d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, \mathbf{y} h(\mathbf{x}-\mathbf{y}) \operatorname{Tr}\rho |\mathbf{x}\rangle\langle\mathbf{x}| \right)^2 \\ &= \langle \hat{\mathbf{x}}^2 \rangle_\rho - \langle \hat{\mathbf{x}} \rangle_\rho^2 + \mathsf{Var}(h) .\end{aligned}$$

It is not anymore expressed only by the mean value of the operator which can be used to evaluate the first moment and by its square. A further contribution $\mathsf{Var}(h)$ appears, which is state independent, and reflects the finite resolution of the equivalence class of apparatuses used for the localization measurement. Note that the usual result is recovered in the limit of a sharply peaked probability density $h(\mathbf{x}) \to \delta^3(\mathbf{x})$. Taking, e.g., a distribution of the form

$$h_\sigma(\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{3}{2}} e^{-\frac{1}{2\sigma^2}\mathbf{x}^2} \xrightarrow{\sigma\to 0} \delta^3(\mathbf{x}),$$

one has that in the limit of an infinite accuracy in the localization measurement of the apparatus exploited the POVM reduces to the standard PVM:

$$\begin{aligned}F^{\mathbf{x}}(M) &= \int_M d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x} \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{3}{2}} e^{-\frac{1}{2\sigma^2}(\mathbf{x}-\mathbf{y})^2} |\mathbf{x}\rangle\langle\mathbf{x}| \\ &\xrightarrow{\sigma\to 0} \int_M d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, \delta^3(\mathbf{x}-\mathbf{y}) |\mathbf{x}\rangle\langle\mathbf{x}| \\ &= \int_M d^3\mathbf{x}\, |\mathbf{x}\rangle\langle\mathbf{x}|.\end{aligned}$$

Analogous results can obviously be obtained for a momentum observable, asking for the corresponding covariance properties.

Position and momentum observable

A more interesting situation appears when considering apparatuses performing both a measurement of the spatial location of a particle as well as of its momentum. As it is well known no observable can be associated to such a measurement in the

framework of standard textbook quantum mechanics. Let us consider the covariance equations of such an observable in the more general framework of POVM. A position and momentum observable should be given by a POVM $F^{\mathbf{x},\mathbf{p}}$ defined on $\mathcal{B}(\mathbb{R}^3 \times \mathbb{R}^3)$ satisfying the following covariance equations under the action of translations, rotations and boosts, respectively:

$$\begin{aligned} U^\dagger(\mathbf{a})F^{\mathbf{x},\mathbf{p}}(M \times N)U(\mathbf{a}) &= F^{\mathbf{x},\mathbf{p}}((M - \mathbf{a}) \times N) && \forall\, \mathbf{a} \in \mathbb{R}^3, \\ U^\dagger(\mathsf{R})F^{\mathbf{x},\mathbf{p}}(M \times N)U(\mathsf{R}) &= F^{\mathbf{x},\mathbf{p}}(\mathsf{R}^{-1}M \times \mathsf{R}^{-1}N) && \forall\, \mathsf{R} \in SO(3), \\ U^\dagger(\mathbf{q})F^{\mathbf{x},\mathbf{p}}(M \times N)U(\mathbf{q}) &= F^{\mathbf{x},\mathbf{p}}(M \times (N - \mathbf{q})) && \forall\, \mathbf{q} \in \mathbb{R}^3. \end{aligned} \tag{6}$$

Such covariance equations, defining a position and momentum observable by means of its operational meaning, do not admit any solution within the set of PVM, while the general solution within the set of POVM is given by

$$F^{\mathbf{x},\mathbf{p}}(M \times N) = \frac{1}{(2\pi\hbar)^3} \int_M d^3\mathbf{x} \int_N d^3\mathbf{p}\, W(\mathbf{x}, \mathbf{p}) S W^\dagger(\mathbf{x}, \mathbf{p}), \tag{7}$$

where S is a trace class operator, positive, with trace equal to one and invariant under rotations

$$S \in \mathcal{T}(\mathcal{H}), \quad S \geq 0, \quad \mathrm{Tr}\, S = 1, \qquad U^\dagger(\mathsf{R}) S U(\mathsf{R}) = S,$$

so that it is in fact a statistical operator, even though it does not have the meaning of a state, while the unitaries

$$W(\mathbf{x}, \mathbf{p}) = e^{-\frac{\mathrm{i}}{\hbar}(\mathbf{x}\cdot\hat{\mathbf{p}} - \hat{\mathbf{x}}\cdot\mathbf{p})}$$

are the Weyl operators built in terms of the canonical position and momentum operators. The covariance of (7) under (6) can be directly checked, together with its normalization, working with the matrix elements of the operator expression. The couple $(U, F^{\mathbf{x},\mathbf{p}})$, where U is the unitary representation of the symmetry group and $F^{\mathbf{x},\mathbf{p}}$ a POVM covariant under its action, is now called system of covariance. The connection with position and momentum observables as well as the reason why such a joint observable can be expressed only in the formalism of POVM, where position observables alone are generally given by smeared versions of the usual position observable, and similarly for momentum, can be understood looking at the marginal observables. Starting from (7) one can in fact consider a measure of position irrespective of the momentum of the particle, thus coming to the marginal position observable:

$$
\begin{aligned}
F^{\mathbf{x}}(M) &= F^{\mathbf{x},\mathbf{p}}(M \times \mathbb{R}^3) \\
&= \int_M d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, |\mathbf{x}\rangle\langle \mathbf{x}-\mathbf{y}|S|\mathbf{x}-\mathbf{y}\rangle\langle \mathbf{x}| \\
&= \int_M d^3\mathbf{y} \int_{\mathbb{R}^3} d^3\mathbf{x}\, h_{S^{\mathbf{x}}}(\mathbf{x}-\mathbf{y})|\mathbf{x}\rangle\langle \mathbf{x}|,
\end{aligned}
$$

where the function

$$
h_{S^{\mathbf{x}}}(\mathbf{x}) = \langle \mathbf{x}|S|\mathbf{x}\rangle \tag{8}
$$

is a well-defined probability density due to the fact that the operator S has all the properties of a statistical operator, so that $\langle \mathbf{x}|S|\mathbf{x}\rangle$ would be the position probability density of a system described by the state S. On similar grounds the marginal momentum observable is given by

$$
\begin{aligned}
F^{\mathbf{p}}(N) &= F^{\mathbf{x},\mathbf{p}}(\mathbb{R}^3 \times N) \\
&= \int_N d^3\mathbf{q} \int_{\mathbb{R}^3} d^3\mathbf{p}\, |\mathbf{p}\rangle\langle \mathbf{p}-\mathbf{q}|S|\mathbf{p}-\mathbf{q}\rangle\langle \mathbf{p}| \\
&= \int_N d^3\mathbf{q} \int_{\mathbb{R}^3} d^3\mathbf{p}\, h_{S^{\mathbf{p}}}(\mathbf{p}-\mathbf{q})|\mathbf{p}\rangle\langle \mathbf{p}|,
\end{aligned}
$$

where again the function

$$
h_{S^{\mathbf{p}}}(\mathbf{p}) = \langle \mathbf{p}|S|\mathbf{p}\rangle \tag{9}
$$

is a well-defined probability density, which corresponds to the momentum probability density of a system described by the statistical operator S. As it appears the marginal observables are given by two POVM characterized by a smearing of the standard position and momentum observables by means of the probability densities $h_{S^{\mathbf{x}}}(\mathbf{x})$ and $h_{S^{\mathbf{p}}}(\mathbf{p})$, respectively. It is exactly this finite resolution in the measurement of both position and momentum, with two probability densities satisfying

$$
\mathsf{Var}_i(h_{S^{\mathbf{x}}})\mathsf{Var}_i(h_{S^{\mathbf{p}}}) \geq \frac{\hbar^2}{4} \text{ for } i = x, y, z
$$

as follows from (8) and (9), that allows for a joint measurement for position and momentum in quantum mechanics, without violating Heisenberg's uncertainty relations. In order to consider a definite example we take S to be a pure state corresponding to a Gaussian of width σ:

$$
\langle \mathbf{x}|\psi\rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{3}{4}} e^{-\frac{1}{4\sigma^2}\mathbf{x}^2},
$$

on which the Weyl operators act as a translation in both position and momentum, leading to

$$\langle \mathbf{x}|W(\mathbf{x}_0,\mathbf{p}_0)|\psi\rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{3}{4}} e^{-\frac{1}{4\sigma^2}(\mathbf{x}-\mathbf{x}_0)^2+\frac{\mathrm{i}}{\hbar}\mathbf{p}_0\cdot(\mathbf{x}-\mathbf{x}_0)} = \langle \mathbf{x}|\psi_{\mathbf{x}_0,\mathbf{p}_0}\rangle.$$

In particular one has

$$h_{\psi^{\mathsf{x}}}(\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{3}{2}} e^{-\frac{1}{2\sigma^2}\mathbf{x}^2}, \qquad \mathsf{Var}_i(h_{\psi^{\mathsf{x}}}) = \sigma^2 \qquad \text{for } i = x, y, z$$

and

$$h_{\psi^{\mathsf{p}}}(\mathbf{p}) = \left(\frac{2\sigma^2}{\pi\hbar^2}\right)^{\frac{3}{2}} e^{-\frac{2\sigma^2}{\hbar^2}\mathbf{p}^2}, \qquad \mathsf{Var}_i(h_{\psi^{\mathsf{p}}}) = \frac{\hbar^2}{4\sigma^2} \qquad \text{for } i = x, y, z$$

so that

$$\mathsf{Var}_i(h_{\psi^{\mathsf{x}}})\mathsf{Var}_i(h_{\psi^{\mathsf{p}}}) = \frac{\hbar^2}{4} \qquad i = x, y, z.$$

The POVM now reads

$$F^{\mathbf{x},\mathbf{p}}(M\times N) = \frac{1}{(2\pi\hbar)^3}\int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0\, |\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}| \tag{10}$$

with marginals

$$F^{\mathbf{x}}(M) = \int_M d^3\mathbf{x}_0 \int_{\mathbb{R}^3} d^3\mathbf{x} \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{3}{2}} e^{-\frac{1}{2\sigma^2}(\mathbf{x}-\mathbf{x}_0)^2} |\mathbf{x}\rangle\langle\mathbf{x}|$$

and

$$F^{\mathbf{p}}(N) = \int_N d^3\mathbf{p}_0 \int_{\mathbb{R}^3} d^3\mathbf{p} \left(\frac{2\sigma^2}{\pi\hbar^2}\right)^{\frac{3}{2}} e^{-\frac{2\sigma^2}{\hbar^2}(\mathbf{p}-\mathbf{p}_0)^2} |\mathbf{p}\rangle\langle\mathbf{p}|$$

for position and momentum, respectively. It is now clear that depending on the value of σ one can have more or less coarse-grained position and momentum observables. No limit on σ can however be taken in order to have a sharp observable for both position and momentum. In the limit $\sigma \to 0$ one has as before $F^{\mathbf{x}} \to E^{\mathbf{x}}$, but the marginal for momentum would identically vanish, intuitively corresponding to a complete lack of information on momentum, and vice versa.

As a last remark we would like to stress that despite the fact that PVM are only a very particular case of POVM, corresponding to most accurate measurements,

self-adjoint operators, which are in one-to-one correspondence to PVM, do play a distinguished and very important role in quantum mechanics as generators of symmetry transformations, according to Stone's theorem.

2.4 Measurements and State Transformations

The formalism presented up to now allows to describe in the most general way the state corresponding to a given preparation procedure and the statistics of the experimental outcomes obtained by feeding a certain registration procedure by such a state. It is however also of interest to have information not only on the statistics of the outcomes, which amounts to provide a classical probability density, but also to specify the state obtained as a consequence of such a measurement, provided the system does not simply get absorbed. This makes it possible to deal, e.g., with repeated consecutive measurements, allowing for a description of continual measurement in quantum mechanics [15], as well as to use the combination of initial preparation and registration apparata altogether as a new preparation apparatus, preparing states according to the value of a certain observable.

State transformations with a measuring character as instruments

The mathematical object characterizing a state transformation as a consequence of a given measurement is called instrument and is generally given by a mapping defined on the outcome space of the measurement, e.g., $\mathcal{B}(\mathbb{R}^3)$ in the examples we have considered, and taking values in the space of bounded mappings acting on the space of trace class operators, obeying the following requirements:

$$\begin{aligned} \mathcal{F}(\cdot) : \mathcal{B}(\mathbb{R}^3) &\to \mathcal{B}(\mathcal{T}(\mathcal{H})), \\ M &\to \mathcal{F}(M), \end{aligned}$$

$$\begin{aligned} &\mathrm{Tr}\,\mathcal{F}(\mathbb{R}^3)[\rho] = \mathrm{Tr}\,\rho, \\ &\mathcal{F}(\cup_i M_i) = \sum_i \mathcal{F}(M_i) \,\text{if}\, M_i \cap M_j = \emptyset \,\text{for}\, i \neq j\,, \end{aligned}$$

where for each $M \in \mathcal{B}(\mathbb{R}^3)$ the mapping $\mathcal{F}(M)$ is completely positive and generally trace decreasing, transforming trace class operators into trace class operators. $\mathcal{F}(M)$ is often called operation, and the mapping $\mathcal{F}$ is normalized in the sense that $\mathcal{F}(\mathbb{R}^3)$ is trace preserving. As for any completely positive mapping for fixed M one has for $\mathcal{F}(M)$ a Kraus representation:

$$\mathcal{F}(M)[\rho] = \sum_i V_i \rho V_i^\dagger .$$

The trace class operator

$$\mathcal{F}(M)[\rho_{\text{in}}]$$

whose trace is generally less than one, describes the subcollection of systems obtained by asking the outcome of the measurement to be in $M \in \mathcal{B}(\mathbb{R}^3)$. After the measurement in fact the transformed system can be sorted according to the outcome of the measurement itself. The transformed state according to a measurement without readout, i.e., without making any selection with respect to the result of the measurement (the so-called a priori state) is given by

$$\rho_{\text{out}} = \mathcal{F}(\mathbb{R}^3)[\rho_{\text{in}}],$$

the mapping $\mathcal{F}(\mathbb{R}^3)$ now only describing the modification on the incoming state ρ_{in} as a consequence of its interaction with the registration apparatus. The transformed state conditioned on the result of the measurement, i.e., the new state obtained by sorting out only the systems for which the outcome of the measurement was in $M \in \mathcal{B}(\mathbb{R}^3)$ (the so-called a posteriori state) is given by

$$\rho_{\text{out}}(M) = \frac{\mathcal{F}(M)[\rho_{\text{in}}]}{\operatorname{Tr} \mathcal{F}(M)[\rho_{\text{in}}]}.$$

Of course an instrument does not only provide the transformed trace class operator describing the system after its interaction with the registration apparatus performing the measurement, but also the statistics of the outcomes. The probability of an outcome $M \in \mathcal{B}(\mathbb{R}^3)$ for a measurement described by the instrument $\mathcal{F}$ on an incoming state ρ is given by the formula:

$$\operatorname{Tr} \mathcal{F}(M)[\rho],$$

which can also be expressed by means of a POVM F uniquely determined by the instrument $\mathcal{F}$ as follows:

$$F(M) = \mathcal{F}'(M)[\mathbb{1}],$$

where $\mathcal{F}'$ denotes the adjoint mapping with respect to the trace evaluation

$$\operatorname{Tr} \mathcal{F}(M)[\rho] = \operatorname{Tr} \mathbb{1}(\mathcal{F}(M)[\rho]) = \operatorname{Tr}(\mathcal{F}'(M)[\mathbb{1}])\rho = \operatorname{Tr} F(M)[\rho].$$

Note that the correspondence between instruments and POVM is not one-to-one. In fact there are different registration apparatuses, possibly leading to quite different transformations on the incoming state, which however all provide a measurement of the same observable. As a consequence while an instrument uniquely defines the associated POVM as outlined above, generally infinitely many different instruments are compatible with a given POVM, corresponding to different macroscopic implementation of measurements of the same observable. Note further that if the mapping $\mathcal{F}$ is reversible it is necessarily given by a unitary transformation and therefore does

not have any measuring character, the system transforms in a reversible way because of its free evolution described by a self-adjoint operator. A very special case of instrument can be obtained starting from the knowledge of an observable given as self-adjoint operator, $A = \sum_i a_i E_i$, where $\{E_i\}$ is a collection of mutually orthogonal projection operators, summing up to the identity, $E_i = E_i^2$ and $\sum_i E_i = \mathbb{1}$. Then the mapping

$$\mathcal{F}(M)[\rho] = \sum_{\{i\,|a_i \in M\}} E_i \rho E_i$$

actually is an instrument describing the state transformation of an incoming state ρ as a consequence of the measurement of the observable A as predicted by von Neumann's projection postulate. If the experimenter detects the value a_i for A the transformed state is given by

$$\mathcal{F}(\{a_i\})[\rho] = E_i \rho E_i .$$

This instrument has the peculiar property of being repeatable, in the sense that from

$$\mathcal{F}(M)[\mathcal{F}(N)[\rho]] = \sum_{\{i\,|a_i \in M \cap N\}} E_i \rho E_i ,$$

it follows

$$\mathcal{F}(\{a_i\})[\mathcal{F}(\{a_i\})[\rho]] = \mathcal{F}(\{a_i\})[\rho] = E_i \rho E_i ,$$

that is to say subsequent measurements of the same observable do always lead to the same result, which implicitly means an absolute precision in the measurement of the observable. As it appears this is a very particular situation, which can only be realized for a measurement in the sense of PVM of an observable with discrete spectrum.

Example of instrument for position and momentum

We now provide an example of an instrument corresponding to the description of a state transformation taking place by jointly measuring position and momentum of a particle in $L^2(\mathbb{R}^3)$, whose uniquely associated POVM is just the one given in (10). Consider in fact the mapping

$$\mathcal{F}^{\mathbf{x},\mathbf{p}}(M \times N)[\rho] = \frac{1}{(2\pi\hbar)^3} \int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0 \, |\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}|\rho|\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}| ,$$

built in terms of the normalized Gaussian wave packets $|\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle$ centered in $(\mathbf{x}_0, \mathbf{p}_0)$. This mapping depends on the interval $M \times N$ as a σ-additive measure, thanks to the fact that it is expressed by means of an operator density with respect to the Lebesgue

measure. Normalization is ensured by the completeness relation for Gaussian wave packets

$$\frac{1}{(2\pi\hbar)^3}\int_{\mathbb{R}^3} d^3\mathbf{x}_0 \int_{\mathbb{R}^3} d^3\mathbf{p}_0\, |\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}| = \mathbb{1},$$

leading to

$$\begin{aligned}\mathrm{Tr}\, \mathcal{F}^{\mathbf{x},\mathbf{p}}(\mathbb{R}^3\times\mathbb{R}^3)[\rho] &= \frac{1}{(2\pi\hbar)^3}\int_{\mathbb{R}^3} d^3\mathbf{x}_0 \int_{\mathbb{R}^3} d^3\mathbf{p}_0\, \langle\psi_{\mathbf{x}_0\mathbf{p}_0}|\rho|\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle,\\ &= \mathrm{Tr}\,\rho,\end{aligned}$$

while the adjoint mapping acting on the identity operator

$$\begin{aligned}\mathcal{F}^{\mathbf{x},\mathbf{p}'}(M\times N)[\mathbb{1}] &= \frac{1}{(2\pi\hbar)^3}\int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0\, |\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}|\mathbb{1}|\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}|\\ &= \frac{1}{(2\pi\hbar)^3}\int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0\, |\psi_{\mathbf{x}_0\mathbf{p}_0}\rangle\langle\psi_{\mathbf{x}_0\mathbf{p}_0}|\\ &= F^{\mathbf{x},\mathbf{p}}(M\times N)\end{aligned}$$

immediately gives the joint position and momentum POVM considered in (10). More generally, one can consider an instrument of the form:

$$\begin{aligned}\mathcal{F}^{\mathbf{x},\mathbf{p}}(M\times N) = \frac{1}{(2\pi\hbar)^3}\int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0\\ \times W(\mathbf{x}_0,\mathbf{p}_0)\sqrt{S}W^\dagger(\mathbf{x}_0,\mathbf{p}_0)\rho W(\mathbf{x}_0,\mathbf{p}_0)\sqrt{S}W^\dagger(\mathbf{x}_0,\mathbf{p}_0),\end{aligned}$$

which is again well defined due to the relation

$$\frac{1}{(2\pi\hbar)^3}\int_{\mathbb{R}^3} d^3\mathbf{x}_0 \int_{\mathbb{R}^3} d^3\mathbf{p}_0\, W(\mathbf{x}_0,\mathbf{p}_0)SW^\dagger(\mathbf{x}_0,\mathbf{p}_0) = \mathbb{1},$$

where S is a positive operator given by a statistical operator invariant under rotations, and whose adjoint mapping applied to the identity

$$\begin{aligned}\mathcal{F}^{\mathbf{x},\mathbf{p}'}(M\times N)[\mathbb{1}] &= \frac{1}{(2\pi\hbar)^3}\int_M d^3\mathbf{x}_0 \int_N d^3\mathbf{p}_0\, W(\mathbf{x}_0,\mathbf{p}_0)SW^\dagger(\mathbf{x}_0,\mathbf{p}_0)\\ &= F^{\mathbf{x},\mathbf{p}}(M\times N)\end{aligned}$$

coincides with the general expression for a covariant position and momentum POVM observable given in (7).

3 Open Systems and Covariance

In the previous section we have outlined the modern formulation of quantum mechanics, understood as a probability theory necessary for the description of the outcomes of statistical experiments involving microscopic systems. In this framework pure states are generally replaced by statistical operators, observables in the sense of self-adjoint operators are substituted by mappings taking values in an operator space, von Neumann's projection postulate is a very special case of mappings with a measuring character describing state transformations as a consequence of measurement. As it appears the notion of mapping taking values in an operator space becomes very natural and of great relevance. The knowledge of an instrument for example allows not only to predict the statistics of the outcomes of a certain measurement, but also the transformation of the state due to the interaction with the measuring apparatus. On the contrary the spontaneous transformation of the state of a closed system due to passing time is described by a very special kind of mappings, unitary time evolutions uniquely determined by fixing a self-adjoint operator. Considering more generally an open system, that is to say a system interacting with some other external system, its irreversible evolution in time as a consequence of this interaction is determined by a suitable mapping, whose characterization is a very intricate and interesting subject, together with the possibility of understanding and describing such an evolution as a measurement effected on the system. While it is relatively easy to state the general properties that should be obeyed by such mappings in order to provide a well-defined time evolution, the characterization of the structure of such mappings in its full generality is an overwhelmingly complicated problem. Important and quite general results can however be obtained considering constraints coming from physical or mathematical considerations. Also in this section we will aim at a brief introduction of key concepts, skipping all mathematical details, referring the reader to [16, 8, 7, 9, 17] for a more exhaustive presentation.

3.1 Constraints on Dynamical Mappings

In Sect. 2 we have already mentioned two important constraints applying to mappings describing how a state transforms in time, both as a consequence of a measurement or of its dynamical evolution. Actually the two situations are not of a completely different nature, even though in the first case the time extension of the interaction between system and measuring apparatus is typically though not necessarily assumed to be very short and neglected, so that the whole transformation is considered as a one-step process. In the case of the evolution of a closed or open system on the contrary the explicit time dependence of the mapping is essential, while a decomposition of the mapping according to the measurement outcome for a certain observable is generally not available.

Complete positivity

The first constraint was the by now well-known requirement of complete positivity of the mapping. It is a mathematical condition, which at the beginning was

somewhat mistrusted by physicists, naturally coming in the foreground in quantum mechanics because of the tensor product structure of the space in which to describe composite systems, thus playing an important role in the theory of entanglement. A completely positive mapping is a mapping which remains positive when extended in a trivial way, i.e., by taking the tensor product with the identity mapping, on a composite Hilbert space. Considering a positive mapping $\mathcal{M}$ in the Schrödinger picture, acting on the space of states $\mathcal{T}(\mathcal{H})$:

$$\begin{aligned} \mathcal{M} : \mathcal{T}(\mathcal{H}) &\to \mathcal{T}(\mathcal{H}) \\ \rho &\to \mathcal{M}[\rho], \end{aligned}$$

complete positivity amounts to the requirement that the mapping

$$\begin{aligned} \mathcal{M}_n : \mathcal{T}(\mathcal{H} \otimes \mathbb{C}^n) &\to \mathcal{T}(\mathcal{H} \otimes \mathbb{C}^n) \\ \rho \otimes \sigma_n &\to \mathcal{M}[\rho] \otimes \sigma_n \end{aligned}$$

is positive for any $n \in \mathbb{N}$, with σ_n a statistical operator in $\mathbb{C}^n$. An equivalent requirement can be formulated on the adjoint mapping $\mathcal{M}'$ in Heisenberg picture acting on the space of observables $\mathcal{B}(\mathcal{H})$

$$\begin{aligned} \mathcal{M}' : \mathcal{B}(\mathcal{H}) &\to \mathcal{B}(\mathcal{H}), \\ B &\to \mathcal{M}'[B]. \end{aligned}$$

$\mathcal{M}'$ is completely positive provided

$$\sum_{i,j=1}^{n} \langle \psi_i | \mathcal{M}'(B_i^\dagger B_j) \psi_j \rangle \geq 0 \text{ for } \{\psi_i\} \subset \mathcal{H} \text{ and } \{B_i\} \subset \mathcal{B}(\mathcal{H})$$

for any $n \in \mathbb{N}$. As it was shown by Kraus, any completely positive mapping can be expressed as follows:

$$\mathcal{M}[\rho] = \sum_i V_i \rho V_i^\dagger$$

with a suitable collection $\{V_i\}$ of operators also called Kraus operators, as already mentioned in Sect. 2.4. As it appears from this fundamental result, of great significance in applications, the condition of complete positivity is quite restrictive, thus allowing for important characterizations.

Covariance

Another important constraint, this time however only arising in the presence of symmetries, is given by the requirement of covariance, already considered in Sec. 2.3. For the case of a mapping defined in an operator space, e.g., in the Schrödinger picture sending statistical operators to statistical operators, given a unitary represen-

tation $U(g)$ of the group G on $\mathcal{H}$, the requirement of covariance can be expressed as follows:

$$\mathcal{M}[U(g)\rho U^\dagger(g)] = U(g)\mathcal{M}[\rho]U^\dagger(g) \qquad \forall g \in G. \tag{11}$$

This condition expresses the fact that the action of the mapping and of the representation of G on $\mathcal{T}(\mathcal{H})$ commute, and automatically implies the same property for the adjoint mapping $\mathcal{M}'$ acting on the space of observables $\mathcal{B}(\mathcal{H})$. Such a condition typically applies when a symmetry, available in the system one is studying, is not spoiled by the transformations brought about on the system by letting it interact with another system, be it a reservoir or a measuring apparatus. The possibility of giving a general solution of the covariance equation (11) obviously depends on the unitary representation of the group and on the class of mappings considered, possibly giving very detailed information on the general structure of such mappings.

Semigroup evolution

For the case of a closed system we know that the mapping giving the reversible time evolution is a one-parameter group of unitary transformations, fixed by a self-adjoint Hamiltonian according to Stone's theorem. A broader class of time evolutions allowing for an irreversible dynamics can be obtained by relaxing the group property to a semigroup composition law, corresponding to the existence of a preferred time direction. In particular one can introduce a so-called quantum-dynamical semigroup, which is a collection of one-parameter mappings $\{\mathcal{U}_t\}_{t\in\mathbb{R}_+}$ such that

$$\begin{aligned} \mathcal{U}_t : \mathcal{T}(\mathcal{H}) &\to \mathcal{T}(\mathcal{H}), \\ \rho &\to \mathcal{U}_t[\rho], \end{aligned}$$

is completely positive and trace preserving for any $t \geq 0$, for $t = 0$ one has the identity mapping, and the following semigroup composition law applies:

$$\mathcal{U}_t = \mathcal{U}_{t-s}\mathcal{U}_s \qquad \forall t \geq s \geq 0. \tag{12}$$

The semigroup condition (12) is sometimes called Markov condition, because it expresses the fact that the time evolution of the system does not exhibit memory effects, in analogy to the notion of Markov semigroup in classical probability theory. It tells us that the evolution up to time t can be obtained by arbitrarily composing the evolution mapping up to an intermediate time s with an evolution mapping depending only on the residual time $t - s$ acting on the state $\rho_s = \mathcal{U}_s[\rho_0]$, not referring to the knowledge of the state at previous times $\{\rho_{t'}\}_{0\leq t'\leq s}$. The requirement of complete positivity for this family of mappings allows for a most important characterization of the so-called generator $\mathcal{L}$ of the quantum-dynamical semigroup, which is the mapping giving the infinitesimal time evolution, defined through the relation:

$$\mathcal{U}_t = e^{t\mathcal{L}}.$$

According to a celebrated result of Gorini, Kossakowski, Sudarshan and Lindblad [18, 19] of enormous relevance in the applications, the general structure of the generator of a quantum-dynamical semigroup is given in the Schrödinger picture by

$$\mathcal{L}[\rho] = -\frac{\mathrm{i}}{\hbar}[H, \rho] + \sum_j [L_j \rho L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, \rho\}], \tag{13}$$

where H is a self-adjoint operator, and the operators L_j are often called Lindblad operators. The expression (13) is also called a master equation, since it provides the infinitesimal time evolution of the statistical operator, according to $d\rho/dt = \mathcal{L}[\rho]$. The key point is now obviously to determine the explicit expression of H and L_j relevant for the reduced dynamics of the physical system of interest, typically depending on the external reservoir and the details of the interaction mechanism. As we shall see further restrictions on $\mathcal{L}$ can arise as a consequence of an available symmetry not destroyed by the interaction. Note that introducing the operator

$$K = \frac{\mathrm{i}}{\hbar} H + \frac{1}{2}\sum_j L_j^\dagger L_j,$$

where the effective Hamiltonian H appears together with an operator $\frac{\hbar}{2}\sum_j L_j^\dagger L_j$ which can be formally seen as an imaginary, optical effective potential, the Lindblad structure (13) can also be written as

$$\mathcal{L}[\rho] = -K\rho - \rho K^\dagger + \sum_j L_j \rho L_j^\dagger,$$

leading to a Dyson expansion of the semigroup evolution

$$\mathcal{U}_t[\rho] = e^{t\mathcal{L}}[\rho] = \mathcal{K}_t[\rho] + \sum_{n=1}^{\infty} \int \dots \int_{0\le t_1\le \dots \le t_n \le t} dt_1 \dots dt_n \, \mathcal{K}_{t_1}[\mathcal{J}[\mathcal{K}_{t_2-t_1} \dots \mathcal{J}[\mathcal{K}_{t-t_n}[\rho]] \dots]], \tag{14}$$

where the superoperators

$$\mathcal{J}[\rho] = \sum_j L_j \rho L_j^\dagger \qquad \mathcal{K}_t[\rho] = e^{-Kt}\rho e^{-K^\dagger t}$$

appear. The formal solution of the time evolution (14) has a quite intuitive physical meaning; in fact it can be seen as a sequence of relaxing evolutions over a time interval t given by the contraction semigroup $\mathcal{K}_t$ interrupted by jumps described by the completely positive mapping $\mathcal{J}$. The sum of over all possible such evolutions gives the final state. Most recently a general characterization has been obtained also for a class of non-Markovian time evolutions which provides a kind of generalization of

the Lindblad result (13), in the sense that the state evolved up to a given time can be expressed as a mixture of subcollections (that is positive trace class operators with trace less than one) each obeying a Lindblad type of master equation, however, with different Lindblad operators [20–24], so that the overall time evolution does no more obey a semigroup composition law as in (12).

3.2 Shift-Covariance and Damped Harmonic Oscillator

As a first example of a master equation corresponding to a completely positive quantum-dynamical semigroup let us consider the well-known master equation for a damped harmonic oscillator [25], describing, e.g., the damping of an electromagnetic field mode in a cavity. The Hilbert space of the single mode is given by $\mathcal{H} = l^2(\mathbb{C})$, the square summable sequences over the complex field, with basis $\{|n\rangle\}$, the ket $|n\rangle$ denoting as usual the eigenvector with eigenvalue $n \in \mathbb{N}$ of the number operator $N = a^\dagger a$, $a^\dagger$ and a being respectively creation and annihilation operators of a photon in the given mode of frequency ω. The master equation then reads

$$\begin{aligned}\mathcal{L}_{\mathrm{DHO}}[\rho] = &-\frac{\mathrm{i}}{\hbar}[H_0(N), \rho] + \eta(N_\beta(\omega)+1)[a\rho a^\dagger - \frac{1}{2}\{a^\dagger a, \rho\}] \\ &+\eta N_\beta(\omega)[a^\dagger \rho a - \frac{1}{2}\{aa^\dagger, \rho\}],\end{aligned} \tag{15}$$

where $N_\beta(\omega)$ denotes the average of the photon number operator over a thermal distribution

$$N_\beta(\omega) = \frac{1}{e^{\beta\hbar\omega}-1} = \frac{1}{2}[\coth(\beta\hbar\omega/2) - 1],$$

$H_0(N) = \hbar\omega N$ is the free Hamiltonian and η the relaxation rate.

Dissipation and decoherence for the damped harmonic oscillator

As it is well known such a master equation describes both classical dissipative effects as well as quantum decoherence effects. To see this let us first focus on dissipative effects, considering the time evolution of the mean amplitude and mean number of quanta in the mode. Considering the adjoint mapping of (15), giving the time evolution in Heisenberg picture

$$\begin{aligned}\mathcal{L}'_{\mathrm{DHO}}[X] = &+\frac{\mathrm{i}}{\hbar}[H_0(N), \rho] + \eta(N_\beta(\omega)+1)[a^\dagger \rho a - \frac{1}{2}\{a^\dagger a, \rho\}] \\ &+\eta N_\beta(\omega)[a\rho a^\dagger - \frac{1}{2}\{aa^\dagger, \rho\}],\end{aligned}$$

one can solve the Heisenberg equations of motion for $X \to a$ and $X \to N = a^\dagger a$, finally obtaining

$$\langle a(t)\rangle = \mathrm{Tr}(a(t)\rho) = \langle a\rangle e^{-\mathrm{i}\omega t - \frac{\eta}{2}t},$$
$$\langle N(t)\rangle = \mathrm{Tr}(N(t)\rho) = \langle N\rangle e^{-\eta t} + N_\beta(\omega)(1 - e^{-\eta t}),$$

where $a(t)$, $N(t)$ denote Heisenberg operators at time t, with $a = a(0)$, $N = N(0)$ the corresponding Schrödinger operators. The mean amplitude of the mode thus rotates in the complex plane, vanishing for long enough times, with a decay rate given by $(\eta/2)^{-1}$; the population of the mode goes from the initial value to a final thermal distribution with a decay rate given by η^{-1}. For the study of decoherence we shall consider the time evolution of an initial state given by a coherent superposition of two coherent states characterized by two amplitudes α and β [9]. Setting

$$\rho_0 = \frac{1}{\mathcal{N}_0}[|\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta| + |\alpha\rangle\langle\beta| + h.c.],$$

one can look for the time evolved state, exploiting the fact that coherent states remain coherent states under time evolution. Working for simplicity at zero temperature, so that $N_\beta(\omega) = 0$ one has

$$\rho_t = \frac{1}{\mathcal{N}_t}[|\alpha(t)\rangle\langle\alpha(t)| + |\beta(t)\rangle\langle\beta(t)| + e^{-\frac{1}{2}|\alpha-\beta|^2(1-e^{-\eta t})}|\alpha(t)\rangle\langle\beta(t)| + h.c.],$$

where $\alpha(t) = \alpha e^{-i\omega t - \frac{\eta}{2}t}$ and similarly for $\beta(t)$. As it immediately appears, the so-called coherences, that is to say the off-diagonal matrix elements of the statistical operator, are suppressed with respect to diagonal ones by a factor that for long enough times is given by

$$e^{-\frac{1}{2}|\alpha-\beta|^2} = |\langle\alpha|\beta\rangle|,$$

that is to say the modulus of the overlap of the two coherent states, a tiny quantity for two macroscopically distinguishable states of the electromagnetic field.

Structure of the mapping and covariance

It is immediately apparent that the master equation (15) is an example of a realization of the Lindblad structure (13) with just two Lindblad operators given by

$$L_1 = \sqrt{\eta(N_\beta(\omega) + 1)}a, \qquad L_2 = \sqrt{\eta N_\beta(\omega)}a^\dagger,$$

with Hamiltonian

$$H_0(N) = \hbar\omega N$$

and stationary solution

$$w \propto e^{-\beta H_0(N)}. \tag{16}$$

More than this it is also covariant under the action of the group $U(1)$. Consider in fact the unitary representation of $U(1)$ on $\mathcal{H} = l^2(\mathbb{C})$ given by

$$U(\theta) = e^{\mathrm{i}\theta N} \text{ with } \theta \in [0, 2\pi],$$

where N is the usual number operator. The master equation (15) is covariant under this unitary representation of the group $U(1)$ according to

$$\mathcal{L}_{\mathrm{DHO}}[U(\theta)\rho U^\dagger(\theta)] = U(\theta)\mathcal{L}_{\mathrm{DHO}}[\rho]U^\dagger(\theta) \tag{17}$$

as can be checked immediately. One can also show that this is the unique structure of master equation bilinear in a and $a^\dagger$ complying with the Lindblad structure, covariant under $U(1)$ and admitting (16) as a stationary state [26].

It is also possible to give a complete characterization of the structure of the generator of a quantum-dynamical semigroup covariant with respect to $U(1)$ as in (17). The general result has been obtained by Holevo [27, 28] and is given by the following expression:

$$\begin{aligned}\mathcal{L}[\rho] = &-\frac{\mathrm{i}}{\hbar}[H(N), \rho] + \sum_j \left[A_{0j}(N)\rho A_{0j}^\dagger(N) - \frac{1}{2}\left\{ A_{0j}^\dagger(N)A_{0j}(N), \rho \right\} \right] \\ &+ \sum_{m=1}^{\infty}\sum_j \left[W^m A_{mj}(N)\rho A_{mj}^\dagger(N)W^{\dagger m} - \frac{1}{2}\left\{ A_{mj}^\dagger(N)A_{mj}(N), \rho \right\} \right] \\ &+ \sum_{m=1}^{\infty}\sum_j \left[W^{\dagger m} A_{-mj}(N)\rho A_{-mj}^\dagger(N)W^m - \frac{1}{2}\left\{ A_{-mj}^\dagger(N)P_m A_{-mj}(N), \rho \right\} \right],\end{aligned}$$

where $A_{mj}(N)$ are functions of the number operator N, the generator of the symmetry; the operator W is given by

$$W = \sum_{n=0}^{\infty} |n+1\rangle\langle n|$$

acting as a shift $|n\rangle \to |n+1\rangle$ on the basis of eigenvectors of the number operator, so that this kind of symmetry is also called shift-covariance, while P_m is the projection on the subspace spanned by $\{|n\rangle\}_{n=m,\ldots,+\infty}$ given by

$$P_m \equiv \sum_{n=m}^{\infty} |n\rangle\langle n| = W^m W^{\dagger m}$$

and one further has

$$U(\theta)W^m = e^{i\theta m} W^m U(\theta), \tag{18}$$

which compared to (28) can be seen as a generalized Weyl relation, expressed by means of the isometric but not unitary operators W^m [29]. Examples of realizations of this general shift-covariant expression are given by the master equation for the damped harmonic oscillator as indicated above, corresponding to the choice

$$A_1(N) = \sqrt{\eta N_\beta(\omega)}\sqrt{N+1}, \qquad A_{-1}(N) = e^{\frac{1}{2}\beta\hbar\omega}\sqrt{\eta N_\beta(\omega)}\sqrt{N},$$
$$A_m(n) = 0\ m = 0, |m| > 1, \qquad H(N) = \hbar\omega N,$$

as can be checked immediately exploiting the polar representation for the creation and annihilation operators:

$$a = W^\dagger\sqrt{N}, \qquad a^\dagger = W\sqrt{N+1}.$$

A more general structure still preserving the stationary solution (16) is given by the choice

$$A_0(n) = \eta_0, \qquad A_m(N) = \sqrt{\eta_m}N_\beta^{\frac{m}{2}}(\omega)\sqrt{\frac{(N+m)!}{N!}},$$
$$A_{-m}(N) = e^{\frac{m}{2}\beta\hbar\omega}\sqrt{\eta_m}N_\beta^{\frac{m}{2}}(\omega)\sqrt{\frac{N!}{(N-m)!}}, \qquad H(N) = \hbar\omega N,$$

corresponding to

$$\mathcal{L}[\rho] = -\frac{\mathrm{i}}{\hbar}[H_0(N), \rho] - \eta_0[N, [N, \rho]]$$
$$+ \sum_{m=1}^{+\infty}\eta_m(N_\beta(\omega)+1)^m[a^m\rho a^{\dagger m} - \frac{1}{2}\{a^{\dagger m}a^m, \rho\}]$$
$$+ \sum_{m=1}^{+\infty}\eta_m N_\beta^m(\omega)[a^{\dagger m}\rho a^m - \frac{1}{2}\{a^m a^{\dagger m}, \rho\}],$$

where a phase damping term is given by a double commutator with the number operator, as well as many photon processes with different decay rates appear.

3.3 Rotation-Covariance and Two-Level System

Another example of a well-known master equation which can be characterized in terms of covariance properties comes from the description of a two-level system interacting with a thermal reservoir, e.g., a two-level atom in the presence of the radiation field or a spin in a magnetic field, so that the Hilbert space is now simply $\mathbb{C}^2$. It corresponds to the so-called Bloch equation and is typically used in quantum optics and magnetic resonance theory. Focussing on a two-level atom with transi-

tion frequency ω and spontaneous emission rate η interacting with the quantized electromagnetic field one has

$$\mathcal{L}_{2\mathrm{LS}}[\rho] = -\frac{\mathrm{i}}{\hbar}[H_0(\sigma_z), \rho] + \eta(N_\beta(\omega)+1)[\sigma_-\rho\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho\}] \\ +\eta N_\beta(\omega)[\sigma_+\rho\sigma_- - \frac{1}{2}\{\sigma_-\sigma_+, \rho\}], \tag{19}$$

where $N_\beta(\omega)$ is the thermal photon number at the transition frequency and as usual $\sigma_\pm = \frac{1}{2}(\sigma_x \pm \mathrm{i}\sigma_y)$ with $\{\sigma_i\}_{i=x,y,z}$, the Pauli matrices.

Dissipation and decoherence for the two-level system

As it is well known this master equation predicts relaxation to a stationary state which is diagonal in the basis of eigenvectors of the free Hamiltonian, with a relative population between ground and excited state determined by the temperature of the bath. This can be immediately seen considering as usual the adjoint mapping of (19):

$$\mathcal{L}'_{2\mathrm{LS}}[X] = +\frac{\mathrm{i}}{\hbar}[H_0(\sigma_z), X] + \eta(N_\beta(\omega)+1)[\sigma_+X\sigma_- - \frac{1}{2}\{\sigma_+\sigma_-, X\}] \\ +\eta N_\beta(\omega)[\sigma_-X\sigma_+ - \frac{1}{2}\{\sigma_-\sigma_+, X\}]$$

and solve the Heisenberg equations of motion for the operator $X \to P_e = |1\rangle\langle 1|$ representing the population in the excited state

$$\langle P_e(t)\rangle = \mathrm{Tr}(P_e(t)\rho) = \langle P_e\rangle e^{-\bar{\eta}t} + \frac{N_\beta(\omega)}{2N_\beta(\omega)+1}(1-e^{-\bar{\eta}t}),$$

which also fixes $\langle P_g(t)\rangle = 1 - \langle P_e(t)\rangle$ due to the normalization condition, where we have denoted as $\bar{\eta} = \eta(2N_\beta(\omega)+1)$ the total transition rate and $P_e = P_e(0)$. Due to the simplicity of the equation also the evolution in time of the coherences is easily determined, now considering the evolution in the Heisenberg picture of the operator $X \to C = |0\rangle\langle 1|$, which is given by

$$\langle C(t)\rangle = \mathrm{Tr}(C(t)\rho) = \langle C\rangle e^{-\mathrm{i}\omega t - \frac{\bar{\eta}}{2}t},$$

where again $C = C(0)$. With elapsing time the populations reach a stationary value, while coherences get suppressed.

pStructure of the mapping and covariance

In strict analogy to the results presented in Sect. 3.2 also the master equation (19) can be immediately recast in Lindblad form (13) with two Lindblad operators given by

$$L_1 = \sqrt{\eta(N_\beta(\omega)+1)}\sigma_-, \qquad L_2 = \sqrt{\eta N_\beta(\omega)}\sigma_+$$

and Hamiltonian

$$H_0(\sigma_z) = \frac{\hbar\omega}{2}\sigma_z\,,$$

admitting the stationary solution

$$w \propto e^{-\beta H_0(\sigma_z)}. \tag{20}$$

The master equation is also covariant under the group $SO(2)$, describing rotations along a given axis and being isomorphic to $U(1)$, which accounts for the similarity between the two equations, even though important differences appear due to the different Hilbert spaces in which the two systems are described and the group represented. Consider in fact the representation of $SO(2)$ in $\mathbb{C}^2$:

$$U(\phi) = e^{\frac{\mathrm{i}}{\hbar}\phi S_z} \qquad \phi \in [0, 2\pi]$$

with $S_z = \frac{\hbar}{2}\sigma_z$: one immediately checks that (19) is covariant in the sense that

$$\mathcal{L}_{2\mathrm{LS}}[U(\phi)\rho U^\dagger(\phi)] = U(\phi)\mathcal{L}_{2\mathrm{LS}}[\rho]U^\dagger(\phi). \tag{21}$$

Also in this case it is possible to provide the explicit expression of the generator of a rotation-covariant quantum-dynamical semigroup in the sense of (21). It takes the simple form [30]:

$$\mathcal{L}[\rho] = -\frac{\mathrm{i}}{\hbar}[H(\sigma_z), \rho] + \sum_{m=0,\pm1} c_m[T_{1m}\rho T_{1m}^\dagger - \frac{1}{2}\{T_{1m}^\dagger T_{1m}, \rho\}],$$

where the c_m are positive constants and T_{1m} are irreducible tensor operators given by

$$T_{11} = -\frac{1}{\sqrt{2}}(\sigma_x + \mathrm{i}\sigma_y), \qquad T_{10} = \sigma_z, \qquad T_{1-1} = \frac{1}{\sqrt{2}}(\sigma_x - \mathrm{i}\sigma_y)$$

or equivalently in terms of σ_z, σ_+ and σ_-, in order to allow for a direct comparison with (19):

$$\begin{aligned}\mathcal{L}[\rho] = &-\frac{\mathrm{i}}{\hbar}[H(\sigma_z), \rho] - \frac{c_0}{2}[\sigma_z, [\sigma_z, \rho]] \\ &+2c_{-1}[\sigma_-\rho\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho\}] + 2c_1[\sigma_+\rho\sigma_- - \frac{1}{2}\{\sigma_-\sigma_+, \rho\}].\end{aligned}$$

The use of the word rotation-covariance hints at the fact that such a master equation applies, e.g., to a spin $1/2$ in an environment with axial symmetry, so that one

has invariance under rotations along a given axis. More generally the full rotation group $SO(3)$ can be considered, and a general characterization also exists for the structure of generators of semigroups acting on $\mathbb{C}^n$ and invariant under $SO(3)$ [31], relevant for the description of relaxation of a spin j under the influence of isotropic surroundings.

3.4 Translation-Covariance and Quantum Brownian Motion

As a further example of the concepts introduced above we will consider the master equation for the description of quantum Brownian motion, which applies to the motion of a massive test particle in a gas of lighter particles. The Hilbert space of relevance is given here by $L^2(\mathbb{R}^3)$, and denoting the usual position and momentum operators as $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ the general structure of the master equation reads

$$\begin{aligned}\mathcal{L}_{\mathrm{QBM}}[\rho] = &-\frac{\mathrm{i}}{\hbar}[H_0(\hat{\mathbf{p}}), \rho] - \frac{\mathrm{i}}{\hbar}\frac{\eta}{2}\sum_i [\hat{x}_i, \{\hat{p}_i, \rho\}] \\ &-\frac{D_{pp}}{\hbar^2}\sum_i [\hat{x}_i, [\hat{x}_i, \rho]] - \frac{D_{xx}}{\hbar^2}\sum_i [\hat{p}_i, [\hat{p}_i, \rho]], \end{aligned} \tag{22}$$

where we have assumed isotropy for simplicity, the index $i = x, y, z$ denoting the different Cartesian coordinates and $H_0(\hat{\mathbf{p}}) = \hat{\mathbf{p}}^2/2M$ is the free Hamiltonian. This and similar types of master equation, bilinear in the position and momentum operators, do appear in different contexts, leading to different microscopic expressions for the coefficients, as well as to the appearance of other terms such as a double commutator $[\hat{x}_i, [\hat{p}_i, \rho]]$ with position and momentum operators [9, 32]. We will here have in mind the dynamics of a test particle interacting through collisions with a homogeneous gas [33–35], so that the coefficients read

$$D_{pp} = \eta\frac{M}{\beta}, \qquad D_{xx} = \eta\frac{\beta\hbar^2}{16M},$$

with M the mass of the test particle.

Dissipation and decoherence for quantum Brownian motion

As in the previous case we want to point briefly out how such a master equation describes both dissipative and decoherence effects. Regarding dissipation one has similarly to the classical case that the mean value of momentum is driven to zero, while the average of the squared momentum goes to the equipartition value fixed by the gas temperature. The effect of decoherence typically manifests itself in the fact that superpositions of spatially macroscopically distinguished states are quickly suppressed. Let us first focus on dissipation, considering the adjoint mapping of (22) in Heisenberg picture:

$$\mathcal{L}'_{\mathrm{QBM}}[X] = +\frac{\mathrm{i}}{\hbar}[H_0(\hat{\mathbf{p}}), X] + \frac{\mathrm{i}}{\hbar}\frac{\eta}{2}\sum_i \{\hat{p}_i, [\hat{x}_i, X]\} - \frac{D_{pp}}{\hbar^2}\sum_i [\hat{x}_i, [\hat{x}_i, X]] - \frac{D_{xx}}{\hbar^2}\sum_i [\hat{p}_i, [\hat{p}_i, X]].$$

As already mentioned the observables of interest are given by the momentum operators $X \to \hat{\mathbf{p}}$ and the kinetic energy $X \to E = \hat{\mathbf{p}}^2/2M$, whose mean values evolve according to

$$\langle \hat{\mathbf{p}}(t)\rangle = \mathrm{Tr}(\hat{\mathbf{p}}(t)\rho) = \langle \hat{\mathbf{p}}\rangle e^{-\eta t},$$
$$\langle E(t)\rangle = \mathrm{Tr}(E(t)\rho) = \langle E\rangle e^{-2\eta t} + \frac{3}{2\beta}(1 - e^{-2\eta t}),$$

where again $\hat{\mathbf{p}}(t)$ and $E(t)$ denote Heisenberg operators at time t. The average momentum thus relaxes to zero with a decay rate η^{-1}, while the mean kinetic energy reaches the equipartition value with a rate $(2\eta)^{-1}$. We now concentrate on the study of decoherence, both for position and momentum. This can be done considering the off-diagonal matrix elements in both position and momentum of a statistical operator evolved in time according to (22). To do this we exploit the knowledge of the exact solution [36], neglecting however the contribution $-\frac{\mathrm{i}}{\hbar}\frac{\eta}{2}\sum_i [\hat{x}_i, \{\hat{p}_i, \rho\}]$ responsible for dissipative effects, that would lead to too a cumbersome expression. Considering an initial state ρ_0 the state up to time t reads in the momentum representation:

$$\langle \mathbf{p}|\rho_t|\mathbf{q}\rangle = e^{-\frac{D_{xx}}{\hbar^2}(\mathbf{p}-\mathbf{q})^2 t} e^{-\frac{1}{12}\frac{D_{pp}}{\hbar^2}\left(\frac{(\mathbf{p}-\mathbf{q})}{M}t\right)^2 t} \times \left(\frac{\hbar^2}{4\pi D_{pp} t}\right)^{3/2} \int d^3\mathbf{k}\, e^{-\frac{\hbar^2\mathbf{k}^2}{4D_{pp}t}} \langle \mathbf{p}-\mathbf{k}|\rho_0|\mathbf{q}-\mathbf{k}\rangle.$$

It appears immediately that off-diagonal matrix elements are quickly suppressed with elapsing time, depending on their separation $(\mathbf{p}-\mathbf{q})^2$. The factor depending on the coefficient D_{xx} is due to the momentum localization term $\sum_i [\hat{p}_i, [\hat{p}_i, \rho]]$, while the factor where the coefficient D_{pp} appears is due to the position localization term $\sum_i [\hat{x}_i, [\hat{x}_i, \rho]]$; this also suppresses coherences in momentum because different momentum states quick-ly lead to spatial separation, so that the position localization mechanism is again of relevance. As far as coherences in position are concerned one has to consider the matrix elements of ρ_t in the position representation. Here it is convenient to express the exact solution ρ_t in terms of the solution ρ_t^S of the free Schrödinger equation. One has the quite cumbersome expression:

$$\langle \mathbf{x}|\rho_t|\mathbf{y}\rangle = e^{-\frac{D_{pp}}{\hbar^2}(\mathbf{x}-\mathbf{y})^2 t\left[1-\frac{D_{pp}}{4M^2}t^2\frac{1}{\left[D_{xx}+\frac{D_{pp}}{3M^2}t^2\right]}\right]}\left(\frac{\hbar^2}{4\pi\left[D_{xx}+\frac{D_{pp}}{3M^2}t^2\right]t}\right)^{3/2}$$
$$\times\int d^3\mathbf{z}\, e^{-\frac{\hbar^2\mathbf{z}^2}{4\left[D_{xx}+\frac{D_{pp}}{3M^2}t^2\right]t}} e^{\frac{\mathrm{i}}{\hbar}\frac{D_{pp}}{2M}\frac{\mathbf{z}\cdot(\mathbf{x}-\mathbf{y})t}{\left[D_{xx}+\frac{D_{pp}}{3M^2}t^2\right]}}\langle \mathbf{x}-\mathbf{z}|\rho_t^S|\mathbf{y}-\mathbf{z}\rangle,$$

which is essentially given by a convolution of the free solution with a Gaussian kernel, multiplied by an exponential factor suppressing off-diagonal matrix elements according to their distance $(\mathbf{x}-\mathbf{y})^2$ in space. Spatially macroscopically distant states are again very quickly suppressed.

Structure of the mapping and covariance

The master equation (22) can also be written manifestly in Lindblad form, as it can be seen introducing a single Lindblad operator for each Cartesian direction

$$L_i = \sqrt{\eta}a_i\,,$$

with

$$a_i = \frac{1}{\sqrt{2}\lambda_{\mathrm{th}}}\left(\hat{x}_i + \frac{\mathrm{i}}{\hbar}\lambda_{th}^2\hat{p}_i\right), \qquad \lambda_{\mathrm{th}} = \sqrt{\frac{\beta\hbar^2}{4M}}, \qquad [a_i, a_j^\dagger] = \delta_{ij}\,,$$

and the effective Hamiltonian

$$H = H_0(\hat{\mathbf{p}}) + \frac{\eta}{2}\sum_i\{\hat{x}_i, \hat{p}_i\},$$

leading to

$$\mathcal{L}_{\mathrm{QBM}}[\rho] = -\frac{\mathrm{i}}{\hbar}\left[H_0(\hat{\mathbf{p}}) + \frac{\eta}{2}\sum_i\{\hat{x}_i, \hat{p}_i\}, \rho\right] + \eta\sum_i[a_i\rho a_i^\dagger - \frac{1}{2}\{a_i^\dagger a_i, \rho\}]$$

with a stationary solution

$$w \propto e^{-\beta H_0(p)}. \tag{23}$$

Also the quantum Brownian motion master equation is characterized by covariance under the action of a symmetry group which in this case is the group $\mathbb{R}^3$ of translations. Given the unitary representation

$$U(\mathbf{a}) = e^{\frac{\mathrm{i}}{\hbar}\mathbf{a}\cdot\hat{\mathbf{p}}} \quad \text{for}\,\mathbf{a}\in\mathbb{R}^3 \tag{24}$$

of the group of translations on $L^2(\mathbb{R}^3)$ where the operators $\hat{\mathbf{p}}$ act as generator of the symmetry, one can indeed check immediately that (22) is covariant under this representation in the sense that

$$\mathcal{L}_{\mathrm{QBM}}[U(\mathbf{a})\rho U^\dagger(\mathbf{a})] = U(\mathbf{a})\mathcal{L}_{\mathrm{QBM}}[\rho]U^\dagger(\mathbf{a}). \tag{25}$$

In this case however, at odds with the case of the master equation for the damping harmonic oscillator, the three requirements of Lindblad structure, covariance under $\mathbb{R}^3$ and a stationary state given by (23) do not uniquely fix the form of a master equation at most bilinear in the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ [26].

A most important general characterization of structures of generators of quantum-dynamical semigroups covariant under translations has been obtained by Holevo, relying on a non-commutative quantum generalization [29, 37, 38] of the classical Lévy-Khintchine formula (see e.g. [9, 39]). In this case the generator can be written as follows:

$$\mathcal{L}[\rho] = -\frac{\mathrm{i}}{\hbar}\,[H(\hat{\mathbf{p}}),\,\rho] + \mathcal{L}_G[\rho] + \mathcal{L}_P[\rho],$$

where $H(\hat{\mathbf{p}})$ is a self-adjoint operator only depending on the momentum operators, while $\mathcal{L}_G$ and $\mathcal{L}_P$ correspond to a Gaussian and a Poisson component, as in the Lévy–Khintchine formula, and are given by

$$\begin{aligned}\mathcal{L}_G[\rho] = &-\frac{\mathrm{i}}{\hbar}\left[\hat{y}_0 + \frac{1}{2\mathrm{i}}\sum_{k=1}^{r}(\hat{y}_k L_k(\hat{\mathbf{p}}) - L_k^\dagger(\hat{\mathbf{p}})\hat{y}_k),\,\rho\right]\\ &+\sum_{k=1}^{r}\left[(\hat{y}_k + L_k(\hat{\mathbf{p}}))\rho(\hat{y}_k + L_k(\hat{\mathbf{p}}))^\dagger - \frac{1}{2}\left\{(\hat{y}_k + L_k(\hat{\mathbf{p}}))^\dagger(\hat{y}_k + L_k(\hat{\mathbf{p}})),\,\rho\right\}\right]\end{aligned} \tag{26}$$

and

$$\begin{aligned}\mathcal{L}_P[\rho] = &\int_{\mathbb{R}^3}\sum_{j=1}^{\infty}\left[U(\mathbf{q})L_j(\mathbf{q},\hat{\mathbf{p}})\rho L_j^\dagger(\mathbf{q},\hat{\mathbf{p}})U^\dagger(\mathbf{q}) - \frac{1}{2}\left\{L_j^\dagger(\mathbf{q},\hat{\mathbf{p}})L_j(\mathbf{q},\hat{\mathbf{p}}),\,\rho\right\}\right]d\mu(\mathbf{q})\\ &+\int_{\mathbb{R}^3}\sum_{j=1}^{\infty}\Big[\omega_j(\mathbf{q})(U(\mathbf{q})\rho L_j^\dagger(\mathbf{q},\hat{\mathbf{p}})U^\dagger(\mathbf{q}) - \rho L_j^\dagger(\mathbf{q},\hat{\mathbf{p}}))\\ &+(U(\mathbf{q})L_j(\mathbf{q},\hat{\mathbf{p}})\rho U^\dagger(\mathbf{q}) - L_j(\mathbf{q},\hat{\mathbf{p}})\rho)\omega_j^*(\mathbf{q})\Big]d\mu(\mathbf{q})\\ &+\int_{\mathbb{R}^3}\sum_{j=1}^{\infty}\left[U(\mathbf{q})\rho U^\dagger(\mathbf{q}) - \rho - \frac{\mathrm{i}}{\hbar}\frac{[\mathbf{q}\cdot\hat{\mathbf{x}},\,\rho]}{1+|\mathbf{q}|^2}\right]|\omega_j(\mathbf{q})|^2 d\mu(\mathbf{q})\end{aligned} \tag{27}$$

respectively, where

$$\hat{y}_k = \sum_{i=1}^{3} a_{ki}\hat{x}_i, \qquad k = 0, \ldots, r \leq 3, \qquad a_{ki} \in \mathbb{R}$$

are linear combinations of the three position operators $\hat{x}_i$, $L_k(\hat{\mathbf{p}})$ and $L_j(\mathbf{q}, \hat{\mathbf{p}})$ and are generally complex functions of the momentum operators,

$$U(\mathbf{q}) = e^{i\mathbf{q}\cdot\hat{\mathbf{x}}}$$

are unitary operators corresponding to a translation in momentum or boost, satisfying together with (24) the Weyl form of the canonical commutation rules

$$U(\mathbf{a})U(\mathbf{q}) = e^{i\mathbf{q}\cdot\mathbf{a}}U(\mathbf{q})U(\mathbf{a}), \tag{28}$$

$\omega_j(\mathbf{q})$ complex functions and $\mu(\mathbf{q})$ a positive measure on $\mathbb{R}^3$ satisfying the Lévy condition

$$\int_{\mathbb{R}^3} \frac{|\mathbf{q}|^2}{1+|\mathbf{q}|^2} \sum_{j=1}^{\infty} |\omega_j(\mathbf{q})|^2 d\mu(\mathbf{q}) < +\infty.$$

As it appears that this is quite a rich structure, allowing for the description of a very broad class of physical phenomena, only having in common invariance under translations. A first application has already been mentioned above when considering the master equation for quantum Brownian motion, describing the motion of a quantum test particle in a gas, close to thermal equilibrium, which corresponds to a particular realization of the Gaussian component (26). Further examples related to more recent research work will be considered in the next section.

3.5 Translation-Covariant Mappings for the Description of Dissipation and Decoherence

In the previous section we have given the general expression of a translation-covariant generator of a quantum-dynamical semigroup according to Holevo's results. This operator expression can be actually distinguished in to two parts, as also happens in the classical Lévy-Khintchine formula, providing the general characterization of the exponent of the characteristic function of a classical Lévy process (see e.g., [9] for a concise presentation from a physicist's standpoint or [39] for a more thorough probabilistic treatment). Very roughly speaking in the presence of translation-covariance the dynamics can be described essentially in terms of momentum exchanges between test particle and reservoir. The Gaussian part corresponds to a situation in which the dynamics is determined by a very large number of very small momentum transfers, which in the case of finite variance leads to a Gaussian process. For example, in the case of quantum Brownian motion the test particle is close to thermal equilibrium, so that typical values of its momentum are

much bigger than that of the gas particles due to its bigger mass, and the momentum changes imparted in single collisions are therefore relatively small. In contrast, the Poisson part can account for a situation in which few interaction events corresponding to significant momentum transfers drive the dynamics, as happens for example in experiments on collisional decoherence, where very few kicks already lead to a significant loss of coherence. In addition to this the general expression can also account for peculiar situations, typical of Lévy processes, where the variance or the mean of the momentum transfers do diverge, so that rare but extremely strong events can give the predominant contributions to the dynamics. We will now try to exemplify such situations referring to recent research work, thus showing how paying attention to covariance properties of dynamical mappings does not only lead to a better and deeper understanding of known results as for the case of damped harmonic oscillator, two-level system and quantum Brownian motion briefly considered in Sects. 3.2, 3.3 and 3.4 respectively, but also provides important insights into the treatment of more complicated problems, allowing for a unified description of apparently quite different situations.

Dissipation and quantum linear Boltzmann equation

The well-known quantum Brownian motion master equation (22) provides, as we have seen, an example of a realization of the Gaussian component (26) of the general structure of generators of translation-covariant quantum-dynamical semigroups specified above. A further example involving the Poisson component (27) can be given, still considering the reduced dynamics of the center of mass of a test particle interacting through collisions with a gas, however, not focussing on the case of a very massive particle close to thermal equilibrium, so that momentum transfers and therefore energy transfers due to collision events between test particle and gas are not necessarily small anymore. This kinetic stage of dynamical description asks for a quantum version of the classical linear Boltzmann equation, the equation being linear in the sense that the gas is supposed to be and remain in equilibrium, so that only the state of the test particle evolves in time. Such a master equation has been obtained recently and its expression in the case in which the scattering cross-section describing the collisions between test particle and gas only depends on the transferred momentum $\mathbf{q}$, which is always true in Born approximation, is given by [33, 40–42]

$$\begin{aligned}\frac{d\rho}{dt} = &-\frac{\mathrm{i}}{\hbar}[H_0, \rho] + \frac{2\pi}{\hbar}(2\pi\hbar)^3 n \int d^3\mathbf{q}\, |\tilde{t}(\mathbf{q})|^2 \times \Big[e^{\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}} \\ &\sqrt{S(\mathbf{q}, E(\mathbf{q}, \hat{\mathbf{p}}))}\rho\sqrt{S(\mathbf{q}, E(\mathbf{q}, \hat{\mathbf{p}}))}e^{-\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}} - \frac{1}{2}\{S(\mathbf{q}, E(\mathbf{q}, \hat{\mathbf{p}})), \rho\}\Big],\end{aligned} \tag{29}$$

where n is the gas density, $\tilde{t}(q)$ is the Fourier transform of the interaction potential, S a two-point correlation function known as dynamic structure factor, depending on momentum transfer $\mathbf{q}$ and energy transfer $E(\mathbf{q}, \mathbf{p})$:

$$E(\mathbf{q}, \mathbf{p}) = \frac{(\mathbf{p}+\mathbf{q})^2}{2M} - \frac{\mathbf{p}^2}{2M}.$$

The dynamic structure factor for a free gas of particles obeying Maxwell–Boltzmann statistics has the explicit expression:

$$S_{\mathrm{MB}}(\mathbf{q}, E) = \sqrt{\frac{\beta m}{2\pi}} \frac{1}{q} e^{-\frac{\beta}{8m}\frac{(2mE+q^2)^2}{q^2}},$$

while the general definition of dynamic structure factor reads

$$S(\mathbf{q}, E) = \frac{1}{2\pi\hbar} \int dt \int d^3\mathbf{x}\, e^{\frac{i}{\hbar}(Et-\mathbf{q}\cdot\mathbf{x})} \frac{1}{N} \int d^3\mathbf{y}\ \langle N(\mathbf{y}) N(\mathbf{x}+\mathbf{y}, t)\rangle\,,$$

that is to say it is the Fourier transform with respect to momentum and energy transfer of the two-point density–density correlation function of the medium, $N(\mathbf{y})$ being the number density operator of the gas. This indirectly tells us that the dynamics of the test particle is indeed driven by the density fluctuations in the fluid, due to its discrete microscopic nature [43]. In the structure of the master equation the unitary operators $e^{\frac{i}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}}$ and $e^{-\frac{i}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}}$ appearing to the left and the right of the statistical operator do account for the momentum transfer imparted to the test particle as a consequence of a certain collision; the rate with which collisions characterized by a certain momentum transfer happen however, do depend on the actual momentum of the test particle described by the operator $\hat{\mathbf{p}}$ through the dependence of the dynamic structure factor S on $\mathbf{p}$. This mechanism accounts for the approach to equilibrium, favoring collisions driving the kinetic energy of the test particle towards the equipartition value. The result can also be generalized to an arbitrary scattering cross section, not necessarily only depending on momentum transfer, by introducing in the master equation instead of the scattering cross section an operator-valued scattering amplitude, averaged over the gas particles momenta [44]. One can indeed check that the quantum linear Boltzmann equation (29) and its generalizations, apart from being in Lindblad form and manifestly covariant, do admit the correct stationary state (23) and drive the kinetic energy to its equipartition value.

Decoherence and Lévy processes

In the present section we will show how translation-covariant quantum-dynamical semigroups can provide a unified theoretical description of quite different decoherence experiments. At odds with the previous section we are not interested in the dynamics of the momentum observable, decoherence due to spatial localization usually takes place on a much shorter timescale than relaxation phenomena. We therefore neglect in the general expressions (26) and (27) of a translation-covariant generator the dependence on the momentum operator, which we take as a classical label, characterized, e.g., by the mean momentum of the incoming test particle. The formulas (26) and (27) drastically simplify to

$$\mathcal{L}_G[\rho] = -\mathrm{i}\sum_{i=1}^{3} b_i\,[\hat{x}_i,\rho] - \frac{1}{2}\sum_{i,j=1}^{3} D_{ij}\left[\hat{x}_i,\left[\hat{x}_j,\rho\right]\right], \tag{30}$$

$$\mathcal{L}_P[\rho] = \int d\mu(\mathbf{q})|\lambda(\mathbf{q})|^2\left[e^{\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}}\rho e^{-\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\hat{\mathbf{x}}} - \rho - \frac{\mathrm{i}}{\hbar}\frac{[\mathbf{q}\cdot\hat{\mathbf{x}},\rho]}{1+|\mathbf{q}|^2}\right], \tag{31}$$

with

$$b_i \in \mathbb{R}, \quad D_{ij} \geq 0, \quad |\lambda(\mathbf{q})|^2 = \sum_{j=1}^{\infty}|L_j^{\dagger}(\mathbf{q}) + \omega_j(\mathbf{q})|^2,$$

whose matrix elements in the position representation simply read

$$\langle\mathbf{x}|\mathcal{L}_G[\rho] + \mathcal{L}_P[\rho]|\mathbf{y}\rangle = -\Psi(\mathbf{x}-\mathbf{y})\langle\mathbf{x}|\rho|\mathbf{y}\rangle$$

with

$$\Psi(\mathbf{x}) = \mathrm{i}\sum_{i=1}^{3} b_i x_i + \frac{1}{2}\sum_{i=1}^{3} D_{ij}x_i x_j - \int d\mu(\mathbf{q})|\lambda(\mathbf{q})|^2\left[e^{\frac{\mathrm{i}}{\hbar}\mathbf{q}\cdot\mathbf{x}} - 1 - \frac{\mathrm{i}}{\hbar}\frac{\mathbf{q}\cdot\mathbf{x}}{1+|\mathbf{q}|^2}\right], \tag{32}$$

which is exactly the general expression of the characteristic exponent of a classical Lévy process [9, 39]. Neglecting the free contribution the equation for the time evolution of the statistical operator in the position representation, which now becomes

$$\frac{d}{dt}\langle\mathbf{x}|\rho|\mathbf{y}\rangle = -\Psi(\mathbf{x}-\mathbf{y})\langle\mathbf{x}|\rho|\mathbf{y}\rangle,$$

has solution

$$\langle\mathbf{x}|\rho_t|\mathbf{y}\rangle = e^{-t\Psi(\mathbf{x}-\mathbf{y})}\langle\mathbf{x}|\rho_0|\mathbf{y}\rangle \equiv \Phi(t,\mathbf{x}-\mathbf{y})\langle\mathbf{x}|\rho_0|\mathbf{y}\rangle, \tag{33}$$

given by multiplying the matrix elements of the initial statistical operator by the characteristic function

$$\Phi(t,\mathbf{x}) = e^{-t\Psi(\mathbf{x})}$$

of a classical Lévy process evaluated at a point $(\mathbf{x}-\mathbf{y})$ given by the difference between the two spatial locations characterizing bra and ket with which the matrix elements of ρ_t are taken. In view of the general properties of the characteristic function Φ, which is the Fourier transform of a probability density, that is to say

$$\Phi(t,0)=1, \qquad |\Phi(t,\mathbf{x}-\mathbf{y})|\leq 1, \tag{34}$$
$$\Phi(t,\mathbf{x}-\mathbf{y}) \xrightarrow{(\mathbf{x}-\mathbf{y})\to\infty} 0, \qquad \Phi(t,\mathbf{x}-\mathbf{y}) \xrightarrow{t\to\infty} 0,$$

the solution (33) actually predicts on general grounds an exponential loss of coherence in position, that is to say diagonalization in the localization basis. In fact according to (34) diagonal matrix elements are left untouched by the dynamics, which together with the fact that a characteristic function is actually a positive definite function accounts for the correct probability and positivity preserving time evolution. For growing time off-diagonal matrix elements are fully suppressed, whatever the distance, while for fixed time evolution the reduction of off-diagonal matrix elements depends on the relative distance $(\mathbf{x}-\mathbf{y})$, leading to a vanishing contribution for macroscopic distances (provided the corresponding classical Lévy process admits a proper probability density). An application of this general theoretical treatment to actual physical systems and in particular to relevant experimental situations relies on a choice of functions and parameters appearing in (32) dictated by actual physical input. This has been accomplished in [45], where this general scheme has been connected to actual experiments on decoherence, as well as theoretical predictions of decoherence effects when the reservoir inducing decoherence is a quantum chaotic system.

Acknowledgments The author would like to thank Prof. Francesco Petruccione for the invitation to hold these lectures, and the Centre for Quantum Technology in Durban for financial support. He is very grateful to Prof. L. Lanz for careful reading of the manuscript and most useful discussions on the subject of the paper. This research work was also supported by MIUR under PRIN05.

References

1. J. von Neumann: *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin 1932) 41
2. A.N. Kolmogorov: *Fundamental Concepts of Probability Theory* (Nauka, Moscow 1974) 41
3. G. Ludwig: *Foundations of Quantum Mechanics. I*, (Springer-Verlag, New York 1983); *Foundations of Quantum Mechanics. II*, (Springer-Verlag, New York 1985) 41, 42
4. A.S. Holevo: *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam 1982) 41
5. K. Kraus: *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin 1983) 41
6. R.F. Streater: J. Math. Phys. **41**(6), 3556–3603 (2000) 41
7. A.S. Holevo: *Statistical Structure of Quantum Theory* (Springer, Berlin 2001) 41, 47, 58
8. R. Alicki and M. Fannes: *Quantum Dynamical Systems* (Oxford University Press, Oxford 2001) 41, 58
9. H.-P. Breuer and F. Petruccione: *The Theory of Open Quantum Systems* (Oxford University Press, Oxford 2007) 41, 58, 63, 68, 71, 72, 75
10. F. Strocchi: *An Introduction to the Mathematical Structure of Quantum Mechanics* (World Scientific, Singapore 2005) 41
11. L. Lanz, B. Vacchini, and O. Melsheimer: J. Phys. A, Math. Theor. **40**(12), 3123–3140 (2007) 41

12. G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Roetteler, H. Weinfurter, R. Werner, and A. Zeilinger: *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments* (Springer, Berlin 2001) 42
13. E.B. Davies: *Quantum Theory of Open Systems* (Academic Press, London 1976) 47
14. P. Busch, M. Grabowski, and P. Lahti: *Operational Quantum Physics* (Springer-Verlag, Berlin 1995) 47
15. A. Barchielli: 'Continual measurements in quantum mechanics and quantum stochastic calculus'. In: *Open Quantum Systems. III, Volume 1882 of Lecture Notes in Mathematics*, pp. 207–292 (Springer, Berlin 2006) 54
16. R. Alicki and K. Lendi: *Quantum Dynamical Semigroups and Applications* (Springer, Berlin 1987) 58
17. R. Alicki: 'Invitation to quantum dynamical semigroups'. In: *Dynamical semigroups: Dissipation, Chaos, Quanta, Volume 597 of Lecture Notes in Physics*, pp. 239–264, edited by P. Garbaczewski and R. Olkiewicz (Springer-Verlag, Berlin 2002) 58
18. V. Gorini, A. Kossakowski, and E.C.G. Sudarshan: J. Math. Phys. **17**(5), 821–825 (1976) 61
19. G. Lindblad: Comm. Math. Phys. **48**(2), 119–130 (1976) 61
20. A.A. Budini: Phys. Rev. E **72**(5), 056106 (2005) 62
21. H.P. Breuer, J. Gemmer, M. Michel: Phys. Rev. E **73**(1), 016139 (2006) 62
22. A.A. Budini: Phys. Rev. A **74**(5), 053815 (2006) 62
23. H.-P. Breuer: Phys. Rev. A **75**(2), 022103 (2007) 62
24. H.-P. Breuer: 'Non-Markovian quantum dynamics and the method of correlated projection super-operators'. In: *Theoretical Foundations of Quantum Information Processing and Communication, Volume 787 of Lecture Notes in Physics* (Springer, Berlin) 62
25. B.-G. Englert and G. Morigi: 'Five lectures on dissipative master equations'.In: *Coherent Evolution in Noisy Environments*, Volume 611 of *Lecture Notes in Physics.*, pp. 55–106 (Springer, Berlin 2002) 62
26. B. Vacchini: J. Math. Phys. **43**(11), 5446–5458 (2002) 64, 71
27. A.S. Holevo: J. Funct. Anal. **131**(2), 255–278 (1995) 64
28. A.S. Holevo: Rep. Math. Phys. **33**(1-2), 95–110 (1993) 64
29. A.S. Holevo: Rep. Math. Phys. **32**(2), 211–216 (1993) 65, 71
30. A.Y. Artemev: Teoret. Mat. Fiz. **79**(3), 323–333 (1989) 67
31. M. Verri and V. Gorini: J. Math. Phys. **19**(9), 1803–1807 (1978) 68
32. A. Săndulescu and H. Scutaru: Ann. Phys. **173**(2), 277–317 (1987) 68
33. B. Vacchini: Phys. Rev. Lett. **84**(7), 1374–1377 (2000) 68, 73
34. R.F. O'Connell: Phys. Rev. Lett. **87**(2), 028901 (2001) 68
35. B. Vacchini: Phys. Rev. Lett. **87**(2), 028902 (2001) 68
36. A. Bassi, E. Ippoliti, and B. Vacchini: J. Phys. A **38**(37), 8017–8038 (2005) 69
37. A.S. Holevo: Izv. Math. **59**(2), 427–443 (1995) 71
38. A.S. Holevo: 'Covariant quantum dynamical semigroups: unbounded generators. In: *Irreversibility and Causality (Goslar, 1996)*, volume 504 of *L*ecture Notes in Physics, pp. 67–81 (Springer, Berlin, 1998) 71
39. W. Feller: *An Introduction to Probability Theory and its Applications. Vol. II.*, Second edition. (John Wiley & Sons Inc., New York 1971) 71, 72, 75
40. B. Vacchini: Phys. Rev. E **63**(6), 066115 (2001) 73
41. B. Vacchini: J. Math. Phys. **42**(9), 4291–4312 (2001) 73
42. B. Vacchini: Phys. Rev. E **66**(2), 027107 (2002) 73
43. F. Petruccione and B. Vacchini: Phys. Rev. E **71**(4), 046134 (2005) 74
44. K. Hornberger: Phys. Rev. Lett. **97**(6), 060601 (2006) 74
45. B. Vacchini: Phys. Rev. Lett. **95**(23), 230402 (2005) 76

Quantum Open Systems with Time-Dependent Control

Robert Alicki

1 Introduction

Any quantum device interacts with an environment and the resulting noise is a main obstacle in the implementation of bold ideas of quantum computation and large-scale quantum communication. The purpose of these notes is to present a general formalism based on the first principles of Hamiltonian models which can be used to describe mathematically and estimate numerically the influence of an environmental noise on quantum devices controlled by time-dependent external forces. We call them controlled quantum open systems (CQOS). The main idea is to eliminate the degrees of freedom of the environment and use approximative expressions for the reduced dynamics of the quantum system (device)[1]. The main problem is related to the existence of multiple timescales in such models which make the reduced dynamics very complicated [2–4]. Therefore, we restrict ourselves to several regimes of the parameters which allow the tractable description. We begin with the discussion of Markovian approximation for the case of time-independent controlling Hamiltonians which leads to the formalism of completely positive quantum dynamical semigroups with Lindblad–Gorini-Kossakowski–Sudarshan (LGKS) generators [5–9]. Spectral and ergodic properties of dynamical semigroups satisfying detailed balance condition are briefly outlined. Then, we describe the case of slowly varying external forces which also leads to inhomogeneous in time Markovian Master Equations (MME) of LGKS type and reproduces the laws of thermodynamics [10, 11]. Finally, the non-Markovian Born approximation is discussed and the corresponding formula for an error is derived [12, 13]. The last approach is illustrated by a generic mode l of the controlled spin-boson system which can be applied to several implementations of controlled qubit [14, 15].

R. Alicki (✉)
Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański, Poland
e-mail: fizra@univ.gda.pl

Alicki, R.: *Quantum Open Systems with Time-Dependent Control*. Lect. Notes Phys. **787**, 79–95 (2010)
DOI 10.1007/978-3-642-02871-7_3 © Springer-Verlag Berlin Heidelberg 2010

2 The Generic Model of CQOS

We consider a "small" open quantum system S with the corresponding Hilbert space $\mathbf{H}_S$ interacting with a "large" quantum reservoir R described by the Hilbert space $\mathbf{H}_R$. The Hilbert space of the total composed system is a tensor product $\mathbf{H}_S \otimes \mathbf{H}_R$. The Hamiltonian of the total system reads

$$H(t) = (H_S^0 + H_C(t)) \otimes I_R + I_S \otimes H_R + H_{SR}, \tag{1}$$

where H_S^0 is a *bare Hamiltonian* of S, $H_C(t)$ describes the control over S, H_R is the reservoir's Hamiltonian and H_{SR} describes the interaction between S and R which is always assumed to be weak. In the following we omit unit operators I_S, I_R. One should notice that in this picture the control is perfect up to certain fluctuations which are entirely described in terms of suitable reservoir models.

2.1 Notation and Definitions

Consider a general case of a quantum system with the Hilbert space $\mathbf{H}$ which satisfies the Schrödinger equation with time-dependent Hamiltonian $H(t)$ (we put $\hbar \equiv 1$):

$$\frac{d}{dt}\psi(t) = -\mathrm{i}H(t)\psi(t)\,. \tag{2}$$

The solution of (2) can be written in terms of the *unitary propagator* $U(t,s)$ on the Hilbert space $\mathbf{H}$ as $\psi(t) = U(t,s)\psi(s)$ and is given by the time-ordered exponential

$$\begin{aligned} U(t,s) &= \mathbf{T}\exp\Big\{-\mathrm{i}\int_s^t H(u)du\Big\} = \\ &= \sum_{n=0}^{\infty}(-\mathrm{i})^n \int_s^t dt_n \int_s^{t_n} dt_{n-1} \cdots \int_s^{t_2} dt_1 H(t_n)H(t_{n-1})\cdots H(t_1) \end{aligned} \tag{3}$$

satisfying the composition rule

$$U(t,u)U(u,s) = U(t,s)\,,\ U(t,t) = I,\ U(s,t) = U^{-1}(t,s) = U^{\dagger}(t,s)\,. \tag{4}$$

Unitary *super-propagators* on operator spaces of either trace class, bounded or Hilbert–Schmidt operators acting on $\mathbf{H}$ are defined in a similar way:

$$\mathcal{U}(t,s) = \mathbf{T}\exp\Big\{-\mathrm{i}\int_s^t \mathcal{H}(u)du\Big\},\ \mathcal{H}X = [H,X], \tag{5}$$

$$\mathcal{U}(t,s)X = U(t,s)XU^{-1}(t,s) = U(t,s)XU^{\dagger}(t,s) = U(t,s)XU(s,t), \tag{6}$$

and satisfy the composition rule $\mathcal{U}(t,u)\mathcal{U}(u,s) = \mathcal{U}(t,s)$.

The perturbation analysis of a quantum dynamics will involve the following *second-order integral equation:*

$$U(t,s) = U_0(t,s) - \mathrm{i}\int_s^t du U_0(t,u)[H(u) - H_0(u)]U_0(u,s) -$$
$$- \int_u^t du \int_s^u dv U_0(t,u)[H(u) - H_0(u)]U_0(u,v)[H(v) - H_0(v)]U(v,s), \quad (7)$$

where

$$\frac{d}{dt}U(t,s) = -\mathrm{i}H(t)U(t,s)\,,\ U(t,t) = I, \quad (8)$$
$$\frac{d}{dt}U_0(t,s) = -\mathrm{i}H_0(t)U_0(t,s)\,,\ U_0(t,t) = I, \quad (9)$$

valid for super-propagators also.

3 Reduced Dynamics

In the regime of a weak coupling to the environment we usually assume the product state structure for the initial state of the total system

$$\rho_{SR}(0) = \rho_S(0) \otimes \rho_R \quad (10)$$

with the fixed stationary state ρ_R of the reservoir and an arbitrary initial state of the system $\rho_S(0)$. The dynamics of the system S alone is obtained by the averaging over the reservoir's degrees of freedom by means of the partial trace operation $\mathrm{Tr}_R(\cdot)$. We first perform the transition to the interaction picture

$$\rho_S(t) = \Lambda(t,0)\rho_S(0) = \mathrm{Tr}_R\big(\mathcal{U}_S(0,t) \otimes \mathcal{U}_R(0,t)\mathcal{U}(t,0)\rho_S(0) \otimes \omega_R\big), \quad (11)$$

where $\mathcal{U}_S(0,t)$ is generated by the renormalized time-dependent *physical Hamiltonian* $H_S(t) = H_S^{\mathrm{corr}}(t) + H_C(t)$, with

$$H_S^{\mathrm{corr}}(t) = H_S^0 + \lambda^2 H_1^{\mathrm{corr}}(t) + \cdots \quad (12)$$

and $\mathcal{U}_R(0,t) = \exp\{i[H_R, \cdot]t\}$ and then apply the partial trace. The interaction picture eliminates short-time oscillations which are, anyway, averaged out during realistic measurement procedures. The corrections to the bare Hamiltonian producing the physical one compensate (in the given order with respect to λ) for the contribution from the interaction with an environment. Therefore, we compute always only the terms corresponding to purely dissipative effects assuming that the physical Hamiltonian is given in advance.

In the following, for the reason of clarity, we restrict ourselves to a simple interaction Hamiltonian $H_{SR} = \lambda S \otimes R$ with the standard assumption

$$\mathrm{Tr}(\rho_R R) = 0. \tag{13}$$

Here and in the following formulae, an S or S_α refers to an operator for the system while an R or R_α refers to an operator for the reservoir. The generalization to generic interactions of the form

$$H_{SR} = \lambda \sum_\alpha S_\alpha \otimes R_\alpha \tag{14}$$

is straightforward.

A useful mathematical tool for the perturbative analysis of the reduced dynamics is the *cumulant expansion*:

$$\Lambda(t,0) = \exp \sum_{n=1}^{\infty} [\lambda^n \mathcal{K}^{(n)}(t)] \,. \tag{15}$$

The first two terms $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ can be easily computed by comparison of the expansions of both sides in (15). For the LHS we use (7) and (11). From (13) it follows that $\mathcal{K}^{(1)} = 0$. In the most interesting cases the reasonable approach consists in taking the *Born approximation* $\mathcal{K}^{(2)} \equiv \mathcal{K}$:

$$\Lambda(t,0) \simeq \exp[\lambda^2 \mathcal{K}(t)]. \tag{16}$$

Here

$$\mathcal{K}(t)\rho_S = \int_0^t ds \int_0^t du \; F(s-u) S(s) \rho_S S(u) + (\text{similar terms}), \tag{17}$$

where *similar terms* are of the type $S(s)S(u)\rho_S$, $\rho_S S(s)S(u)$, and their final contribution is always computed at the end from the trace preservation condition. The reservoir's autocorrelation function F is given by the *spectral density* G:

$$F(t) = \mathrm{Tr}(\rho_R R(t) R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\omega t} G(\omega) d\omega, \tag{18}$$

where $R(t)$ evolves according to H_R and $S(t)$ according to $H_S(t)$.

The double integral structure of (17) implies the quadratic behavior $\mathcal{K}(t) \sim t^2$ for short times and hence a Gaussian decay which is a basic assumption in the theory of the *Zeno effect*.

4 Markovian Dynamics

For longer times the decay of the autocorrelation function $F(t)$ combined with the rapid oscillations present in the time-dependent operators $S(t)$ causes the superoperator $\mathcal{K}(t)$ (17) to increase linearly in time. This is the main mechanism which allows to use different types of *Markovian approximations*. In the following we shall use a rather general definition of the Markovian approximation based on the approximative relation

$$\mathcal{K}(t) \simeq \int_0^t \mathcal{L}(s)ds, \tag{19}$$

where $\mathcal{L}(s)$ is the Lindblad–Kossakowski–Gorini–Sudarshan (LKGS) generator. We recall that any LKGS generator can be written in the form

$$\mathcal{L}\rho_S = -\mathrm{i}[H, \rho_S] + \frac{1}{2}\sum_\alpha \Big([V_\alpha, \rho_S V_\alpha^\dagger] + [V_\alpha \rho_S, V_\alpha^\dagger]\Big), \tag{20}$$

and for any $\tau \geq 0$ the map $e^{\mathcal{L}\tau}$ is completely positive (CP) and trace preserving. The LKGS assumption can be justified as follows. The condition (19) implies that up to the higher order corrections the Born approximation for the cumulant expansion (16) coincides with the propagator for the following (generally inhomogeneous in time) MME in the interaction picture

$$\frac{d\rho_S(t)}{dt} = \mathcal{L}(t)\rho_S(t)\ . \tag{21}$$

Therefore, the propagator $\Lambda(t, s) = \mathbf{T}\exp\int_s^t \mathcal{L}(u)du$ can be written as the limit of a product of completely positive maps:

$$\Lambda(t, s) = \lim_{n\to\infty} e^{\mathcal{L}(u_n)(t-u_n)} \cdots e^{\mathcal{L}(u_1)(u_2-u_1)} e^{\mathcal{L}(s)(u_1-s)}\ . \tag{22}$$

For any CP map $\Lambda_k = e^{\mathcal{L}(u_k)(u_{k+1}-u_k)}$ we can construct a *unitary dilation*, i.e., a model of a reservoir with the Hilbert space $\mathbf{H}_{\mathbf{R_k}}$, the initial state ρ_{R_k} and the total unitary dynamics U_k acting on $\mathbf{H_S} \otimes \mathbf{H}_{\mathbf{R_k}}$ such that

$$\Lambda_k \rho_S = \mathrm{Tr}_{R_k}\Big(U_k \rho_S \otimes \rho_{R_k} U_k^\dagger\Big)\ . \tag{23}$$

It follows that the reduced dynamics $\Lambda(t, s)$ can be seen as an effect of the coupling to a permanently "refreshed" reservoir such that the actual state of the composed system is always a tensor product. This picture corresponds to the most general idea of a memoryless (Markovian) model of the open system. One should notice that, formally, the general choice of $\mathcal{L}(s)$, which does not satisfy the LKGS condition,

is always possible. It leads to the so-called *convolutionless approach* to reduced dynamics which, generally, has nothing to do with the Markovian modelling.

4.1 White-Noise Model

We begin with a naive Markov approximation which consists in replacing the environmental observable $R(t)$ by the *white noise*, i.e., a Gaussian stationary stochastic process with the second moment given by

$$F(t) = a\delta(t)\,,\, a > 0\,. \tag{24}$$

Then we obtain

$$\mathcal{L}(s)\rho_S = aS(s)\rho_S S(s) + (\text{similar terms}), \tag{25}$$

which after transition to the Schrödinger picture yields the following MME:

$$\frac{d\rho_S}{dt} = -\mathrm{i}[H_S(t), \rho_S] - \frac{1}{2}\lambda^2 a[S, [S, \rho_S]]\,. \tag{26}$$

The important feature of this equation is that the dissipative part is independent of the Hamiltonian one. Although the "double commutator" generators are used to describe initial stages of decoherence or continuous measurement processes they cannot be considered as generic ones. There exist serious physical objections to the naive Markovian approximation (24). First for many cases of reservoirs (e.g., electromagnetic field, phonons in solids) we have the following behavior in the domain of low frequencies:

$$G(\omega) \sim \omega^d\,,\; G(0) = 0 = \int_{-\infty}^{+\infty} F(t)dt\,, \tag{27}$$

and then the substitution (24) is inconsistent. Secondly, for heat baths the Kubo–Martin–Schwinger (KMS) condition holds

$$G(-\omega) = e^{-\beta\omega}G(\omega)\,, \quad \text{where} \quad \beta = \frac{1}{kT}, \quad k \text{ denotes Boltzmann constant}, \tag{28}$$

which for low temperatures T excludes even approximately a constant spectral density characterizing the white noise. Usually, this phenomenon is attributed to the *thermal memory time* $\hbar/kT$. Therefore, strictly speaking, the white noise approximation makes sense for $T \to \infty$ only.

5 Weak-Coupling Limit for Constant Hamiltonian

We discuss briefly the derivation of the MME for the case of time-independent controlling Hamiltonian $H_C(t) = H_C$ based on the rigorous approach of Davies in terms of the weak-coupling limit (WCL). We assume that the total physical Hamiltonian H_S possesses well-separated *Bohr frequencies*, i.e., differences of the quantised energies, which form a set $\{\omega\}$.

We use the following decomposition into Fourier components:

$$S(t) = \exp(\mathrm{i}H_S t)S\exp(-\mathrm{i}H_S t) = \sum_{\omega} S_\omega \exp(-\mathrm{i}\omega t)\ ,\ \ S_{-\omega} = S_\omega^\dagger. \tag{29}$$

The operators S_ω are eigenvectors of the super-operator $\mathcal{H}_S$:

$$\mathcal{H}_S S_\omega \equiv [H_S, S_\omega] = -\omega S_\omega\ . \tag{30}$$

Then the Born approximation (17) reads

$$\mathcal{K}(t)\rho_S = \sum_{\omega,\omega'} S_\omega \rho_S S_{\omega'}^\dagger \int_0^t e^{\mathrm{i}(\omega-\omega')u}du \int_{-u}^{t-u} F(\tau)e^{\mathrm{i}\omega\tau}d\tau + (\text{similar terms})\ . \tag{31}$$

The main idea of WCL is to use the *coarse-grained timescale* $t \to t/\lambda^2$ and formally go with the coupling λ to zero (van Hove limit). Physically, the idea of WCL corresponds to the following approximations involving different timescales in the system, and the decay properties of multitime correlation functions:

1. The integral

$$\int_0^t e^{\mathrm{i}(\omega-\omega')u}du \approx t\delta_{\omega\omega'}, \tag{32}$$

   which makes sense for $t \gg \max\{1/(\omega-\omega')\}$.
2. Limits in the integral

$$\int_{-u}^{t-u} F(\tau)e^{\mathrm{i}\omega\tau}d\tau \tag{33}$$

   are replaced by the infinities

$$\int_{-u}^{t-u} F(\tau)e^{\mathrm{i}\omega\tau}d\tau \approx G(\omega) = \int_{-\infty}^{\infty} F(\tau)e^{\mathrm{i}\omega\tau}d\tau\ . \tag{34}$$

3. Multitime correlation functions of the reservoir's variables decay sufficiently fast to justify the approximation (16).

Applying those approximations we obtain

$$\mathcal{K}(t)\rho_S = t\sum_{\omega} G(\omega)S_\omega\rho_S S_\omega^\dagger + (\text{similar terms}), \tag{35}$$

and hence an MME of the Davies–LGKS form in the Schrödinger picture

$$\frac{d\rho_S}{dt} = -\mathrm{i}[H_S, \rho_S] + \frac{1}{2}\lambda^2\sum_{\omega} G(\omega)([S_\omega, \rho_S S_\omega^\dagger] + [S_\omega\rho_S, S_\omega^\dagger]) \,. \tag{36}$$

Remark One should notice that the standard justification of the Markovian approximation based on the assumption that the *bath's correlation time* is short in comparison with the other timescales is not generally true. Quite often the decay of correlations of bath's variables F is power-like and does not possess any natural timescale. In these cases Markovian behavior is a consequence of a "cooperation" between system and bath dynamics.

For a generic interaction Hamiltonian (14) we obtain the following Davies–LKGS generator:

$$\mathcal{L}\rho_S = \frac{1}{2}\lambda^2\sum_{\alpha,\beta}\sum_{\omega} G_{\alpha\beta}(\omega)\big([S_\alpha(\omega), \rho_S S_\beta(\omega)^\dagger] + [S_\alpha(\omega)\rho_S, S_\beta(\omega)^\dagger]\big), \tag{37}$$

with the positive matrix of spectral densities $G_{\alpha\beta}(\omega)$

$$\mathrm{Tr}(\rho_R R_\alpha(t)R_\beta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}\omega t}G_{\alpha\beta}(\omega)d\omega \,. \tag{38}$$

It is often very useful to use the Heisenberg picture of the time evolution in terms of time-dependent observables such that for any observable A_S:

$$\mathrm{Tr}(\rho_S(t)A_S) = \mathrm{Tr}(\rho_S A_S(t)).$$

For the MME with the generator (37) its Heisenberg version reads

$$\frac{dA_S}{dt} = \mathrm{i}\mathcal{H}_S A_S + \mathcal{L}_* A_S \,,\ \mathcal{H}_S A_S = [H_S, A_S], \tag{39}$$

where

$$\mathcal{L}_* A_S = \frac{1}{2}\lambda^2\sum_{\alpha,\beta}\sum_{\omega} G_{\alpha\beta}(\omega)\Big(S_\beta(\omega)^\dagger[A_S, S_\alpha(\omega)] + [S_\beta(\omega)^\dagger, A_S]S_\alpha(\omega)\Big) \,. \tag{40}$$

5.1 Properties of MME for a Heat Bath

Assume that our system is coupled to a heat bath, i.e., a single reservoir in the thermodynamic equilibrium state at the temperature T. Using the form (40) of the Heisenberg generator with the KMS condition (28) one can prove by direct calculation the following properties:

1. The stationary state of MME in the Schrödinger picture is a Gibbs state at the temperature $T(\mathcal{L}\rho_T = 0)$:

$$\rho_T = \frac{e^{-H_S/kT}}{\mathrm{Tr}e^{-H_S/kT}} \,. \tag{41}$$

2. The Heisenberg generator is a normal (diagonalizable) operator on the Hilbert space L^2_T of operators with the scalar product

$$\langle A, B\rangle_T = \mathrm{Tr}(\rho_T A^\dagger B) \tag{42}$$

such that $\mathcal{H}$ commutes with $\mathcal{L}_*$ and the *quantum detailed balance* holds

$$\langle \mathcal{H}_S A, B\rangle_T = \langle A, \mathcal{H}_S B\rangle_T,\ \langle \mathcal{L}_* A, B\rangle_T = \langle A, \mathcal{L}_* B\rangle_T \,. \tag{43}$$

3. Dissipativity condition is satisfied, i.e.,

$$\mathcal{L}_* \le 0 \tag{44}$$

as an operator on L^2_T.

4. Spectral decomposition holds ($\langle X_\mu, X_{\mu'}\rangle_T = \delta_{\mu\mu'}$):

$$A_S(t) = \sum_{\{\mu\}} e^{-\mu t}\langle X_\mu, A_S\rangle_T X_\mu \,, \mathrm{Im}\,\mu = \omega,\ \mathrm{Re}\,\mu \ge 0 \,. \tag{45}$$

5. Under the conditions $[G_{\alpha\beta}(\omega)] > 0$ and

$$\{A; [H_S, A] = [S_\alpha(\omega), A] = 0, \text{for all } \omega\} = \mathbf{C}I,$$

the dynamical semigroup is exponentially relaxing, i.e.,

in the Schrödinger picture

$$\lim_{t\to\infty} \exp\big\{(-\mathrm{i}\mathcal{H} + \mathcal{L})t\big\}\rho_S = \rho_T \,, \tag{46}$$

and in the Heisenberg picture

$$\lim_{t\to\infty} \exp\big\{(\mathrm{i}\mathcal{H} + \mathcal{L}_*)t\big\}A_S = \mathrm{Tr}(\rho_T A_S)I \,. \tag{47}$$

6 Adiabatic Control in the WCL

From the discussion of the approximations involved in the WCL derivation of the MME (36) it follows that this method does not work for a general $H_S(t)$. However, there exists an *adiabatic window* described by the following conditions concerning the relevant timescales:

1.

$$\text{typical}\left\{\frac{1}{\omega-\omega'}\right\} \ll \text{variation time of } H_S(t), \tag{48}$$

2.

$$\text{variation time of } H_S(t) \ll \text{relaxation time of } S\ . \tag{49}$$

Under the conditions (48) and (49) one can use the MME with the time-dependent generator

$$\frac{d\rho_S(t)}{dt} = -\mathrm{i}[H_S(t),\, \rho_S(t)] + \mathcal{L}(t)\rho_S(t), \tag{50}$$

where $\mathcal{L}(t)$ is obtained by the WCL procedure (37) with respect to $H_S(t)$.

6.1 A System Coupled to a Few Heat Baths

We consider now a quantum system controlled by slowly varying external fields such that the conditions (48) and (49) are fulfilled. We assume also that its environment can be represented by several independent heat baths at different temperatures, in general. Then the corresponding time-dependent generator satisfies

$$\mathcal{L}(t) = \sum_m \mathcal{L}_m(t)\ ,\ \ \mathcal{L}_m(t)\rho_{T_m}(t) = 0, \tag{51}$$

with

$$\rho_{T_m}(t) = \frac{e^{-H_S(t)/kT_m}}{\mathrm{Tr}e^{-H_S(t)/kT_m}}\ . \tag{52}$$

6.2 The Laws of Thermodynamics

The model described by (51) and (52) can be used to derive the laws of phenomenological thermodynamics. First, we define the basic quantities—energy (E), heat (Q), work (W):

$$E(t) = \mathrm{Tr}(\rho_S(t)H_S(t)), \tag{53}$$

$$\begin{aligned}\frac{dQ(t)}{dt} &= \mathrm{Tr}\Big\{\frac{d\rho_S(t)}{dt}H_S(t)\Big\},\\ \frac{dW(t)}{dt} &= \mathrm{Tr}\Big\{\rho_S(t)\frac{dH_S(t)}{dt}\Big\},\end{aligned} \tag{54}$$

and entropy

$$S(t) \equiv S(\rho_S(t)) = -k\mathrm{Tr}\big(\rho_S(t)\log\rho_S(t)\big). \tag{55}$$

We compute the time derivative of the entropy

$$\begin{aligned}\frac{dS}{dt} &= -k\sum_m \mathrm{Tr}\big([\mathcal{L}_m\rho_S]\log\rho_S\big) =\\ &= -k\sum_m \mathrm{Tr}\big([\mathcal{L}_m\rho_S][\log\rho_S - \log\rho_{T_m}]\big) + \sum_m \frac{1}{T_m}\frac{dQ_m}{dt},\end{aligned} \tag{56}$$

where

$$\frac{dQ}{dt} = \mathrm{Tr}\Big\{\frac{d\rho_S}{dt}H_S\Big\} = \sum_m \mathrm{Tr}\big(H_S\mathcal{L}_m\rho_S\big) \equiv \sum_m \frac{dQ_m}{dt} \tag{57}$$

and Q_m is a heat supplied by the m-th heat bath.

We define the relative entropy for a pair of density matrices ρ, σ as

$$S(\rho|\sigma) = \mathrm{Tr}(\rho\log\rho - \rho\log\sigma) \geq 0\,. \tag{58}$$

For any CP trace preserving map Λ the following relation holds:

$$S(\Lambda\rho|\Lambda\sigma) \leq S(\rho|\sigma)\,. \tag{59}$$

Take $\Lambda = e^{\tau\mathcal{L}_m}$ and $\sigma = \rho_{T_m}$ then the following inequality holds:

$$0 \geq \frac{d}{d\tau}S(e^{\tau\mathcal{L}_m}\rho|e^{\tau\mathcal{L}_m}\rho_{T_m})|_{\tau=0} = \mathrm{Tr}\big([\mathcal{L}_m\rho_S][\log\rho_S - \log\rho_{T_m}]\big)\,. \tag{60}$$

Now using Eqs. (53), (54), (56), and (60) we obtain:

1. First law of thermodynamics:

$$\frac{dE}{dt} = \frac{dQ}{dt} + \frac{dW}{dt}. \tag{61}$$

2. Second law of thermodynamics:

$$\frac{dS}{dt} = \kappa + \sum_m \frac{1}{T_m}\frac{dQ_m}{dt}, \tag{62}$$

where $\kappa \geq 0$ is the entropy production.

7 CQOP in Non-Markovian–Born Regime

The aim of this section is to derive the expression for an error in the final state of the controlled quantum device due to the generic weak interaction with an environment (14). Comparing the actual dynamics $\Lambda(t,s)$ with the ideal one $\mathcal{U}_S(t,s)$ acting on the initial pure state ψ we define the error accumulated between the initial time $t_{\rm in} = -\tau$ and the final one $t_{\rm fin} = \tau$ as

$$\epsilon = 1 - \langle U_S(\tau,-\tau)\psi, \Lambda(\tau,-\tau)(|\psi><\psi|)U_S(\tau,-\tau)\psi\rangle\ . \tag{63}$$

Using the formulas for the propagators (3) and the second-order formula (7) we obtain the dynamical map Γ in the Schrödinger picture describing the evolution of S from $t_{\rm in} = -\tau$ to $t_{\rm fin} = \tau$:

$$\Gamma\rho_S = \mathcal{U}_S(\tau,-\tau)\rho_S - \mathrm{Tr}_R\Big\{\int_{-\tau}^{\tau} ds \int_{-\tau}^{s} du\, \mathcal{U}_0(\tau,s)\,\mathcal{H}_{SR}\,\mathcal{U}_0(s,u)\,\mathcal{H}_{SR}\,\mathcal{U}(u,-\tau)\rho_S\otimes\rho_R\Big\}. \tag{64}$$

The Born approximation consists in the replacement

$$\mathcal{U}(u,-\tau)\rho_S\otimes\rho_R \to \mathcal{U}_S(u,-\tau)\rho_S\otimes\rho_R\ .$$

$\Gamma\rho_S$ in the Born approximation possesses "almost CP structure" (CP up to the higher orders with respect to the interaction)

$$\Gamma\rho = \mathcal{U}_S\Big(\rho_S + \Phi(\rho_S) - \frac{1}{2}\{B,\rho_S\}\Big), \tag{65}$$

where $\mathcal{U}_S \equiv \mathcal{U}_S(\tau,-\tau)$. The CP map Φ is called *error map*:

$$\Phi\rho_S = \int_{-\tau}^{\tau} ds \int_{-\tau}^{\tau} du \mathrm{Tr}\big(\rho_R R_\alpha R_\beta(s-u)\big) S_\beta(s,-\tau)\rho_S S_\alpha(u,-\tau), \tag{66}$$

where $S_\alpha(s,u) = \mathcal{U}_S(u,s)S_\alpha$, $R_\alpha(t) = e^{{\rm i}t\mathcal{H}_R}R_\alpha$.

The operator B, as mentioned before, can be computed at the end using the trace preservation conditions which leads to the expression $B = \Phi_* I \geq 0$.

It is convenient to introduce the *Fourier picture* with spectral densities ($R_{\alpha\beta}(\omega) \equiv \frac{1}{2\pi}G_{\alpha\beta}(\omega)$ of Eq. (38))

$$\mathrm{Tr}\big(\rho_R R_\alpha R_\beta(t)\big) = \int_{-\infty}^{\infty} R_{\alpha\beta}(\omega) e^{-\mathrm{i}\omega t} d\omega \tag{67}$$

and defining

$$Y_\alpha(\omega) = \int_{-\tau}^{\tau} S_\alpha(s, -\tau) e^{-\mathrm{i}\omega s} ds \tag{68}$$

to obtain

$$\Phi\rho_S = \sum_{\alpha,\beta} \int_{-\infty}^{\infty} d\omega\, R_{\alpha\beta}(\omega)\, Y_\beta(\omega) \rho_S Y_\alpha^\dagger(\omega)\,. \tag{69}$$

The relation (69) can be seen as a particular example of the *quantum fluctuation–dissipation theorem*. Inserting Eqs. (69) and (65) into (63) we obtain the error formula in the Born approximation in terms of the Fourier transforms:

$$\begin{aligned}\epsilon &= \langle\psi, B\psi\rangle - \langle\psi, \Phi(|\psi\rangle\langle\psi|)\psi\rangle \\ &= \sum_{\alpha,\beta} \int_{-\infty}^{\infty} d\omega\, R_{\alpha\beta}(\omega) \Big(\langle\psi, Y_\alpha^\dagger(\omega) Y_\beta(\omega)\psi\rangle - -\langle\psi, Y_\alpha^\dagger(\omega)\psi\rangle\langle\psi, Y_\beta(\omega)\psi\rangle \Big)\,.\end{aligned} \tag{70}$$

It is convenient to define

$$\langle\psi, Y_\alpha^\dagger(\omega) Y_\beta(\omega)\psi\rangle - \langle\psi, Y_\alpha^\dagger(\omega)\psi\rangle\langle\psi, Y_\beta(\omega)\psi\rangle = 2\tau S_{\alpha\beta}(\omega)\,. \tag{71}$$

In order to find the meaning of $S_{\alpha\beta}(\omega)$ we introduce for any two random variables (classical or quantum) their *correlator* $\mathcal{F}(\omega) = [F_{\alpha\beta}(\omega)]$ as

$$\begin{aligned}\lim_{\tau\to\infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} ds\, \Big(\langle f_\alpha(t+s) f_\beta(s)\rangle_{av} - \langle f_\alpha(t+s)\rangle_{av} \langle f_\beta(s)\rangle_{av} \Big) \\ = \int_{-\infty}^{\infty} F_{\alpha\beta}(\omega) e^{\mathrm{i}\omega t} d\omega\,.\end{aligned} \tag{72}$$

Putting $f_\alpha \equiv R_\alpha$ and $\langle\cdot\rangle_{av} \equiv \mathrm{Tr}(\omega_R \cdot)$ we obtain $\mathcal{F}(\omega) \equiv \mathcal{R}(\omega)$.

Similarly, $\mathcal{F}(\omega) \equiv \mathcal{S}(\omega)$ for $f_\alpha \equiv S_\alpha$, $\langle\cdot\rangle_{av} \equiv \langle\psi, \cdot\psi\rangle$ only if τ is long enough. Therefore, the error formula (70) can be expressed in terms of two correlators $\mathcal{R}(\omega) = [R_{\alpha\beta}(\omega)]$ and $\mathcal{S}(\omega) = [S_{\alpha\beta}(\omega)]$ multiplied by the operation time 2τ:

$$\epsilon = 2\tau \int_{-\infty}^{\infty} d\omega\, \mathrm{Tr}\big(\mathcal{R}(\omega)\mathcal{S}(\omega)\big)\,. \tag{73}$$

Equation (73) provides a useful starting point for the analysis of different error reduction strategies.

7.1 Error Reduction Strategies

The error formula (73) has a simple physical meaning. In the small error regime the error grows linearly with time and is proportional to the overlap of two "spectral densities" corresponding to the dynamics of the system and the reservoir, respectively. This overlap is related to the energy conservation principle—the transmission of energy and information is possible between the "modes" of the system and the reservoir with equal frequencies only. Therefore, the fundamental strategies of decoherence reduction are the following:

1. *Slow gates* – applicable for environments with low frequency behavior

$$\mathcal{R}(\omega) \sim \omega^{\kappa} \ , \ \kappa > 1 \ . \tag{74}$$

   Indeed, for slow gates the support of $\mathcal{S}(\omega)$ is concentrated in the regime of low frequencies, where $\mathcal{R}(\omega)$ is small.
2. *Fast gates* – applicable for environments with flat $\mathcal{R}(\omega) \simeq \mathcal{R}$ which leads to the error proportional to the duration of the gate. If $\mathcal{R}(\omega)$ rapidly decreases above certain cut-off frequency then the fast gate strategy is even more effective because $\mathcal{S}(\omega)$ is concentrated in the regime of high frequencies.
3. *Decoherence-free subspaces (subalgebras)* – applicable for systems with a symmetric coupling to an environment. Due to a certain symmetry of the system-bath coupling there exists a subspace of ψs such that $\mathcal{S}(\omega) = 0$ for all relevant frequencies. From (70) it implies

$$Y_{\alpha}(\omega)\psi = \lambda_{\alpha}(\omega)\psi \ , \ \lambda_{\alpha}(\omega) \in \mathbf{C} \ . \tag{75}$$

   Such initial states evolve without errors in the lowest order approximation. A linear subspace spanned by those ψs is called *decoherence-free* and the typical example is given by the superradiance phenomena where the symmetry corresponds to the approximative invariance of the coupling to the electromagnetic field with respect to permutations of atoms.

8 Controlled Spin-Boson System

The simplest model of CQOP consists of a qubit, representing different systems like spin-1/2, "2-level atom," 2-configurations of a molecule, quantum dot, etc., and described by the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ as basic observables. The qubit is coupled to a reservoir modelled by the (infinite) system of harmonic oscillators (bosonic field). The bath can be treated as a gas of noninteracting bosons (photons, phonons, etc.) described by the family of annihilation and creation operators satisfying canonical commutation relations:

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = [a^{\dagger}_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}'}] = 0 \ , \ [a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'} \ . \tag{76}$$

Here $\mathbf{k}$ is typically the wave vector of the given mode and $\omega_{\mathbf{k}}$ its frequency.

We restrict ourselves to the following Hamiltonian describing, for example, an optically controlled exciton in a quantum dot coupled to acoustical phonons:

$$H''(t) = \frac{1}{2}\Omega\sigma_z + \frac{1}{2}E(t)\big(e^{\mathrm{i}\Omega t}\sigma_+ + e^{-\mathrm{i}\Omega t}\sigma_-\big) + \sum_{\mathbf{k}} \omega_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sigma_z \otimes \Big[\sum_{\mathbf{k}} f_{\mathbf{k}}^* a_{\mathbf{k}} + f_{\mathbf{k}} a_{\mathbf{k}}^\dagger\Big]. \tag{77}$$

Here $E(t)$ is the envelope of the laser pulse used to perform gates. The state $|0\rangle$ of the qubit represents the absence of an electron in a dot and the state $|1\rangle$ its presence.

In the first step we eliminate oscillations with the frequency Ω by the transition to the interaction picture Hamiltonian (with respect to $\frac{1}{2}\Omega\sigma_z$):

$$H'(t) = E(t)\sigma_x + H_R + H_{SR}\,. \tag{78}$$

The coupling to the environment cannot be treated as a weak one. The electromagnetic interaction of an electron with a lattice strongly modifies the local state of the later. Therefore, before applying any "weak coupling" type of approximation one should find a new decomposition of the total system into S and R introducing the notion of a dressed system (*dressed qubit*). Define *Weyl unitary operators* $W(g)$ acting on the Hilbert space $\mathbf{F}$ of the harmonic oscillators bath as

$$g \equiv \{g_{\mathbf{k}}\}\,,\ W(g) = \exp\sum_{\mathbf{k}}\big\{g_{\mathbf{k}}^* a_{\mathbf{k}} - g_{\mathbf{k}} a_{\mathbf{k}}^\dagger\big\} \tag{79}$$

with the properties

$$W(-g) = W(g)^\dagger\,,\ W(g)W(h) = e^{\frac{\mathrm{i}}{2}\mathrm{Im}\langle g,h\rangle}W(g+h), \tag{80}$$

$$W(g)a_{\mathbf{k}}W^\dagger(g) = a_{\mathbf{k}} + g_{\mathbf{k}}\,,\ W(g)a_{\mathbf{k}}^\dagger W^\dagger(g) = a_{\mathbf{k}}^\dagger + g_{\mathbf{k}}^*\,. \tag{81}$$

Then we introduce the *dressing unitary transformation* acting on the total Hilbert space $\mathbf{C}^2 \otimes \mathbf{F}$:

$$\mathbf{W} = |0\rangle\langle 0| \otimes W(f/\omega) - |1\rangle\langle 1| \otimes W^\dagger(f/\omega), \tag{82}$$

where $f/\omega = \{f_{\mathbf{k}}/\omega_{\mathbf{k}}\}$.

The *dressed Hamiltonian* $H(t) = \mathbf{W}H'(t)\mathbf{W}^\dagger$ can be easily computed:

$$H(t) = \frac{1}{2}E(t)\sigma_x \otimes \cos\Big\{2\mathrm{i}\sum_{\mathbf{k}}\Big(\frac{f_{\mathbf{k}}}{\omega_{\mathbf{k}}}a_{\mathbf{k}}^\dagger - \frac{f_{\mathbf{k}}^*}{\omega_{\mathbf{k}}}a_{\mathbf{k}}\Big)\Big\} + \frac{1}{2}E(t)\sigma_y \otimes \sin\Big\{2\mathrm{i}\sum_{\mathbf{k}}\Big(\frac{f_{\mathbf{k}}}{\omega_{\mathbf{k}}}a_{\mathbf{k}}^\dagger - \frac{f_{\mathbf{k}}^*}{\omega_{\mathbf{k}}}a_{\mathbf{k}}\Big)\Big\} + H_R\,. \tag{83}$$

One should notice that the interaction between the dressed qubit and the remaining degrees of freedom is present during gate operation only!

Now the weak-coupling approximation can be used and we restrict ourselves to the first order terms with respect to f/ω:

$$H(t) \simeq E(t)\sigma_x + H_R + E(t)\sigma_y \otimes \mathrm{i}\sum_{\mathbf{k}}\Big(\frac{f_{\mathbf{k}}}{\omega_{\mathbf{k}}}a^{\dagger}_{\mathbf{k}} - \frac{f^{*}_{\mathbf{k}}}{\omega_{\mathbf{k}}}a_{\mathbf{k}}\Big)\,. \tag{84}$$

8.1 Remarks

1. The dynamics governed by (83) or (84) is a strongly non-Markovian reduced dynamics for the dressed qubit and hence produces non-Lorentzian spectrum (observed in experiments).
2. Electromagnetic bath, not included in our model, leads to an independent, essentially Markovian recombination process with a relatively long time τ_M. Therefore, we expect here a certain trade-off between fast and slow gate strategies of error reduction.

8.2 Error for a Single Bit Flip

The formulas for the error in Born approximation derived in the previous section can be applied to compute an averaged error of a single bit flip gate in our model. We assume a Gaussian pulse of the duration τ_g corresponding to a bit flip:

$$E(t) = \frac{\sqrt{\pi}}{\sqrt{2}\tau_g}e^{-\frac{1}{2}(t/\tau_g)^2}\,. \tag{85}$$

The zero-temperature spectral density of the acoustical phonons behaves like

$$R(\omega) = R_0\omega^3\,,\ \omega \in [0,\omega_D]\,,\ \ R(\omega) = 0 \text{ otherwise}\,, \tag{86}$$

where ω_D denotes the Debye frequency. The straightforward estimation based on the formula (73) and including the Markovian recombination process gives

$$\epsilon \simeq \frac{\pi^2}{3}R_0\frac{1}{\tau_g^2} + \frac{\tau_g}{\tau_M} \tag{87}$$

with a trade-off between short and long τ_g . More realistic behavior of $\epsilon(\tau_g)$ involving experimental data for $R(\omega)$, τ_M, finite temperature and the cut-off parameters (quantum dot size) can be obtained numerically.

Acknowledgment This work is supported by the POLAND/SA COLLABORATION PROGRAMME of the National Research Foundation of South Africa and the Polish Ministry of Science and Higher Education.

References

1. H.-P. Breuer and F. Petruccione: *The Theory of Open Quantum Systems* (Oxford University Press Oxford, 2002) 79
2. E.B. Davies and H. Spohn: J. Stat. Phys. **19**, 511 (1978) 79
3. R. Alicki: 'Invitation to Quantum Dynamical Semigroups'. In: *Dynamics of Dissipation*, ed. by P. Garbaczewski and R. Olkiewicz (LNP 597 Springer-Verlag, Berlin, 2002) 79
4. R. Alicki, D.A. Lidar, and P. Zanardi: Phys. Rev. A **73**, 052311 (2006) 79
5. G. Lindblad: Commun. Math. Phys. **48**, 119 (1976) 79
6. V. Gorini, A. Kossakowski and E.C.G. Sudarshan: J. Math. Phys. **17**, 821 (1976) 79
7. E.B. Davies: Commun. Math. Phys. **39**, 91 (1974) 79
8. R. Alicki and K. Lendi: *Quantum Dynamical Semigroups and Applications*, 2nd edn. (LNP 717 Springer-Verlag, Berlin, 2007) 79
9. R. Alicki and M. Fannes: *Quantum Dynamical Systems* (Oxford University Press, Oxford, 2001) 79
10. J.L. Lebowitz and H. Spohn: Adv. Chem. Phys. **38**, 109 (1978) 79
11. R. Alicki: J. Phys. A **12**, L103 (1979) 79
12. R. Alicki: 'Controlled Quantum Open Systems'. In: *Irreversible Quantum Dynamics*, ed. by F. Benatti and R. Floreanini (LNP 622, Springer Verlag Berlin, 2003) 79
13. R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki: Phys. Rev. A **65**, 062101 (2002) 79
14. R. Alicki, M. Horodecki, P. Horodecki, R. Horodecki, L. Jacak, and P. Machnikowski: Phys. Rev. A **65**, 062101 (2002) 79
15. R. Alicki: Open Sys. Inf. Dyn. **11**, 53 (2004) 79

Five Lectures on Quantum Information Applications of Complex Many-Body Systems

Sougato Bose

1 Introduction

In recent years it has been realized that utilizing the full potential of quantum mechanics enables one to perform certain particularly important computational tasks much faster than ever possible with conventional (classical) computing technology [1]. In addition, quantum mechanics enables completely novel forms of cryptograpy and communication [2]. For the above tasks it is essential to encode information in quantum mechanical versions of bits or "qubits" (quantum mechanical two-state systems), typical examples being the polarization of a photon or the spin of an electron. For quantum computation in particular, very large collections of qubits are required. In general, larger the collection of qubits, more powerful the quantum computer. While few qubit quantum processors are already available, increasing the number of qubits by significant amounts is a problem that needs to be solved before truly powerful quantum computers come into being. It is from this point of view that the current set of lectures can be motivated. Complex many-body systems, a typical example being magnets, are each natural collections of several continuously interacting qubits or other identical systems. Can this *natural* resource of qubits be exploited for quantum computation or at least for tasks allied to quantum computation such as building small automata for logic gates or even simply for constructing a connection bus between two quantum computers? These are some of the questions that we will try to address in our lectures.

The topic of these lectures can be motivated in yet another way. An ideal quantum computer itself is the ultimate example of an engineered and controllable many-body quantum system. An example is provided in Fig. 1(a), which shows a designed system of regular structure, each of whose elements (which are qubits) can be individually controlled and measured. Additionally, one must be able to switch on and off the interaction between any two qubits of a quantum computer controllably at any time. So one can ask what happens if some of this control is relaxed? Suppose the interactions are frozen or permanent or can be switched on only all at once?

S. Bose (✉)
Department of Physics and Astronomy, University College London, UK
e-mail: theory.phys.ucl.ac.uk

Bose, S.: *Five Lectures on Quantum Information Applications of Complex Many-Body Systems*. Lect. Notes Phys. **787**, 97–124 (2010)
DOI 10.1007/978-3-642-02871-7_4 © Springer-Verlag Berlin Heidelberg 2010

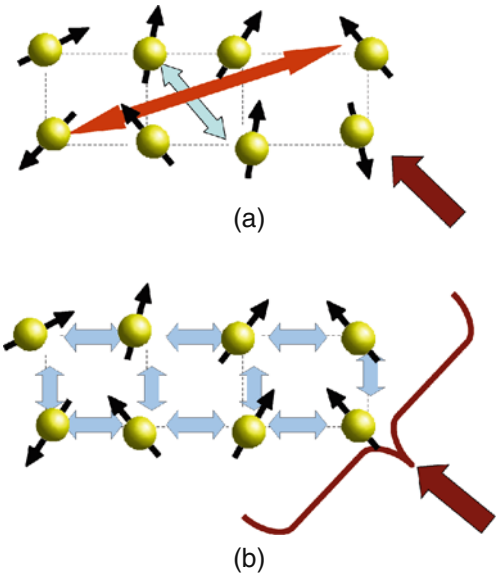


Fig. 1 Part (**a**) of the figure depicts an ideal quantum computer as the ultimate example of an engineered and controllable many-qubit system. The double-ended block arrows show interactions that should be switchable at will and the single-ended block arrow depicts the control/observation of a single qubit by external fields/probes. Part (**b**) of the figure depicts a typical (more natural, less designed) quantum many-body system in which interations (*double-ended block arrows*) between the qubits are permanent and external fields/probes (*single-ended block arrow* ending with the brace denoting the region influenced by it) can only address/access many of the qubits at once

Suppose the qubits cannot be individually manipulated but only in a coarse grained sense, i.e., as many qubits at once? This reduced control situation is depicted in Fig. 1(b). Under such reduced control, the system becomes essentially equivalent, as far as its Hamiltonian (and thereby its equilibrium states and dynamical evolution) is concerned to natural quantum many-body systems such as a magnet. In a magnet, for example, interactions are constant and uncontrollable and too many spins are packed together to be individually addressed or manipulated. This is how the topic of whether one can accomplish any useful quantum information processing with a many-body system automatically arises. It may come as a surprise that even with the above reduced degree of control, one can indeed accomplish non-trivial tasks such as quantum communication, and this will be one of the topic of lectures 3 and 4. For designed many-body systems such as arrays of quantum dots or Josephson junctions, it might be possible to individually address and manipulate quantum systems, but controllable switching of interactions may yet be impossible as the interaction depends on the placement of the elements of the array and may be frozen in time. These slightly more controllable quantum many-body systems can be made to accomplish quantum computation, as we will see in Lecture 5. In Lecture 2, we will take the modest step of examining whether quantum many-body systems can merely serve as a resource for truly "quantum" correlations or entanglement. This will provide the impetus for the more ambitious program of the later lectures. In this

very first lecture, we will present the key notions from quantum information that will be required in the subsequent lectures. We will also introduce typical examples of many-body systems and the type of Hamiltonians that govern them.

2 Lecture 1

In this lecture, we will introduce those basic notions of quantum information that are required for following the rest of the lectures.

2.1 Quantum Entanglement

Entanglement is the truly "quantum" part of the correlations that exist between two systems when they are in states such as $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$ which can never be written down as a pure product of states of the individual systems such as in the form $|\chi\rangle|\phi\rangle$ and *not even in the form of a probablility distribution over such products* such as $\sum_i p_i |\chi_i\rangle\langle\chi_i| \otimes |\phi_i\rangle\langle\phi_i|$ (to take into account the possibility of mixed states). The study of entanglement is a huge area of quantum information science, and we refer the reader to a review such as Ref. [3]. An important point to note is that entanglement is quantifiable. Simply stating, the "harder" it is to approximate a state as a probability distribution over products of pure states, the "higher" is its entanglement. For example, for the case of two qubits, the state $|\psi^-\rangle$ is the most entangled, whose amount is generally set to unity, while a product state of the form $|\chi\rangle|\phi\rangle$ or mixed states of the form $p|\chi_1\rangle\langle\chi_1| \otimes |\phi_1\rangle\langle\phi_1| + (1-p)|\chi_2\rangle\langle\chi_2| \otimes |\phi_2\rangle\langle\phi_2|$ have zero entanglement. In the following lectures, we will require quantifications of entanglement. This is for two reasons. First, being an uniquely quantum feature, more the entanglement between constituents of a complex many-body system, more is its true quantum nature evident. Thus we will need measures of entanglement in Lecture 2, when we analyze how truly quantum different many-body systems are—setting the background for quantum information applications of such systems. Secondly, the *more* entangled a quantum state is, it is *better* for uniquely quantum communication applications such as quantum teleportation [4], whereby one can use an entangled state shared between well-separated parties to transmit an arbitrary quantum state from one party to another using measurements and sending only classical bits from one party to another. Thus in Lectures 3 and 4, when we want to use a many-body system as a quantum communication channel, we can investigate how it transfers entanglement. The higher the entanglement that can be transmitted through such a channel, the better it is at quantum communications—thus we will again require a quantification of entanglement.

For the purposes of this lecture series, it suffices to note how, given a density matrix ρ of two qubits (in an arbitrary and generally mixed state), its entanglement can be computed [5]. The procedure is to first compute the matrix

$$\tilde{\rho} = \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y, \tag{1}$$

where the complex conjugate ρ^* of ρ is taken in the basis $|00\rangle$, $|01\rangle$,$|10\rangle$,$|11\rangle$. Then the entanglement can be quantified by a number called concurrence $\mathcal{E}$ given by [5]

$$\mathcal{E} = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \tag{2}$$

where the λ_is, $i = 1, \ldots, n$, are the square roots of the eigenvalues of $\rho\tilde{\rho}$ in decreasing order.

A second measure of entanglement that we will use in these lectures applies to the entanglement of two systems of arbitrary dimensions (such as the case when each system has several qubits, rather than a single qubit), but only for pure states. This is computed by first computing the reduced density matrix ρ_A of system A from the state ρ_{AB} of the total system using the procedure $\rho_A = \mathrm{Tr}_B(\rho_{AB})$, where $\mathrm{Tr}_B(\ )$ denotes partial tracing over system B. The entanglement of the two systems A and B in the pure state ρ_{AB} is then given by the von Neumann entropy [6]

$$\mathcal{S} = -\mathrm{Tr}\{\rho_A \log_2 \rho_A\}, \tag{3}$$

which in terms of the eigenvalues η_i of ρ_A, is $\mathcal{S} = -\sum_i \eta_i \log_2 \eta_i$.

2.2 Fidelity

In order to judge how well a quantum state is transferred by a many-body system (or more specifically a spin chain) in Lecture 3, we will have to use a figure of merit. Suppose the state that is put in a channel is the pure state $|\psi_{\mathrm{in}}\rangle$ and the state that comes out of the channel is ρ_{out} (the output state is depicted by a density operator to allow for the possibility for it to be a mixed state). Then a measure of the quality of the transfer is defined by the fidelity:

$$F = \langle\psi_{\mathrm{in}}|\rho_{\mathrm{out}}|\psi_{\mathrm{in}}\rangle, \tag{4}$$

which is always between 0 and 1, with higher value meaning better transfer (it is unity for perfect transfer). A fidelity of $2/3$ can already be obtained measuring a state and simply reconstructing it from this data. Thus F needs to be greater than $2/3$ for any quantum communication scheme for it to be any better than straightforward classical communication.

2.3 Operations Required to Build a Quantum Computer

Qubits, which can exist in an arbitrary superposition of states $|0\rangle$ and $|1\rangle$, are generally assumed to be the information bearing degrees of a quantum computer. Two types of operations on a collection of qubits are required in order to build a universal quantum computer. First, one must be able to rotate the state of a qubit, which

is represented as a point on a sphere (the Bloch sphere) to anywhere else in the Bloch sphere. In other words, one must be able to act on a qubit in any initial state $\alpha|0\rangle + \beta|1\rangle$ and obtain the state $\alpha'|0\rangle + \beta'|1\rangle$. These first class of required operations can be called *arbitrary local unitary operations*. Second, one should be able to perform at least one unitary operation which can entangle any two qubits of the quantum computer [7]. This second class of operations can be called *entangling unitary operations*, and along with arbitrary local unitary operations, is sufficient to perform universal quantum computation. To clarify a little, given any quantum algorithm (which is just a unitary operation on a large collection of qubits), one can implement it through a series of arbitrary local unitary operations alternating with several entangling unitary operations on pairs of qubits of the collection. These elementary operations are called quantum gates. In fact, if the reader has heard about certain canonical two-qubit unitary operations such as the controlled NOT (or CNOT) and controlled phase gates, these are specific examples of two-qubit entangling unitary operations (any such two-qubit entangling operation suffices). For those readers not familiar with such gates, we give an example of the controlled phase or CZ gate which converts: $|00\rangle \rightarrow |00\rangle$,$|01\rangle \rightarrow |01\rangle$,$|10\rangle \rightarrow |10\rangle$,$|11\rangle \rightarrow -|11\rangle$.

In Lecture 5, when we want to show that a certain many-body system, more specifically a spin chain, can be made to work as a quantum computer, all we will have to show is that both arbitrary single-qubit gates and a certain entangling two-qubit gate can be accomplished on/between qubits of such a quantum computer.

2.4 Many-Body Spin Systems: Examples

Spins are systems endowed with tiny quantized magnetic moments. Bulk materials often have a large collection of spins permanently coupled to each other. The mutual interactions of these spins makes them prefer alignment or anti-alignment with respect to each other, resulting in diverse phenomena such as ferromagnetism and anti-ferromagnetism. A spin chain models a large class of such materials in which the spins are arranged in a one-dimensional lattice and permanently coupled to each other, usually with an interaction strength decreasing with distance. A common form of the model Hamiltonian for the interaction between the ith and the jth spin is written as

$$\mathbf{H}_{ij} = J_{ij}\, \mathbf{S}^i \cdot \mathbf{S}^j, \tag{5}$$

where $\mathbf{S}_i \cdot \mathbf{S}_j \equiv S_x^i S_x^j + S_y^i S_y^j + S_z^i S_z^j$ and S_x^i, S_y^i, S_z^i are the operators for the component of the ith spin along the x, y and z directions, respectively. In particular, when all the spins are spin-1/2 systems, S_x, S_y and S_z stand for the familiar Pauli matrices σ_x, σ_y and σ_z. A Hamiltonian of the above form is termed as an *exchange interaction* as it can arise in from the pure exchange electrons between neighboring atoms in a metal. It is also called the *Heisenberg interaction* after its inventor. In

particular, the specific Hamiltonian we have written above is called the *isotropic* exchange interaction. It is, in fact, the most natural form of interaction between spins. In this chapter we will also encounter a variant of the above interaction which has an anisotropy

$$\mathbf{H}_{ij}^{XY} = J_{ij}\ (S_x^i S_x^j + S_y^i S_y^j), \tag{6}$$

which is called the XY interaction. We will be primarily concerned with chains of spin-1/2 systems in this chapter. At an opposite extreme of anisotropy, there is also another type of spin chain called the Ising spin chain with the Hamiltonian:

$$\mathbf{H}_{ij}^{I} = J_{ij}\ S_z^i S_z^j. \tag{7}$$

3 Lecture 2

In this lecture, we are going to ask the simplest possible question of how much entanglement exists "inside" a condensed matter system in its stationary states such as the ground state and the thermal states at various temperatures. Why is this question important? This is because the ultimate aim of these series of lectures is to show how many-body systems may be made to process quantum information. Naturally, the constituents, such as individual spins, of these many-body systems will be involved in the processing of quantum information. Thus one needs to be sure that these are bona-fide quantum mechanical systems. One way to be sure is to compute that entanglement between them in most natural states of the system such as ground and thermal states. The amount of this entanglement will tell how much quantum mechanically the constituents of a many-body system are behaving, and indeed thus they fulfill the basic criterion of being truly quantum, so that one can think of information processing with them storing quantum information. We should note that just proving that they are behaving quantum mechanically (namely being highly entangled with each other) in a stationary state in equilibrium with the environment does not mean that this state itself serves as a good starting state for quantum information processing. Neither does it prove that during dynamics they will continue to behave truly quantum mechanically despite their interaction with their environment. However, computing entanglement is a good starting point for thinking quantum mechanically about the ingredients of a many-body system. Additionally, the entangled state of multiple particles inside a many-body system may serve as a resource extractable and usable in various applications where shared entanglement is necessary.

Historically, the entanglement inside complex many-body systems was the first issue to be examined and so much work has been done that it is literally impossible to cover them in one lecture. For an extensive review, one can look up Ref.[8]. Here I will take a single simple example and try to illustrate all the various kinds of entanglement measures generally computed for quantum many-body systems and offer what we think are reasonable motivations for studying these measures. The

model will be an open-ended chain of four spin-1/2 particles coupled through a nearest-neighbor isotropic Heisenberg interaction, so that the Hamiltonian is

$$\mathbf{H} = \sum_{i=1}^{3} \boldsymbol{\sigma}^i \cdot \boldsymbol{\sigma}^{i+1}. \tag{8}$$

The ground state of this chain

$$|GS\rangle = (\sqrt{\frac{2}{3}}|\psi^-\rangle|\psi^-\rangle - 0.1494(|00\rangle|11\rangle - |01\rangle|01\rangle - |10\rangle|10\rangle + |11\rangle|00\rangle) \tag{9}$$

is manifestly an entangled state (just looking at the form of the state is enough to spot that) and for a long time condensed matter physicists have known that the ground states of such systems are indeed entangled. What is more important, though, is the "amount" of entanglement between two spins of the system. This had not been computed till very recently, and in fact, well after the advent of quantum information. To compute the entanglement, we first obtain the reduced density matrix ρ_{ij} of the spins i and j. From the expression of $|GS\rangle$, it is clear that ρ_{12}, for example, is a mixed state with a significant proportion (2/3) of the maximally entangled state $|\psi^-\rangle$. The entanglement between spins i and j is computed as the concurrence $\mathcal{E}$ from the formula given in Section 2.1. However, it is worth mentioning here that because of certain symmetries of the Heisenberg model, the concurrence $\mathcal{E}$ reduces to a very simple formula and one need not involve all elements of the reduced density matrix ρ_{ij} for the calculation of concurrence. Note that all the states involved in the expression for $|GS\rangle$ have the same number of zeros. This is a consequence of the commutation of $\mathbf{H}$ with $\sum_i \sigma_z^i$ and holds for all eigenstates of $\mathbf{H}$, and consequently also for their mixtures such as thermal states. It is then easy for the reader to verify (I leave it as an exercise here) that the density matrix ρ_{ij} cannot have any off-diagonal terms (or coherence) between spaces with different values of $\sigma_z^i + \sigma_z^j$. In the standard basis $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, ρ_{ij} will thus be of the form

$$\begin{pmatrix} x & 0 & 0 & 0 \\ 0 & y_1 & z & 0 \\ 0 & z^* & y_2 & 0 \\ 0 & 0 & 0 & w \end{pmatrix}.$$

For such a simple form of ρ_{ij}, the concurrence is given by the simple formula $\mathcal{E} = 2\max\{0, |z| - \sqrt{xw}\}$.

The concurrence between spins 1 and 2 is found to be 0.866, which is quite high (the highest possible value, for a maximally entangled state, being 1). These are nearest neighbors. On the other hand, no entanglement exists between 2 and 3, though they are nearest neighbors (the entanglement pattern of the chain is dimerized because of its open ends [9]) and there is no entanglement between any of the non-nearest neighbor spins (such as 1 and 3 or 1 and 4). Interestingly, the

entanglement between 1 and 2 does not die down completely even when the system is put in a thermal state (in a thermal equilibrium with its environment at a temperature T) given by

$$\rho = e^{-\mathbf{H}/K_B T}/Z, \tag{10}$$

where K_B is the Boltzmann constant and Z is the partition function given by $Z = \sum_k e^{-E_k/K_b T}$, where E_k are the eigenvalues of $\mathbf{H}$. It is, of course, trivial that when the system is at a temperature such that $K_B T \ll 1$ (as the strength of the Hamiltonian $\mathbf{H}$ is 1), only the state $|GS\rangle$ is significantly populated, and entanglement is the same as that in the state $|GS\rangle$. What is however surprising is the fact that even when $K_B T \sim 3$ (the effect of temperature is not only comparable, but also higher than the strength of the Hamiltonian), the entanglement between spins 1 and 2 still persists, and the concurrence equals 0.06. This kind of entanglement which exists between the constituents of a physical system in its thermal state is called *thermal entanglement* [10].

An additional interesting feature is that the entanglement between further neighbors in the ground state can also be switched on by applying a magnetic field to the system. For example, for a magnetic field $B = 2$, the ground state becomes

$$|GS(B=2)\rangle = 0.2706(|0001\rangle - |1000\rangle) + 0.6533(-|0010\rangle + |0100\rangle). \tag{11}$$

The concurrence of spins 1 and 2 and of spins 1 and 3 (next nearest neighbor) are both 0.3536, while that of 1 and 4 (farthest neighbors) is 0.1464. One can call this entanglement *magnetic entanglement* [10]. In fact, it has been shown that a magnetic field may be used to prepare a multi-spin entangled state as a ground state of a Heisenberg Hamiltonian in which each pair of spins in a chain, irrespective of their distance, are entangled with each other [11].

In the context of the above kind of entanglement, namely that between two spins in a spin chain system in its ground/thermal state *without* doing/performing anything on the system, calculations have actually been done for long chains using analytic results known about the correlation functions, which have for long been calculated in condensed matter physics. This is possible because, in the end, all the elements of the density matrix are given in terms of the expectation values of $\langle \sigma_\alpha^i \rangle$ (magnetization) and $\langle \sigma_\alpha^i \sigma_\beta^j \rangle$ (correlation functions), where α/β run over x, y and z. Thus analytic computations of the concurrence is possible from the correlation functions, but generically, only a combination of them. One should be warned, for example, that two states which have the same amount of correlations in a certain direction do not necessarily have the same amount of entanglement. For example, the states $|\psi^-\rangle$ (ground state of a 2-spin Heisenberg antiferromagnet) and $|01\rangle$ (one of the ground states of a 2-spin Ising antiferromagnet) both have correlation $\langle \sigma_z^1 \sigma_z^2 \rangle = -1$, but only the first state is entangled. By computing entanglement from combinations of correlation functions, it has been shown that near *quantum phase transitions* [12]—when the ground state of certain quantum many-body systems undergo a sudden "qualitative" change due to the variation of a parameter of the Hamiltonian,

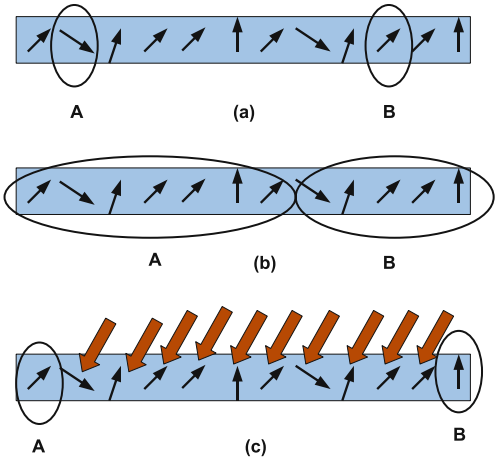


Fig. 2 A pictorial representation of the different types of entanglement usually computed for many-body systems. Part (**a**) shows the simplest form of entanglement that one can study, namely that between two spins A and B of a many-spin system. Part (**b**) shows two subsystems A and B, each entire block of spins by themselves, between which entanglement can also be studied. Part (**c**) shows the process of measuring several spins of a system to localize entanglement in the unmeasured spins A and B

the entanglement between two spins i and j, usually nearest and next nearest neighbors, can peak [13, 14]. Speaking very roughly and qualitatively, the ground state near a quantum phase transition is a highly entangled state because of a competition between different ordering tendencies of different terms of a Hamiltonian. On either side of the transition different tendencies win and impose their order, while at the transition neither can win and only an entangled state can be the lowest energy state.

The entanglement between two constituents of a many-body system is not the only type of entanglement which can be computed to characterize the true quantum aspects of a many-body system. In fact, Fig. 2 shows three different types of entanglement that have been computed for many-body systems. In Fig. 2(a), the systems A and B between which entanglement is computed are depicted to be individual spins. This was the case we were discussing above, and one of the latest developments in this category is the observation that if spins A and B are weakly coupled to the rest of the spin system, they could even be maximally entangled, such as in a state $|\psi^-\rangle$, in the ground ($T = 0$) state [15]. This type of entanglement has been termed as *long distance entanglement*.

Generically, the entanglement between two spins of a many-spin system ignores a lot of the entanglement which is present within the system. Such as, in our above example, when we compute the entanglement of spins 1 and 4 of our four spin system described by $\mathbf{H}$, or separately that between spins 2 and 3, this entanglement does not tell us fully about the entanglement between two blocks, one containing 1 and 2 and the other containing 3 and 4. Thus the computation of entanglement between

blocks has also recently been a popular area of research, a situation depicted in Fig. 2(b), where the blocks are collections of spins encircled as A and B. As the full system is in a pure state, the entanglement can easily be computed as the von Neumann entropy of either the block A or the block B, as in Eq. (3). For example, for the state $|GS\rangle$, the entanglement between block A consisting of spins 1 and 2 and block B consisting of spins 3 and 4 is 0.4608. Clearly the entanglements between the individual spins do not reveal this entanglement in the sense that neither of the spins in block A are individually entangled with either of the spins in block B. Incidentally, the entanglement between the blocks A and B increases to unity by applying a magnetic field to reach the ground state $|GS(2)\rangle$. However, this is not the limit of the entanglement between the blocks as they are each two-state systems. For example, by making the coupling between spins 2 and 3 much stronger that the other couplings in the system, the entanglement between the blocks can be made to approach their maximal value of 2. This happens because the ground state assumes a form where 1 and 4 are in a singlet and 2 and 3 are in another singlet—the mechanism here is the same as that leading to long distance entanglement of spins weakly coupled to a spin chain [15]. For an order of magnitude difference of the couplings, the entanglement is about 1.97 and it grows to 1.99 as the couplings are made two orders different. There is though a price to pay, for each order of change of the ratio of the couplings, the energy separation of the ground state with the next energy eigenstate decreases—so that the entanglement would be demise fast at finite temperature. In general, the block entanglement, not restricted in amount by the dimensions of individual spins, may grow with lengths of the blocks and may even diverge with the length in certain circumstances, though this may come at the price of the ground state becoming very close to the lowest excited state (the reason for these states to come close may differ from case to case). This is precisely the situation near to a quantum phase transition (or a quantum critical point) where the entanglement between blocks of contiguous spins is found to diverge as $\log N$, where N is the length of a block [16–18]. On the other hand, for gapped systems, such as certain spin-1 chains, the entanglement of two blocks saturates to a finite value and remains so irrespective of the length of the blocks [19].

Yet another type of entanglement which has been computed in spin chains is called *localizable entanglement* [20] and its definition can be understood by looking at Fig. 2(c). Here the intermediate spins of a many-spin system are individually measured to entangle the spins most distant from each other. The highest amount of entanglement that can be established between the unmeasured spins, such as spins A and B in Fig. 2(c) by *optimizing* the measurement of the other spins, is called the localizable entanglement. In this context, it is worth pointing out that there exist curious many-qudit (quantum d-dimensional systems) systems in whose ground state the entanglement establishable between any two of the qudits is always unity (equivalent to that of a two-qubit maximally entangled state) even when all the other qudits are measured completely randomly (i.e., in random bases obtaining random outcomes) [21].

Before concluding this section we should point out that in addition to spin chains, harmonic oscillator chains (in both ground [22] and thermal [23] states) has been another system studied widely for the entanglement that exists naturally inside them.

4 Lecture 3

In this lecture we will study the possibility of using spin chains as channels for quantum communication. The main motivation for this is to do away with optics when connecting solid state based quantum registers. The array of spins itself serves as an "all solid-state" bus for connecting the registers.

4.1 A Simple Spin Chain Quantum Communication Protocol

Let us now present one of the simplest possible protocols for quantum communication through spin chains. This is based on the scheme presented in Ref.[24], where the idea of using a spin chain as a communication channel was first introduced. The simplest possible spin chain is one composed of qubits. We start our protocol with the spin chain initialized in a very simple state, such as the one in which all spins are in the state $|0\rangle$. We will have to choose the couplings J_{ij} in Eq. (5) in such a manner that initialization of the spin chain to such a state is easy. Accordingly, we choose $J_{ij} < 0$, which means that the spin chain describes a *ferromagnet*. The ground state of a ferromagnet in a magnetic field, however weak, is a *symmetry broken* state in which all spins align with the direction of the field. For example, all spins could be pointing down. For communication, Alice places an arbitrary quantum state at one end of the spin chain in such an "all down" state. This is depicted in the upper part of Fig. 3, where Alice has placed an arbitrary state on the first spin of the chain, while all the other spins are still in the down state. Due to the natural evolution of the chain this state both disperses and propagates along the chain. As a result of this

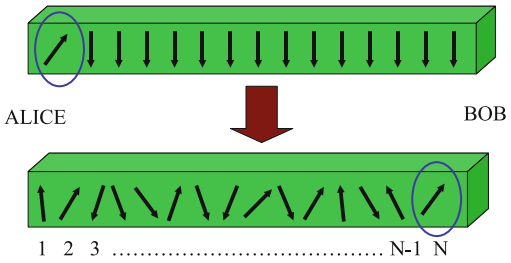


Fig. 3 The simplest spin chain communication protocol. A spin-1/2 ferromagnetic spin chain with all spins facing down is the quantum channel. Alice simply places a quantum state at one end of the chain and Bob simply picks up a close approximation of this state from his end of the spin chain after waiting a while

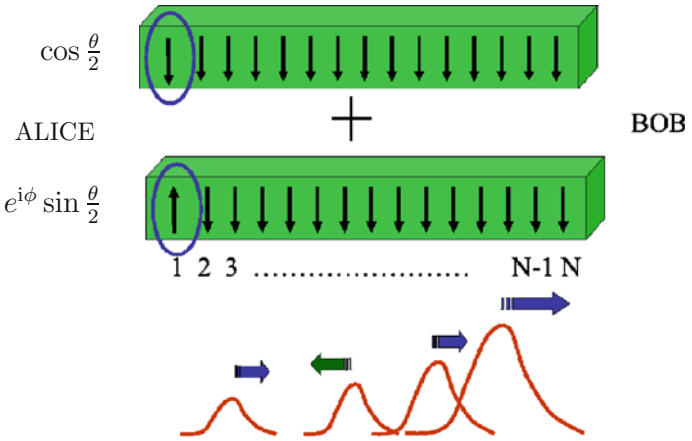


Fig. 4 The figure shows how quantum information is transmitted down a spin chain in the simplest protocol. The chain goes to a superposition of its ground state and a time-evolving state. The time-evolving state transmits a spin flip as a series of wave-packets which travels towards Bob while dispersing at the same time

evolution, the state of the spin at Bob's end of the chain will vary with time (see Fig. 4). Bob now chooses an optimal moment of time in a long interval (as long as he can afford to wait!) to receive Alice's state. This moment of time is carefully chosen so that the state of the spin at Bob's end of the chain is as close as possible to the one that Alice intended to transmit. At this optimum time, Bob simply picks up the spin at his end of the chain to conclude the communication protocol.

4.2 Formula for Fidelity of the Above Quantum Communication Protocol

Say there are N spins in the chain and these are numbered $1, 2, \ldots, N$ as shown in Fig. 3. The Hamiltonian is given by only nearest neighbor interactions, the interactions are uniform with equal strength J at each site and in a uniform magnetic field, i.e.,

$$\mathbf{H} = \sum_i J\, \boldsymbol{\sigma}^i \cdot \boldsymbol{\sigma}^{i+1} - \sum_{i=1}^{N} B\sigma_z^i. \tag{12}$$

In the above equation $\boldsymbol{\sigma}^i = (\sigma_x^i, \sigma_y^i, \sigma_z^i)$ in which $\sigma_{x/y/z}^i$ are the Pauli matrices for the ith spin, $B_i > 0$ are *static* magnetic fields and $J < 0$ are coupling strengths. $\mathbf{H}$ describes a ferromagnet with isotropic Heisenberg interactions. As mentioned above, we initialize the ferromagnet in its ground state $|\mathbf{0}\rangle = |000\ldots0\rangle$ where $|0\rangle$ denotes the spin down state (i.e., spin aligned along $-z$ direction) of a spin. This can be accomplished easily for a ferromagnetic system by cooling and applying a magnetic field, however small. We will set the ground state energy $E_0 = 0$ (i.e.,

redefine $\mathbf{H}$ as $E_0 + \mathbf{H}$) for the rest of this chapter. We also introduce the class of states $|\mathbf{j}\rangle = |00\ldots010\ldots.0\rangle$ (where $\mathbf{j} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{N}$) in which the spin at the jth site has been flipped to the $|1\rangle$ state. We now assume that the state sender Alice is located closest to the first spin and the state receiver Bob is located closest to the Nth spin. As mentioned before, to start the protocol, Alice simply places the state she wants to transmit to Bob on the first spin at time $t = 0$. Let this state be $|\psi_{\text{in}}\rangle = \cos(\theta/2)|0\rangle + e^{\text{i}\phi}\sin(\theta/2)|1\rangle$. We can then describe the state of the whole chain at this instant (time $t = 0$) as

$$|\Psi(0)\rangle = \cos\frac{\theta}{2}|\mathbf{0}\rangle + e^{\text{i}\phi}\sin\frac{\theta}{2}|\mathbf{1}\rangle. \tag{13}$$

Bob now waits for a specific time till the initial state $|\Psi(0)\rangle$ evolves to a final state which is as close as possible to $\cos\frac{\theta}{2}|\mathbf{0}\rangle + e^{\text{i}\phi}\sin\frac{\theta}{2}|\mathbf{N}\rangle$. As $[\mathbf{H}, \sum_{i=1}^{N}\sigma_z^i] = 0$, the state $|\mathbf{1}\rangle$ only evolves to states $|\mathbf{j}\rangle$ and the evolution of the spin chain (with $\hbar = 1$) is

$$|\Psi(t)\rangle = \cos\frac{\theta}{2}|\mathbf{0}\rangle + e^{\text{i}\phi}\sin\frac{\theta}{2}\sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{N}}\langle\mathbf{j}|e^{-\text{i}\mathbf{H}t}|\mathbf{s}\rangle|\mathbf{j}\rangle. \tag{14}$$

The state of the Nth spin will, in general, be a mixed state, and can be obtained by tracing off the states of all other spins from $|\Psi(t)\rangle$. This means that the density operator ρ_{out} of the output state is obtained by $\text{Tr}_{1,2,\ldots,N-1}(|\Psi(t)\rangle\langle\Psi(t)|)$, where $\text{Tr}_{1,2,\ldots,N-1}$ means tracing over the states of the systems 1 to $N-1$. This evolves with time as

$$\rho_{\text{out}}(t) = P(t)|\psi_{\text{out}}(t)\rangle\langle\psi_{\text{out}}(t)| + (1 - P(t))|0\rangle\langle 0|, \tag{15}$$

with

$$|\psi_{\text{out}}(t)\rangle = \frac{1}{\sqrt{P(t)}}(\cos\frac{\theta}{2}|0\rangle + e^{\text{i}\phi}\sin\frac{\theta}{2}f_N(t)|1\rangle), \tag{16}$$

where $P(t) = \cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}|f_N(t)|^2$ and $f_N(t) = \langle\mathbf{N}|\exp\{-\text{i}\mathbf{H}t\}|\mathbf{1}\rangle$. Note that $f_N(t)$ is just the transition amplitude of an excitation (the $|1\rangle$ state) from the first to the Nth site of a graph of N spins.

Now suppose it is decided that Bob will pick up the Nth spin (and hence complete the communication protocol) at a predetermined time $t = t_0$. The fidelity of quantum communication through the channel averaged over all pure input states $|\psi_{\text{in}}\rangle$ in the Bloch-sphere $((1/4\pi)\int\langle\psi_{\text{in}}|\rho_{\text{out}}(t_0)|\psi_{\text{in}}\rangle d\Omega)$ is then

$$F = \frac{|f_N(t_0)|\cos\gamma}{3} + \frac{|f_N(t_0)|^2}{6} + \frac{1}{2}, \tag{17}$$

where $\gamma = \arg\{f_N(t_0)\}$. To maximize the above average fidelity, we must choose the magnetic fields B_i such that γ is a multiple of 2π. Assuming this special choice

of magnetic field value (which can always be made for any given t_0) to be a part of our protocol, we can simply replace $f_N(t_0)$ by $|f_N(t_0)|$ in (16).

4.3 Performance of the Protocol

We now want to examine the performance of the spin chain described in the last subsection, namely the nearest-neighbor isotropic Heisenberg Hamiltonian **H**, as a quantum channel. The eigenstates relevant to our problem are

$$|\tilde{m}\rangle_{\mathbf{L}} = a_m \sum_{j=1}^{N} \cos\{\frac{\pi}{2N}(m-1)(2j-1)\}|\mathbf{j}\rangle, \tag{18}$$

where $m = 1, 2, \ldots, N$, $a_1 = 1/\sqrt{N}$ and $a_{m\neq 1} = \sqrt{2/N}$ with energy (on setting $E_0 = 0$) given by $E_m = 2B + 2J(1 - \cos\{\frac{\pi}{N}(m-1)\})$. In this case, $f_N(t_0)$ is given by

$$f_N(t_0) = \sum_{m=1}^{N} \langle \mathbf{N}|\tilde{m}\rangle\langle\tilde{m}|\mathbf{1}\rangle e^{-\mathrm{i}E_m t_0} = \mathcal{C}(v_m), \tag{19}$$

where $v_m = a_m \cos\{\frac{\pi}{2N}(m-1)(2N-1)\}e^{-\mathrm{i}E_m t_0}$ and

$$\mathcal{C}(v_m) = \sum_{m=1}^{N} a_m v_m \cos\{\frac{\pi}{2N}(m-1)\} \tag{20}$$

is the first element of the inverse discrete cosine transform of the vector $\{v_{m,r}\}$.

We now want to study the performance of our protocol for various chain lengths N (Alice and Bob at opposite ends of the chain as shown in Fig. 3). In general, Bob has to wait for different lengths of time t_0 for different chain lengths N, in order to obtain a high fidelity of quantum state transfer.

Using (17) and (19), we can numerically evaluate the maximum of $|f_N(t_0)|$ for various chain lengths from $N = 2$ to $N = 80$ when Bob is allowed to choose t_0 within a finite (but long) time interval of length $T_{\max} = 4000/J$. This evaluation is fast because (19) allows us to use numerical packages for the discrete cosine transform. Taking a finite $T_{\max}$ is physically reasonable, as Bob cannot afford to wait indefinitely. It is to be understood that within $[0, T_{\max}]$, the time t_0 at which optimal quantum communication occurs varies with N. The maximum fidelities as a function of N and the maximum amounts of entanglement shareable (both rounded to 3 decimal places) are shown in Fig. 5.

Figure 5 shows various interesting features of our protocol. The plot also shows that in addition to the trivial case of $N = 2$, $N = 4$ gives perfect ($F = 1.000$) quantum state transfer to 3 decimal places and $N = 8$ gives near perfect

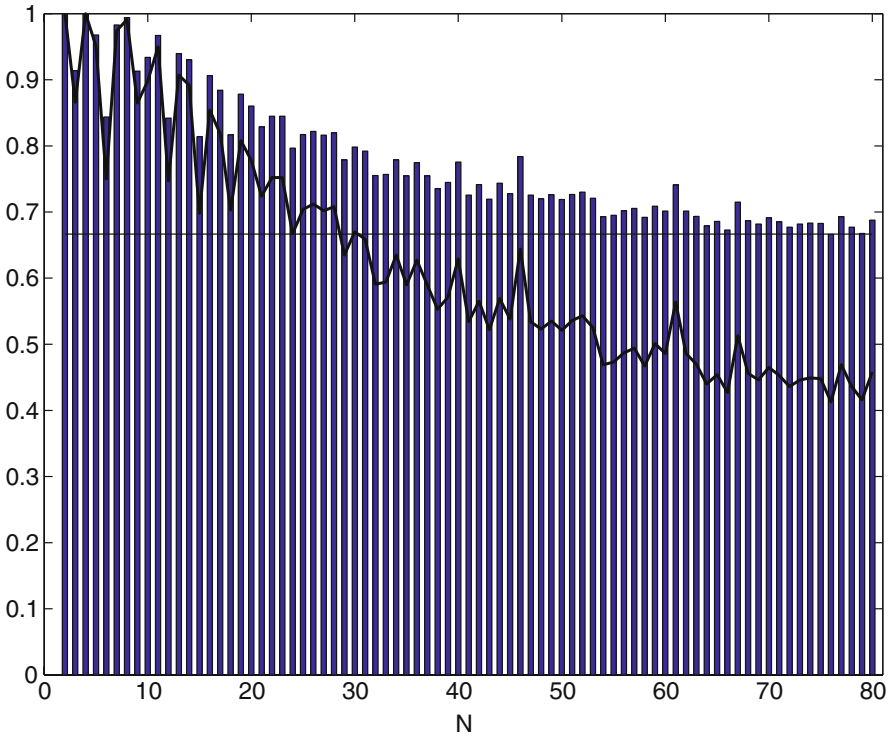


Fig. 5 The bar plot shows the maximum fidelity F of quantum communication and the curve shows the maximum sharable entanglement $\mathcal{E}$ achieved in a time interval $[0, 4000/J]$ as a function of the chain length N from 2 to 80. The time t_0 at which this maxima is achieved varies with N. The straight line at $F = 2/3$ shows the highest fidelity for classical transmission of a quantum state

($F = 0.994$). The fidelity also exceeds 0.9 for $N = 7, 10, 11, 13$ and 14. Till $N = 21$ we observe that the fidelities are lower when N is divisible by 3 in comparison to the fidelities for $N+1$ and $N+2$. While we do not have a clear cut explanation of this effect, it is obviously a link between number theory and constructive interference in a line. The plot also shows that a chain of N as high as 80 exceeds the highest fidelity for classical transmission of the state, i.e., $2/3$ in the time interval probed by us.

4.4 Transmission of Entanglement

As seen above, the simplest protocol with the simplest chain does not enable one to do a quantum communication with unit fidelity for reasonable lengths. However, this is precisely what is required for linking distinct quantum registers, which is our ultimate aim! If we want to stick to such a simple scheme, whether it be motivated by the ease of experimentation or otherwise, we thus have to transmit entanglement

instead. Of course, even the transmission of this entanglement will not be perfect (in fact, "how much" entanglement our channel can transmit will be the subject of the current subsection). After several transmissions of an imperfectly entangled state, one can perform entanglement distillation [25]—a scheme that uses only local operations and classical communication on a set of partially entangled pairs of particles to obtain a smaller set of nearly maximally entangled pairs of particles (i.e., in a state $|\psi^+\rangle$, for example). This maximally entangled pair of particles can subsequently be used for perfect quantum communications using teleportation. With this motivation, we will now examine how entanglement can be transmitted through such a chain.

We will look at the transmission of the state of one member of a pair of particles in the entangled state $|\psi^+\rangle$ through the spin chain channel. This is the usual procedure for sharing entanglement between separated parties through any channel. Alice prepares two qubits in the state $|\psi^+\rangle$, holds one of them (say, A) in her hand and places the other on site 1 of the chain. After waiting for an optimal time t_0, Bob picks up the qubit N from the chain. The joint state shared by Alice's qubit A and Bob's qubit N at this time is given by

$$\rho_{1N}(t_0) = \frac{1}{2}\{(1-|f_N(t_0)|^2)|00\rangle\langle 00|_{1N} + (|10\rangle + |f_N(t_0)||01\rangle)(\langle 10| + |f_N(t_0)|\langle 01|)_{1N}\}.$$

The entanglement $\mathcal{E}$ of the above state, as quantified by its concurrence [5], is given by

$$\mathcal{E} = |f_N(t_0)|. \tag{21}$$

Thus, for any non-zero $f_N(t_0)$ (however small), some entanglement can be shared through the channel. This entanglement, being that of a 2×2 system, can also always be subject to entanglement distillation [26]. Thus several copies of the entangled state $\rho_{1N}(t_0)$ are to be shared in parallel between Alice and Bob and then it is converted to a smaller collection of pure maximally entangled states $|\psi^+\rangle$ shared between Alice and Bob (only through local actions by Alice and Bob and classical communication between them).

5 Lecture 4

In this lecture, we will describe the various methods that have been suggested to perfect communication schemes using spin chains. For connecting two solid state quantum registers, which was our main motivation, it is only perfect or nearly perfect state transfer which is relevant. The case where one transmits the state of one member of a pair of entangled qubits through the channel can be made useful when appended with entanglement distillation procedures. Basically, one has to first use the spin chain channel repeatedly to make Alice and Bob share several copies of

the partially entangled state $\rho_{1N}(t_0)$. Then entanglement distillation is performed to obtain a smaller number of pairs of particles in the state $|\psi^+\rangle$ shared between Alice and Bob. The particles in the state $|\psi^+\rangle$ can now be used to transmit a state perfectly, do quantum gates between qubits in well-separated locations, and so forth. Clearly, it would be better if we could do perfect quantum state transfer through a spin chain without invoking an additional entanglement distillation process. With this view in mind, several schemes have been proposed such as the following ones.

5.1 Engineering the Couplings in a Chain

The couplings J_{ij} can be carefully chosen (even when there are only nearest-neighbor couplings, i.e., $i = j \pm 1$) to obtain a spin chain which accomplishes perfect quantum state transfer [27]. If the couplings are mirror symmetric about the center of the spin chain (*i.e.,* $J_{j,j+1} = J_{N-j,N-j+1}$), then the alternate eigenstates in the single excitation sector (i.e., in a sector when there is only one spin flip in a fully polarized background) have opposite parity (much like a particle in a box of elementary quantum mechanics). More precisely, the energy eigenstates $|\phi_k\rangle$ of the spin chain in the first excitation sector with $k = 0, 1, \ldots, N-1$ satisfy $\langle \mathbf{N} - \mathbf{j}|\phi_k\rangle = (-1)^k \langle \mathbf{j}|\phi_k\rangle$. Additionally, the couplings are so chosen that the spectrum of the Hamiltonian is commensurate (the energies are proportional to integers), for example lets assume that $E_k \propto k^2$. Then, due to time evolution of *any* state $|\psi(t=0)\rangle = \sum_k c_k |\phi_k\rangle$ of the chain for a special time $t = \tau$ such that $E_k \tau = k^2 \pi$ we have

$$\langle \mathbf{j}|\psi(t=\tau)\rangle = \sum_k c_k (-1)^k \langle \mathbf{j}|\phi_k\rangle = \langle \mathbf{N} - \mathbf{j}|\psi(t=0)\rangle. \tag{22}$$

Thus there is a time τ in which the complete state in the first excitation sector *mirror inverts* about the center of the chain. Naturally, a state $|\mathbf{1}\rangle$ of the chain will evolve to $|\mathbf{N}\rangle$, which implies perfect state transfer. Before ending we add (and this is something that will be used in Lecture 5) that in Ref.[28] it has been shown that the mirror inversion works for any excitation sector of the spin chain (i.e., any number of ones and zeros in the chain).

5.2 Encoding in Many-Spin States (Wave-Packets)

It has been suggested [29] that one can use "truncated" Gaussian wave-packet states

$$|G(j_0, k_0)\rangle = \sum_j e^{-(j-j_0)^2/L^2} e^{-\mathrm{i}k_0 j} |\mathbf{j}\rangle$$

centered at the site $j = j_0$ (and defined over L sites around the site j_0) and with velocity $\propto k_0$ to encode the logical $|1\rangle$ state of the qubit to be transmitted. In a block of L spins near a site j_A, Alice encodes the $|1\rangle$ state of the qubit she wants to transmit

on $|G(j_A, k_0)\rangle$ (the $|0\rangle$ is encoded as in the previous protocols). For appropriate choice of k_0 (see Ref.[29]) one can choose $L \sim N^{1/3}$, so that this wave-packet travels with a dispersion which is negligible and remains constant no matter how far it travels down the chain (strictly speaking, in Ref.[29], a ring of spins was used, but Ref.[30] has also modified this scheme so that it is applicable to other graphs such as an open-ended chain). Bob located at any distant site along the chain/ring can catch almost the entire wave-packet by using a sufficiently long block of spins ($\sim N^{1/3}$) to receive the state.

5.3 Coupling Qubits Weakly to a Quantum Many-Body System

Another approach is to couple the sending and receiving qubits weakly to a quantum many-body system [31–33]. Say the many-body system is an arbitrary graph of spins which interact with each other with a coupling strength $J \sim 1$, while the sending and receiving qubits are coupled to the system through a coupling ϵ where $\epsilon \ll 1$. Moreover, assume all the couplings to be of XY or Heisenberg type (other interactions would also do as long as they can enable the transfer of an excitation through the graph from the sending to the receiving qubit). Then, one can derive *effective* XY or Heisenberg Hamiltonians between the two qubits when there are *no* eigenstates of the many-body system whose energy is close to 0 [31, 33]. This is possible, for example, when the many-body system is in its ground state and has a finite energy gap Δ between the ground and the first excited state (such as a spin ladder [31]). Effectively, a Hamiltonian of the form $\epsilon^2 \mathbf{H}_{sr}^{XY}$ or $\epsilon^2 \mathbf{H}_{sr}$ acts on them, where s and r stand for the sending and the receiving qubits, respectively. The other case is when the many-body system has exactly *one* available state $|\lambda\rangle$ of zero energy of the type in which a single spin is flipped from the ground state, and beyond that, there is a gap Δ to all other states of the single flip type. Then a "resonant" transfer [33] through the many-body system takes place with an effective Hamiltonian $\epsilon \mathbf{H}_{sM}^{XY} + \epsilon \mathbf{H}_{Mr}^{XY}$, where σ_M^x, σ_M^y and σ_M^z are defined for a *delocalized qubit* with $\sigma_M^z = +1$ corresponding to the many-body system being in $|\lambda\rangle$ and $\sigma_M^z = -1$ corresponding to the many-body system being in its ground state [32, 33]. This resonant effective Hamiltonian is just a three qubit XY spin chain which can perfectly transfer states in a timescale $t \sim 1/\epsilon$ [27].

5.4 Ising Chain with Global Pulses

Another approach is to use a spin chain Hamiltonian with a different type of coupling, namely a nearest neighbor Ising coupling, as given by

$$\mathbf{H}^{\text{Ising}} = \sum_{j=1}^{N} J \sigma_j^z \sigma_{j+1}^z, \tag{23}$$

in conjunction with "global" pulses (pulses that act on each spin of the chain in exactly the same way) at regular intervals to perfectly transport a state from one of its ends to the other [34]. To understand this, we will need two unitary operations, one called the Hadamard (denoted by H) which acts on a single qubit to change $|0\rangle$ to $|+\rangle = |0\rangle + |1\rangle$ and $|1\rangle$ to $|-\rangle = |0\rangle - |1\rangle$ and the other the CZ operation discussed in Lecture 1. It is shown in Ref.[34] that an Ising chain evolving on its own for a time $\pi/4J$ followed by fast (instantaneous) operations on the chain by global pulses at time $\pi/4J$ (and some extra operations, also fast, at the very ends of the spin chain) accomplishes the operation S which is equivalent to a CZ between all adjacent pairs of spins followed by an H on each spin. The entire time evolution with interruptions by the instantaneous pulses at regular intervals is then equivalent to a series of applications of S. Now imagine a $N = 4$ spin chain to be initialized in the state $(\alpha|0\rangle_1 + \beta|1\rangle_1)|+\rangle_2|0\rangle_3|+\rangle_4$. Then the reader can easily verify that the successive applications of S operations accomplishes the evolution

$$S^{\otimes 4}(\alpha|0\rangle_1 + \beta|1\rangle_1)|+\rangle_2|0\rangle_3|+\rangle_4 = |0\rangle_1|+\rangle_2|0\rangle_3(\alpha|+\rangle_4 + \beta|-\rangle_4).$$

Thus an extra H operation on each qubit (accomplishable by global pulses) after the above evolution has completely transferred the state at site 1 to site 4. The authors of Ref.[34] show that in general, for a N spin chain, one is able to transfer a quantum state from one end to the other by N applications of S (i.e., evolution of the chain till time $N\pi/4J$ interrupted by instantaneous pulses at regular intervals) and a H on each qubit at the end of the evolution. Moreover, they also show that such a transfer can be accomplished by *any* starting state of the spin chain and that not only does the above protocol accomplish state transfer but also a mirror inversion of the state of the chain about its center.

5.5 Adiabatic Passage

Some very recent works show that if some degree of slow modulation of the couplings are allowed, then adiabatic passage can also be used to transfer quantum states perfectly through a spin chain channel [35, 36].

5.6 Dual-Rail Encoding and Heralded Perfect Transfer

In this protocol one uses *two* spin chains in parallel as opposed to a single chain [37]. The couplings in the spin chains need not be either uniform or specially engineered and could even be mildly random [38] (with the reasonable assumptions that the chains are *similar* to each other and permit state transfer). A switchable two-qubit interaction at the ends of the parallel chains suffices for the encoding and decoding of this scheme. It accomplishes a "heralded" perfect quantum state transmission, where conditional on a positive outcome of a measurement (which happens with a

certain probability), Bob can conclude that he has accurately (i.e., with unit fidelity) received the state transmitted to him. In other words, what was an imperfect fidelity earlier has now been converted to a probability of success.

The idea [37] is to use two spin chains I and II in parallel as a single communication channel as shown in Fig. 6. As in previous protocols, this protocol is also restricted to a sector in which each spin chain has at most one spin flipped in a background of spins in the $|0\rangle$ state. We will thus use a similar notation as before for spin chain states, namely $|\mathbf{0}\rangle^{(I)}$ and $|\mathbf{0}\rangle^{(II)}$ denoting the ferromagnetic ground states (all spins in the $|0\rangle$ state) of the chains I and II, respectively, and $|\mathbf{j}\rangle^{(I)}$ and $|\mathbf{j}\rangle^{(II)}$ denoting the jth spin flipped to the $|1\rangle$ state in the chains I and II, respectively. For the moment, we assume the chains to be non-interacting, identical copies of each other and coupled by uniform Heisenberg or XY interactions. The first spin of each chain is controlled by Alice, while the Nth spin of each chain is controlled by Bob. Initially, the spin chains are assumed to be in the states $|\mathbf{0}\rangle^{(I)}$ and $|\mathbf{0}\rangle^{(II)}$. When Alice intends to transmit a qubit state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$, she encodes this into the two spins that she controls as $|\tilde{\phi}\rangle = \alpha|01\rangle + \beta|10\rangle$. This encoding can be accomplished by a simple two-qubit quantum gate [37] involving the qubits that Alice controls and can be accomplished inside the quantum computer. This encoding places the entire system of two spin chains in the quantum state

$$|\Phi(0)\rangle = \alpha|\mathbf{0}\rangle^{(I)}|\mathbf{1}\rangle^{(II)} + \beta|\mathbf{1}\rangle^{(I)}|\mathbf{0}\rangle^{(II)}, \tag{24}$$

which evolves with time as

$$|\Phi(t)\rangle = \sum_{j=1}^{N} f_{1j}(t)(\alpha|\mathbf{0}\rangle^{(I)}|\mathbf{j}\rangle^{(II)} + \beta|\mathbf{j}\rangle^{(I)}|\mathbf{0}\rangle^{(II)}). \tag{25}$$

The time varying density operator $\varrho(t)$ of the two spins which Bob controls is found to be

$$\varrho(t) = (1 - |f_{1N}(t)|^2)|00\rangle\langle 00| + |f_{1N}(t)|^2|\tilde{\phi}\rangle\langle\tilde{\phi}|. \tag{26}$$

Bob now measures the "total" spin component of his spins in the z direction *without* measuring any of the spins individually. Such a measurement gives a value -1 for $|00\rangle$ and the value 0 for any superposition of $|01\rangle$ and $|10\rangle$. When Bob obtains the outcome 0, which happens with probability $|f_{1N}(t)|^2$, his spins are projected to the state $|\tilde{\phi}\rangle$. Bob can now simply apply the inverse of the quantum gate that Alice used for encoding to his spins to obtain a decoded state $|\phi\rangle$, which corresponds to his perfect retrieval of the state transmitted by Alice. For long Heisenberg and XY chains, thus, Bob's success probability in this heralded scheme scales as $|f_{1N}(t)|^2 \sim 1/N^{2/3}$ in a time $t \sim O(N/J)$. Interestingly, if one was willing to wait till a time $O(N^{1.67}/J)$ (a polynomial scaling with N), then the protocol could be modified (essentially to one in which Bob simply repeats his measurement again and again

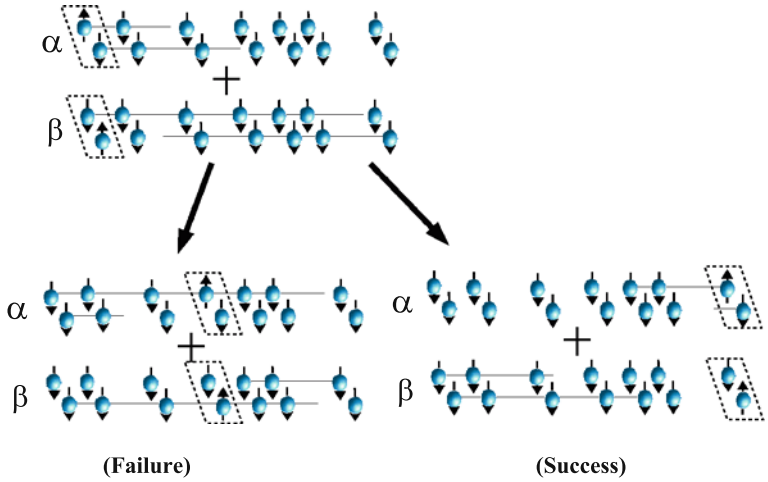


Fig. 6 The dual-rail protocol for perfect quantum communications through spin chain channels. The upper part of the figure shows an arbitrary superposition of the logical qubit states being encoded at one end of the parallel chains (the dotted box shows the encoded states of a qubit). The lower parts of the figure show two possible states of the chain after the passage of some time and Bob's measurement. Bob's success corresponds to the state being received perfectly on the spins at his end of the chain. His failure corresponds to the information being retained in parts of the chain not accessed by Bob, as shown in the lower left hand side of the figure. Strictly speaking the state corresponding to failure will be a superposition of all states of the form shown in the lower left hand side, i.e., it will be a superposition of all those states in which the dotted box is at sites other than N.

on successive failures without resetting the chain) so as to obtain the state with a probability of success arbitrarily close to unity [37].

It is worth noting that it is possible to use a single chain of higher dimensional quantum systems, such as qutrits (quantum three level systems) with levels $|+1\rangle$, $|0\rangle$ and $|-1\rangle$, instead of two parallel chains [37]. For our protocol, the qutrits should be coupled by the natural generalization of an exchange (or isotropic Heisenberg) interaction to higher dimensions given by a Hamiltonian $H = \sum_i P_{i,i+1}$ where

$$P_{i,j}|\psi\rangle_i|\phi\rangle_j = |\phi\rangle_i|\psi\rangle_j. \tag{27}$$

From this state, one generates the states $|+\mathbf{j}\rangle$ and $|-\mathbf{j}\rangle$ of the chain in which the jth qutrit is flipped to the $|+1\rangle$ and $|-1\rangle$ states respectively. Then the dual-rail protocol described in the previous section can be exactly adapted to the chain of qutrits with the mappings $|\mathbf{0}\rangle^{(I)}|\mathbf{0}\rangle^{(II)} \rightarrow |\mathbf{0}\rangle$,$|\mathbf{0}\rangle^{(I)}|\mathbf{j}\rangle^{(II)} \rightarrow |+\mathbf{j}\rangle$ and $|\mathbf{j}\rangle^{(I)}|\mathbf{0}\rangle^{(II)} \rightarrow |-\mathbf{j}\rangle$. One can check that Bob's measurement will now be mapped to a measurement which finds out whether his qutrit is in the state $|0\rangle$ or not (without ascertaining whether the qutrit is in the state $|+1\rangle$ or $|-1\rangle$) and success is when he obtains the result "not $|0\rangle$." In a similar manner, if one had exchange coupled $d+1$ level systems, one could

use one of those levels as the $|0\rangle$ state, and use the others to transmit a d-dimensional system perfectly with 0.99 probability of success in a time $\sim O(N^{1.67}/J)$ through the chain.

5.7 Controllable Coupling of a Memory Qubit at Bob's End

Suppose Bob possessed a memory qubit which could be controllably coupled to the spin at his end of the chain. Then he can receive a single qubit transmitted by Alice with near perfect fidelity in a timescale of the order $\sim O(N^{1.67}/J)$ [39]. Bob lets this memory qubit interact with the spin at his end of the chain at regular intervals, but for different durations of time during each interaction. These time durations are so chosen that the entire amplitude of Bob's spin to be in the $|1\rangle$ state is transferred to the memory qubit. In this way, the chain will finally be left in the state $|\mathbf{0}\rangle$ (all information erased), and the memory qubit will end up in the state $\alpha|0\rangle + \beta|1\rangle$. One positive feature of this scheme is that the memory qubit may itself be a part of the spin chain, say an extra $N + 1$th spin attached to the Nth spin of the chain, with its interaction with the chain being switchable through a local magnetic field [39]. This scheme is essentially a simplification of an earlier scheme where a larger (multi-qubit) memory is required to be held by Bob [40].

The above are by no means exhaustive descriptions of possible schemes of using spin chains as perfect channels for communicating quantum information. However, before closing this lecture, I would like to point to the reader that this is not the only interesting facet of studying information transfer by spin chains or by quantum many-body systems in general. Another interesting aspect, and perhaps more so, is to study how different condensed matter systems respond to the input of quantum information and indeed how well they transfer it. The quality and mode of transfer of quantum information through the system is then just another property of such a system in the same manner as its susceptibility and other response functions are. In some sense, this area of study can be classified as one where one studies the "quantum response" (the state as a function of time) to the "quantum stimulus" (placement of a quantum state at some point on a spin chain). In this context, transfer through antiferromagnets [41–43] and chains of higher dimensional spins [41, 44] have both been studied. In the next lecture we proceed to describe the use of spin chains for quantum computation.

6 Lecture 5

In this lecture we approach the ultimate aim, namely to investigate the possibility of quantum computation using many-body systems. There are quite a few celebrated approaches to this aim which I will not cover in this review. One such approach suggests the use of two-dimensional condensed matter systems with anyonic excitations for fault tolerant quantum computation [45–47]. Another approach suggests using

measurements on certain highly entangled states of many-body systems (could be a state prepared dynamically) [48–51] to perform quantum computation—this is, in fact, a completely different approach to quantum computation called measurement-based quantum computation. The reader interested in these topics should refer the above references. Instead, I will focus on using a given many-body system, such as a one-dimensional magnet, for quantum computation with its spins (or collections of its spins) as qubits and performing sequences of logic gates on these qubits to accomplish quantum computation.

6.1 Quantum Computation with a One-Dimensional Heisenberg Chain

There are two main problems when one wants to use a many-body system as a quantum computer. First, its interactions are generally not tunable—constant in time, and if we want to use its individual spins as qubits, their states will be continuously affected by the interactions with their neighbors. If you have read the earlier lectures, you know that a single qubit placed at one end of certain classes of spin chains, for example, does not stay there, but drifts to the other end—in fact, this could even be used for communications! Clearly thus it is difficult to have isolated qubits on which local unitaries can be performed, or isolated pairs of qubits on which two-qubit gates can be performed. Secondly, it may be difficult to distinctly address a single qubit or even a few qubits—in the ultimate limit only a global addressing of the whole chain may be possible. In Ref.[52] a method of avoiding both the problems was suggested for an uniformly coupled nearest-neighbor isotropic Heisenberg chain of N spins given by

$$\mathbf{H} = \sum_{i=1}^{N} J\ \boldsymbol{\sigma}^i \cdot \boldsymbol{\sigma}^{i+1} - \sum_{i=1}^{N} B_i \sigma_z^i$$

in a magnetic field which can potentially vary from spin to spin (is B_i for the ith spin and for simplicity we have assumed an open-ended chain, though this is not compulsory). The above model is the particular context in which we will now discuss how to use a quantum many-body system as a quantum computer. To keep matters simple, we will not touch upon the second problem with many-body systems, namely, the fact that individual spins, being closely packed, may not be separately addressable (though this can also be circumvented, as the reader may check in Ref.[52]).

Suppose we want to use the individual spins of this system as qubits. Then we will need to show you how to keep such qubits isolated from each other when only single-qubit gates (or no gates) are being performed on them. Additionally we need to show how only a pair of them can be made to talk to each other when we want a two-qubit gate between them, while the pair is kept isolated from the rest. We will additionally assume the ability to apply fields (more specifically, magnetic fields, as we are dealing with spins) to an individual spin. We first discuss how to initialize

such a spin chain as a quantum computer. This involves both an initialization of the Hamiltonian to a desired form, where the isolation mentioned above becomes possible, as well as an initialization of the state of the chain in a certain form for the required isolation. To initialize the Hamiltonian to a desired form, we first apply a magnetic field with B_i having a value $B_{\rm odd}$ for $i = 1, 3, 5, \ldots$, (odd numbered spins) and $B_{\rm even}$ for $i = 2, 4, 6, \ldots$ (even numbered spins). If $|B_{\rm odd} - B_{\rm even}|$ is large compared to J, then, due to mismatching of the Zeeman energies of nearest-neighbor spins, the Hamiltonian undergoes the following Heisenberg to Ising transformation (for a proof see [53]):

$$\sum_{i=1}^{N} J\, \boldsymbol{\sigma}^i \cdot \boldsymbol{\sigma}^{i+1} - \sum_{i=\rm odd}^{N} B_{\rm odd}\sigma_z^i - \sum_{i=\rm even}^{N} B_{\rm even}\sigma_z^i \rightarrow \sum_{i=1}^{N} J\, \sigma_z^i \sigma_z^{i+1}. \qquad (28)$$

Next is the initialization of the spin chain to an ordered state, which is the state at the very start of any computation. This is readily done for J both positive and negative for the above Ising chain, by simply cooling the system to its ground state, where it either goes to the antiferromagnetic Neel state in which alternate spins face opposite to each other or the all-aligned ferromagnetic state. In both these cases, we identify alternate spins as qubits (labeled by X, Y, Z, etc. in the Fig. 7) and *barrier spins* (labeled as B in Fig. 7). This initialized state is shown in Fig. 7(a) for the case of a ferromagnet. Thus all the barrier spins will be aligned to each other at the start. Without the loss of generality, one can assume any one of $B_{\rm odd}$ or $B_{\rm even}$ to be $2J$ and suppose that this magnetic field $2J$ is the one on the location of the qubits X, Y, Z,

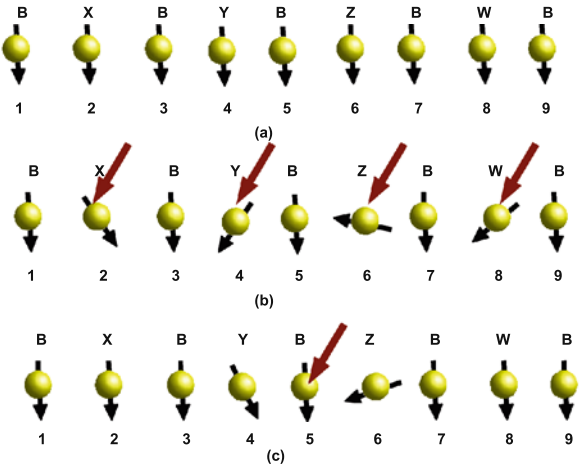


Fig. 7 The figure shows one way of using the spins of a Heisenberg spin chain as qubits for quantum computation. Part (**a**) shows the arrangement of qubits labeled X, Y, Z, etc. and barrier spins B between them. Part (**b**) shows the mechanism of performing single-qubit gates by local fields on the qubits. Part (**c**) shows how one can make a two-qubit gate between qubits Y and Z by tuning the field on the barrier spin between them

etc. In the initialized state, thus, each qubit faces an effective constant magnetic field $-2J$ from its neighboring barrier spins due to the Ising interaction (taking the down direction to correspond to $\sigma_z = -1$), which is exactly canceled by the magnetic field $2J$ on it. Each qubit is thus completely isolated from the effects of its neighbors. Figure 7(b) shows the situation when single-qubit gates are being performed on the qubits. This can simply be done by applying arbitrary local magnetic fields on the qubits X, Y, Z, etc. (these fields have to be "in addition" to the constant magnetic field of $2J$ which plays a central role in the isolation of the qubit from its neighbors).

Having demonstrated the possibility of arbitrary single-qubit gates, we must now demonstrate that entangling two-qubit gates should be possible between any two qubits of the chain in order to prove the feasibility of universal quantum computation with our system. For this, we only require to show that two-qubit entangling gates are possible between *neighboring* qubits. This is because once an entangling two-qubit gate between neighboring qubits is possible, it can, with additional local operations on the qubits, be used to swap the states of the qubits. Thus states of qubits far from each other, such as qubits X and W of Fig. 7 can be brought to neighboring positions Y and Z by repeated swaps for a required two-qubit gate and taken back to their original positions after the gate operation through another series of swaps. Figure 7(c) shows how a two-qubit entangling gate can be accomplished on the neighboring qubits Y and Z. The magnetic field on the barrier spin between Y and Z is suddenly tuned to the value J. Then the three spin segment Y, Z and the barrier between them become equivalent to a three spin open-ended Heisenberg chain in a constant magnetic field J. Evolving such a system for $J = 1$, we get that at time $t_g = \pi/3$ (let us call it "gate time"), the following gate occurs between qubits Y and Z (we leave it to the reader to explicitly verify this by evolving a three spin finite Heisenberg chain):

$$\begin{pmatrix} e^{\mathrm{i}t_g} & 0 & 0 & 0 \\ 0 & -e^{\mathrm{i}t_g}/2 & e^{\mathrm{i}t_g}/2 - 1 & 0 \\ 0 & e^{\mathrm{i}t_g}/2 - 1 & -e^{\mathrm{i}t_g}/2 & 0 \\ 0 & 0 & 0 & e^{-\mathrm{i}t_g} \end{pmatrix}.$$

To be more specific, at the gate time t_g, the barrier spin between qubits Y and Z returns to its original pure state, i.e., decoupled from the state of Y and Z so that it is possible to ascribe a unitary evolution U_g that connects the initial states of qubits Y and Z to their state at t_g. Note that indeed this is an *entangling gate* because if the qubits Y and Z started in the state $|01\rangle$, they would evolve to the entangled state $-e^{\mathrm{i}t_g}/2|01\rangle + (e^{\mathrm{i}t_g}/2 - 1)|10\rangle$.

The above description is a method for accomplishing quantum computation in a one-dimensional spin chain according to a proposal of Simon Benjamin and myself a few years ago [52, 53]. In those papers, the reader will also find how the same protocol can be adapted to a setting where all the qubits of the chain are only globally accessible (a genuine possibility for a many-body system). Benjamin has subsequently demonstrated that a two-dimensional generalization of the protocol

is also possible [54], which can make quantum error correction feasible for such computers. Crucially in the protocol described above, we needed only one barrier spin to separate the qubits as our interactions were only nearest neighbor. Similar protocols have now been shown to be possible for longer range interactions as well [55]. Additionally, the possibility of a similar scheme, even when the couplings of the chain are random, have also been demonstrated [56].

6.2 Logic Gates Through Engineered XY Spin Chains

We now ask the question that whether it is possible for a quantum gate to occur between widely separated qubits such as the qubits at the opposite ends of a spin chain. Indeed such a thing has been shown to be possible [57], but only for the engineered XY spin chains discussed in Lecture 4. These spin chains, as was discussed, mirror invert their states due to their natural time evolution. To give an example, a state $|000000\rangle$ and $|100001\rangle$ of a six spin chain will remain unchanged, while the states $|100000\rangle$ and $|000001\rangle$ will interchange. However, we want much more than simply mirror inversion — we want a quantum gate between the qubits 1 and 6.

For this purpose, let us delve into an important feature of the XY model, namely, that it is equivalent to a model of free fermions hopping in a lattice. These fermions, called Jordan–Wigner fermions [12], are described by operators $c_k^\dagger$ which create a ferm-ion in the kth site of the lattice. They can be written in terms of the raising operator $\sigma_+^k = \sigma_x^k + \mathrm{i}\sigma_y^k$ of the kth spin and a string of σ_z^j operators with $j = 1, \ldots, k-1$, but their specific form, to be found in several condensed matter textbooks [12], need not concern us here. We only need to note that raising a spin from the $|0\rangle$ to $|1\rangle$ at the kth site is equivalent (apart from some phase factors) to the action of the operator $c_k^\dagger$ on the initial state, i.e., equivalent to the creation of a fermion at the kth site. Thus the states of the six spin chain in the above paragraph can be interpreted as the zero fermion (or vacuum) state $|000000\rangle$, one fermion states $|100000\rangle \equiv c_1^\dagger|000000\rangle$ and $|000001\rangle \equiv c_6^\dagger|000000\rangle$, and the two fermion state $|100001\rangle = c_1^\dagger c_6^\dagger|000000\rangle$. Now the engineered spin chain discussed in Lecture 4, which accomplishes mirror inversion, will map $c_1^\dagger c_6^\dagger|000000\rangle$ back to itself after the period of inversion except for a sign change due to the exchange of the two fermions. In other words, the state $|100001\rangle$ will evolve to $-|100001\rangle$ in a certain time T due to the exchange of fermions in a two fermion state. The zero and one fermion states, on the other hand, will have no such sign change, but the pair of one fermion states will transform to each other. The gate acting on the qubits 1 and 6 is then

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

which is an entangling gate, which when appended with local operations on these qubits enables any quantum gate on them. More recent work has shown that such

engineered chains even enable multi-qubit gates involving all the qubits of the chain [58], in particular a multi-qubit gate relevant to the quantum Fourier transform. Before ending the section on quantum computation using spin chain systems, I would like to point out that for Ising spin chains, global pulses could be used for universal quantum computation [34]. Moreover, spin systems with certain couplings can accomplish automata such as those for single-spin measurement [59].

To summarize, through the above set of lectures, I have tried to present an overview of the use of spin chains for quantum information applications. The work in the field is much more than I have been able to cover in these lectures and the lectures have mostly been based on the work with which I have personally been involved. However, I hope that the reader will have obtained a flavor of the activities in this rapidly growing field and will be inspired to contribute to it.

Acknowledgments I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) UK for an Advanced Research Fellowship and for support through the Quantum Information Processing Interdisciplinary Research Collaboration (QIPIRC) grant GR/S82176/01 and the support of the Royal Society and the Wolfson foundation.

References

1. M.A. Nielsen and I.L. Chuang: *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000) 97
2. C.H. Bennett and D.P. DiVincenzo: Nature (London) **404**, 247 (2000) 97
3. M.B. Plenio and V. Vedral: Contemp. Phys. **39**, 431 (1998) 99
4. C.H. Bennett et al.: Phys. Rev. Lett. **70**, 1895 (1993) 99
5. W.K. Wootters: Phys. Rev. Lett. **80**, 2245 (1998) 99, 100, 112
6. S. Popescu and D. Rohrlich: Phys. Rev. A **56**, R3319 (1997) 100
7. M.J. Bremner et al: Phys. Rev. Lett. **89**, 247902 (2002) 101
8. A. Osterloh, L. Amico, R. Fazio, and V. Vedral: arXiv:quant-ph/0703044v1 (2007) 102
9. T. Wang, X. Wang, Z. Sun, arXiv:quant-ph/0607117 (2006) 103
10. M.C. Arnesen, S. Bose, and V. Vedral, Phys. Rev. Lett. **87**, 017901 (2001) 104
11. D. Bruss, N. Datta, A. Ekert, L.C. Kwek, C. Macchiavello: Phys. Rev. A **72**, 014301 (2005) 104
12. S. Sachdev: *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2001) 104, 122
13. A. Osterloh, L. Amico, G. Falci, and R. Fazio: Nature **416**, 608 (2002) 105
14. T.J. Osborne and M.A. Nielsen, Phys. Rev. A **66**, 032110 (2002) 105
15. L. Campos Venuti, C. Degli Esposti Boschi and M. Roncaglia, Phys. Rev. Lett. **99**, 060401 (2007) 105, 106
16. G. Vidal, J.I. Latorre, E. Rico, A. Kitaev: Phys. Rev. Lett. **90**, 227902 (2003) 106
17. V.E. Korepin: Phys. Rev. Lett. **92**, 096402 (2004) 106
18. P. Calabrese and J.L. Cardy: J. Stat. Mech. P06002 (2004) 106
19. H. Fan, V. Korepin, and V. Roychowdhury: Phys. Rev. Lett. **93**, 227203 (2004) 106
20. F. Verstraete, M. Popp, and J.I. Cirac: Phys. Rev. Lett. **92**, 027901 (2004) 106
21. C. Hadley and S. Bose: arXiv:quant-ph/0602139v4 (2006) 106
22. K. Audenaert, J. Eisert, M.B. Plenio, and R.F. Werner: Phys.Rev.A **66**, 042327 (2002) 107
23. J. Anders and A. Winter: arXiv:0705.3026v1 (2007) 107
24. S. Bose: Phys. Rev. Lett. **91**, 207901(2003) 107
25. C.H. Bennett et al.: Phys. Rev. A **54**, 3824 (1996) 112
26. M. Horodecki, P. Horodecki, and R. Horodecki: Phys. Rev. Lett. **78**, 574 (1997) 112
27. M. Christandl, N. Datta, A. Ekert and A.J. Landahl: Phys. Rev. Lett. **92**, 187902 (2004) 113, 114

28. C. Albanese, M. Christandl, N. Datta and A. Ekert: Phys. Rev. Lett. **93**, 230502 (2004) 113
29. T.J. Osborne and N. Linden: Phys. Rev. A **69**, 052315 (2004) 113, 114
30. H.L. Haselgrove: Phys. Rev. A **72**, 062326 (2005) 114
31. Y. Li, T. Shi, B. Chen, Z. Song, and C.-P. Sun: Phys. Rev. A **71**, 022301 (2005) 114
32. M.B. Plenio and F.L. Semiao: New J. Phys. **7**, 73 (2005) 114
33. A. Wojcik, T. Luczak, P. Kurzynski, A. Grudka, T. Gdala, M. Bednarska: Phys. Rev. A **75**, 022330 (2007) 114
34. J. Fitzsimons and J. Twamley: Phys. Rev. Lett. **97**, 090502 (2006) 115, 123
35. T. Ohshima, A. Ekert, D.K.L. Oi, D. Kaslizowski, and L.C. Kwek: quant-ph/0702019 115
36. K. Eckert, O. Romero-Isart and A. Sanpera: New J. Phys. **9**, 155 (2007) 115
37. D. Burgarth and S. Bose: Phys. Rev. A **71**, 052315 (2005) 115, 116, 117
38. D. Burgarth and S. Bose: New J. Phys. **7**, 135 (2005) 115
39. D. Burgarth, V. Giovannetti and S. Bose: Phys. Rev. A **75**, 062327 (2007) 118
40. V. Giovannetti and D. Burgarth: Phys. Rev. Lett. **96**, 030501 (2006) 118
41. O. Romero-Isart, K. Eckert, and A. Sanpera: Phys. Rev. A **75**, 050303(R) (2007) 118
42. L. Campos Venuti, C. Degli Esposti Boschi, M. Roncaglia: Phys. Rev. Lett. **99**, 060401 (2007). 118
43. A. Bayat and S. Bose: arXiv:0706.4176v2 [quant-ph] (2007) 118
44. A. Bayat and V. Karimipour: Phys. Rev. A **75**, 022321 (2007) 118
45. A. Kitaev: Ann. Physics **303**, 2 (2004) 118
46. M. Freedman, C. Nayak, K. Shtengel: Phys. Rev. Lett. **94**, 066401 (2005) 118
47. J. K. Pachos: Int. J. Quantum Inf. **4**, 947 (2006) 118
48. R. Raussendorf and H.J. Briegel: Phys. Rev. Lett. **86**, 5188 (2001) 119
49. F. Verstraete and J.I. Cirac: Phys. Rev. A **70**, 060302(R) (2004) 119
50. D. Gross and J. Eisert: Phys. Rev. Lett. **98**, 220503 (2007) 119
51. S.D. Bartlett and T. Rudolph: Phys. Rev. A **74**, 040302(R) (2006) 119
52. S.C. Benjamin and S. Bose: Phys. Rev. Lett. **90**, 247901 (2003) 119, 121
53. S.C. Benjamin and S. Bose: Phys. Rev. A **70**, 032314 (2004) 120, 121
54. S.C. Benjamin: New J. Phys. **6** 61 (2004) 122
55. Y. Hu, Z.-W. Zhou and G.-C. Guo: New J. Phys. **9**, 27 (2007) 122
56. C.F. Lee and N.F. Johnson: Phys. Rev. A **70**, 052322 (2004) 122
57. M.-H. Yung and S. Bose: Phys. Rev. A **71**, 032310 (2005) 122
58. M.-H. Yung, S.C. Benjamin and S. Bose: Phys. Rev. Lett. **96**, 220501 (2006) 123
59. A. Kay: Phys. Rev. Lett. **98**, 010501 (2007) 123

Non-Markovian Quantum Dynamics and the Method of Correlated Projection Super-Operators

Heinz-Peter Breuer

1 Introduction

Relaxation and decoherence processes are key features of the dynamics of open quantum systems [1]. In the standard approach one tries to develop appropriate master equations for the open system's reduced density matrix ρ_S which is given by the partial trace taken over the environmental variables coupled to the open system. Invoking the weak-coupling assumption one can formulate in many cases of physical interest a Markovian quantum master equation for the reduced density matrix, expressing the dynamical laws for the irreversible motion of the open system.

However, the theoretical description of quantum mechanical relaxation and decoherence processes often leads to a non-Markovian dynamics which is determined by pronounced memory effects. Strong system–environment couplings [2, 3], correlations and entanglement in the initial state [4, 5], interactions with environments at low temperatures and with spin baths [6], finite reservoirs [7, 8], and transport processes in nano-structures [9] can lead to long memory times and to a failure of the Markovian approximation.

Here, we will review the most important features of a systematic approach to non-Markovian quantum dynamics which is known as projection operator technique [10–13]. This technique is based on the introduction of a certain projection super-operator $\mathcal{P}$ which acts on the states of the total system. The super-operator $\mathcal{P}$ expresses in a formal mathematical way the idea of the elimination of degrees of freedom from the complete description of the states of the total system. Namely, if ρ is the full density matrix of the composite system, the projection $\mathcal{P}\rho$ serves to represent a certain approximation of ρ which leads to a simplified effective description of the dynamics through a reduced set of relevant variables.

With the help of the projection operator techniques one derives closed dynamic equations for the relevant variables $\mathcal{P}\rho$. We will discuss two different approximation schemes. The first one is based on the Nakajima–Zwanzig equation [10, 11] which

H.-P. Breuer (✉)
Physikalisches Institut, Universität Freiburg, Freiburg, Germany
e-mail: breuer@physik.uni-freiburg.de

Breuer, H.-P.: *Non-Markovian Quantum Dynamics and the Method of Correlated Projection Super-Operators*. Lect. Notes Phys. **787**, 125–139 (2010)
DOI 10.1007/978-3-642-02871-7_5 © Springer-Verlag Berlin Heidelberg 2010

represents an integro-differential equation for $\mathcal{P}\rho$ with a certain memory kernel. The second scheme employs a time-convolutionless master equation for $\mathcal{P}\rho$, i.e., a time-local differential equation with a time-dependent generator [14–19]. These equations are used as starting point for the derivation of effective master equations through a systematic perturbation expansion. In the standard approach to the dynamics of open systems one chooses a projection super-operator which is defined by the expression $\mathcal{P}\rho = \rho_S \otimes \rho_0$, where ρ_0 is some fixed environmental state. A super-operator of this form projects the total state ρ onto a tensor product state, i.e., onto a state without any statistical correlations between system and environment. Many examples for this product-state projection are known in the fields of quantum optics, decoherence, quantum Brownian motion, quantum measurement theory, and coherent and optimal quantum control. It is typically applicable in the case of weak system–environment couplings. The corresponding perturbation expansion is usually restricted to the second order (known as Born approximation), from which one derives, with the help of certain further assumptions, a Markovian quantum master equations in Lindblad form [20–22].

A possible approach to large deviations from Markovian behavior consists in carrying out the perturbation expansion to higher orders in the system–environment coupling. However, this approach is often limited by the increasing complexity of the resulting equations of motion. Moreover, the perturbation expansion may not converge uniformly in time, such that higher orders only improve the quality of the approximation of the short-time behavior, but completely fail in the long-time limit [23].

We will discuss here a further strategy for the treatment of highly non-Markovian processes which is based on the use of a correlated projection super-operator [24–29]. By contrast to the product-state projection, a correlated projection super-operator projects the total state ρ onto a system–environment state that contains statistical correlations between certain system and environment states. We will discuss a representation theorem for a large class of such projections, which are appropriate for the application of the projection operator techniques, and develop a corresponding non-Markovian generalization of the Lindblad equation.

2 The Standard Projection Operator Method

We investigate an open quantum system S that is coupled to some environment E. The corresponding Hilbert spaces are denoted by $\mathcal{H}_S$ and $\mathcal{H}_E$, respectively. The state space of the composite system is thus given by the tensor product space

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E. \tag{1}$$

The states of the composite system are represented by density matrices ρ on $\mathcal{H}$ satisfying the physical conditions of the positivity and the normalization:

$$\rho \geq 0, \qquad \mathrm{tr}\rho = 1, \tag{2}$$

where tr denotes the trace taken over the total state space $\mathcal{H}$. The partial traces over $\mathcal{H}_S$ and $\mathcal{H}_E$ will be denoted by tr_S and tr_E.

2.1 Nakajima–Zwanzig Projection Operator Technique

A central goal of the theory is to develop efficient strategies for the description of the behavior of the reduced density matrix which is determined by the partial trace over the environmental state space,

$$\rho_S = \mathrm{tr}_E \rho. \tag{3}$$

The basic idea of the projection operator techniques is to regard the operation of taking the partial trace over E formally as a map $\mathcal{P}$ defined by

$$\mathcal{P}\rho = (\mathrm{tr}_E \rho) \otimes \rho_0. \tag{4}$$

For a fixed environmental state ρ_0 this defines a linear transformation which maps any density operator ρ on the total state space $\mathcal{H}$ to a density operator $\mathcal{P}\rho$ on $\mathcal{H}$ and has the property of a projection operator:

$$\mathcal{P}^2 = \mathcal{P}. \tag{5}$$

Being a map acting on operators, $\mathcal{P}$ is often called a projection super-operator. The complementary projection is defined by

$$\mathcal{Q} = I - \mathcal{P}, \tag{6}$$

I being the identity map. Note that according to Eq. (4) the state of the reduced system is obtained from the projection $\mathcal{P}\rho$ by taking the partial trace over the environment:

$$\rho_S = \mathrm{tr}_E\{\mathcal{P}\rho\}. \tag{7}$$

The Hamiltonian of the composite system is of the form

$$H = H_0 + H_I, \tag{8}$$

where H_0 denotes the unperturbed part, usually given by the sum of a system Hamiltonian H_S and an environmental Hamiltonian H_E, and H_I represents the interaction. In many cases it is convenient to formulate the dynamics in the interaction picture with respect to H_0 in which the density matrix $\rho(t)$ of the total system is governed by the von Neumann equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = -\mathrm{i}[H_I(t), \rho(t)] \equiv \mathcal{L}(t)\rho(t). \tag{9}$$

The operator

$$H_I(t) = \mathrm{e}^{\mathrm{i}H_0 t} H_I \mathrm{e}^{-\mathrm{i}H_0 t} \tag{10}$$

represents the Hamiltonian in the interaction picture and $\mathcal{L}(t)$ the corresponding Liouville super-operator.

The Nakajima–Zwanzig (NZ) projection operator technique yields a closed equation of motion for the relevant part $\mathcal{P}\rho(t)$ of the density matrix and, hence, for the reduced density matrix $\rho_S(t)$. To simplify the presentation we assume that the condition

$$\mathcal{P}\mathcal{L}(t_1)\mathcal{L}(t_2)\ldots\mathcal{L}(t_{2n+1})\mathcal{P} = 0 \tag{11}$$

holds true. This condition is in fact satisfied in many applications. Moreover, we suppose that the initial state satisfies $\mathcal{P}\rho(0) = \rho(0)$. The projection $\mathcal{P}\rho(t)$ is then governed by a homogeneous integro-differential equation, the Nakajima–Zwanzig equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho(t) = \int_0^t \mathrm{d}t_1 \mathcal{K}(t, t_1)\mathcal{P}\rho(t_1). \tag{12}$$

The memory kernel $\mathcal{K}(t, t_1)$ is given by

$$\mathcal{K}(t, t_1) = \mathcal{P}\mathcal{L}(t)\,\mathrm{T}\exp\left[\int_{t_1}^t \mathrm{d}t_2\,\mathcal{Q}\mathcal{L}(t_2)\right]\mathcal{Q}\mathcal{L}(t_1)\mathcal{P}, \tag{13}$$

where T denotes the chronological time ordering.

The memory kernel $\mathcal{K}(t, t_1)$ is in general a very complicated super-operator whose determination is in most cases as complicated as the solution of the dynamics of the full system. Therefore, one usually tries to determine it by a perturbation expansion in powers of the strength of the system–environment coupling. The lowest order contribution is given by the second-order equation of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho(t) = \int_0^t \mathrm{d}t_1 \mathcal{P}\mathcal{L}(t)\mathcal{L}(t_1)\mathcal{P}\rho(t_1). \tag{14}$$

Higher orders are obtained with the help of the general expression (13) for the memory kernel.

2.2 Time-Convolutionless Projection Operator Technique

There exists an alternative expansion technique based on the projection super-operator $\mathcal{P}$ which is known as time-convolutionless (TCL) projection operator method. By contrast to the NZ approach, the TCL method leads to an equation of motion for the relevant part of the density matrix which represents a time-local differential equation of the general form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho(t) = \mathcal{K}(t)\mathcal{P}\rho(t). \tag{15}$$

Here, $\mathcal{K}(t)$ is a time-dependent super-operator, called the TCL generator. It should be stressed that the TCL equation (15) describes non-Markovian dynamics, although it is local in time and does not involve an integration over the past of the system. In fact, the TCL equation takes into account all memory effects through the explicit time-dependence of the generator $\mathcal{K}(t)$.

To obtain the time-local form of the TCL equation one eliminates the dependence of the future time evolution on the history of the system through the introduction of the backward propagator into the Nakajima–Zwanzig equation. This enables one to express the density matrix at previous times $t_1 < t$ in terms of the density matrix at time t and to derive an exact time-local equation of motion. We remark that the backward propagator and, hence, also the TCL generator may not exist, typically at isolated points of the time axis. This may happen for very strong system–environment couplings and/or long integration times; an example is discussed in [1].

Again, one can develop a systematic perturbation expansion for the TCL generator which takes the form $\mathcal{K}(t) = \mathcal{K}_2(t) + \mathcal{K}_4(t) + \cdots$. The various orders of this expansion can be expressed through the ordered cumulants [30–33] of the Liouville super-operator $\mathcal{L}(t)$. For instance, the contributions of second and fourth order to the TCL generator are given by [1]

$$\mathcal{K}_2(t) = \int_0^t \mathrm{d}t_1 \mathcal{P}\mathcal{L}(t)\mathcal{L}(t_1)\mathcal{P},$$

and

$$\begin{aligned}\mathcal{K}_4(t) = \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \\ \times\Big[\mathcal{P}\mathcal{L}(t)\mathcal{L}(t_1)\mathcal{L}(t_2)\mathcal{L}(t_3)\mathcal{P} - \mathcal{P}\mathcal{L}(t)\mathcal{L}(t_1)\mathcal{P}\mathcal{L}(t_2)\mathcal{L}(t_3)\mathcal{P} \\ -\mathcal{P}\mathcal{L}(t)\mathcal{L}(t_2)\mathcal{P}\mathcal{L}(t_1)\mathcal{L}(t_3)\mathcal{P} - \mathcal{P}\mathcal{L}(t)\mathcal{L}(t_3)\mathcal{P}\mathcal{L}(t_1)\mathcal{L}(t_2)\mathcal{P}\Big].\end{aligned}$$

In second order the TCL master equation takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho(t) = \int_0^t \mathrm{d}t_1 \mathcal{P}\mathcal{L}(t)\mathcal{L}(t_1)\mathcal{P}\rho(t), \tag{16}$$

which should be contrasted to the NZ equation (14).

It is important to realize that the NZ and the TCL technique lead to equations of motion with entirely different structures and that, therefore, also the mathematical structure of their solutions are quite different in any given order [34]. It is difficult to formulate general conditions that allow to decide for a given model whether the NZ or the TCL approach is more efficient. The assessment of the quality of the approximation obtained generally requires the investigation of higher orders of the expansion, or else the comparison with numerical simulations or with certain limiting cases that can be treated analytically. It turns out that in many cases the degree of accuracy obtained by both methods are of the same order of magnitude. In these cases the TCL approach is of course to be preferred because it is technically much simpler to deal with.

In the NZ equation (12) as well as in the TCL equation (15) we made use of the initial condition $\mathcal{P}\rho(0) = \rho(0)$. According to the definition (4) of the projection $\mathcal{P}$ this condition is equivalent to the assumption that $\rho(0)$ represents an uncorrelated tensor product initial state, $\rho(0) = \rho_S(0) \otimes \rho_0$. For a correlated initial state one has to add a certain inhomogeneity to the right-hand side of the NZ or the TCL equation which involves the initial conditions through the complementary projection $\mathcal{Q}\rho(0) = (I - \mathcal{P})\rho(0)$. A general method for the treatment of such correlated initial states within the TCL technique is described in [1]; for a recent study on their influence in weakly coupled systems see also Refs. [35, 36].

2.3 Markovian Limit and Quantum Dynamical Semigroups

With the standard projection defined in Eq. (4), the TCL equation (16) is equivalent to the following master equation for the reduced density matrix,

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_S(t) = -\int_0^t \mathrm{d}t_1 \mathrm{tr}_E\{[H_I(t), [H_I(t_1), \rho_S(t) \otimes \rho_0]]\}. \tag{17}$$

This equation provides an appropriate starting point for an approximation scheme which is known as Born–Markov approximation and which eventually leads to a Markovian quantum master equation in Lindblad form:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_S(t) &= \mathcal{K}\rho_S(t) \\ &= -\mathrm{i}\,[H_S, \rho_S(t)] + \sum_\lambda \left(R_\lambda \rho_S(t) R_\lambda^\dagger - \frac{1}{2}\left\{ R_\lambda^\dagger R_\lambda, \rho_S(t) \right\} \right). \end{aligned} \tag{18}$$

Here, $\mathcal{K}$ is a time-independent generator, the Lindblad generator, involving a Hermitian operator H_S and arbitrary system operators R_λ. Therefore, it generates state transformations of the form

$$\Phi_t : \rho_S(0) \mapsto \rho_S(t), \qquad \Phi_t = \mathrm{e}^{\mathcal{K}t}. \tag{19}$$

Φ_t is called a quantum dynamical map and the set of transformations

$$\{\Phi_t | t \geq 0\}$$

is referred to as quantum dynamical semigroup. Under certain technical conditions, it can be shown that the form of the Lindblad generator guarantees the preservation of the positivity and normalization of the density matrix, as well as the complete positivity of the dynamical transformation Φ_t. Vice versa, any completely positive quantum dynamical semigroup has a generator of the form (18). This is the well-known Gorini–Kossakowski–Sudarshan–Lindblad theorem [20, 21].

The microscopic derivation of the master equation (18) from the TCL equation (17) requires the validity of several approximations, the most important one being the so-called Markov approximation. This approximation presupposes a rapid decay of the two-point correlation functions of those environmental operators that describe the system–environment coupling. More precisely, if τ_E describes the temporal width of these correlations and τ_R the relaxation time of the system, the Markov approximation demands that

$$\tau_E \ll \tau_R. \tag{20}$$

This means that the environmental correlation time τ_E is short compared to the relaxation time τ_R of the open system.

The Markov approximation is justified in many cases of physical interest. Examples of application are the quantum optical master equation describing the interaction of radiation with matter and the master equation for a test particle in a quantum gas [37–39]. However, strong couplings or interactions with low-temperature reservoirs can lead to large correlations resulting in long memory times and in a failure of the Markov approximation. In the following, the quantum dynamics of an open system is said to be non-Markovian if the time-evolution of its reduced density matrix cannot be described (to the desired degree of accuracy) by means of a closed master equation with a (possibly time-dependent) generator in Lindblad form.

If the two-point environmental correlation functions do not decay rapidly in time the second order of the expansion cannot, in general, be expected to give an accurate description of the dynamics. For instance, this situation arises for the spin star model discussed in Ref. [23], where the second-order generator of the master equation increases linearly with time such that the Born–Markov approximation simply does not exist.

More importantly, the standard Markov condition (20) alone does *not* guarantee, in general, that the Markovian master equation provides a reasonable description of the dynamics. This situation can occur for finite and/or structured reservoirs that cannot be represented by a Bosonic field or a collection of harmonic oscillator modes. In such cases a detailed investigation of the influence of higher-order correlations is indispensable in order to judge the quality of a given order. The model discussed in Ref. [25] represents an example for which the standard Markov condition

is satisfied although the expansion based on the projection (4) completely fails if one truncates the expansion at any finite order. In such cases strong non-Markovian dynamics is induced through the behavior of higher-order correlation functions.

We conclude that in general one can judge the quality of a given projection super-operator and a given expansion technique that is based on it only by an investigation of the structure of higher orders. The standard projection and the corresponding Lindblad equation are not reliable if higher orders lead to contributions that are not bounded in time, signifying the non-uniform convergence of the perturbation expansion [25].

3 Correlated Projection Super-Operators

The performance of the projection operator techniques depends of course on the properties of the microscopic model under study, in particular on the structure of the correlation functions of the model. However, it also depends strongly on the choice of the super-operator $\mathcal{P}$. Several extensions of the standard projection (4) and modifications of the expansion technique have been proposed in the literature (see, e.g., Refs. [40–42]).

The projection defined by (4) projects any state ρ onto a tensor product $\rho_S \otimes \rho_0$ that describes a state without statistical correlations between the system and its environment. Here, we introduce a more general class of projection super–operators that project onto correlated system–environment states and are therefore able to describe strong correlations and non-Markovian effects [24].

3.1 General Conditions

We assume that our new class of maps $\mathcal{P}$ represent super-operators with the property of a projection, i.e., $\mathcal{P}^2 = \mathcal{P}$. As a consequence, the whole machinery of the projection operator techniques described in Sect. 2 can be applied also to the new class of correlated maps.

Within the projection operator techniques the projection $\mathcal{P}\rho$ should represent a suitable approximation of ρ. We therefore require that for any physical state ρ the projection $\mathcal{P}\rho$ is again a physical state, i.e., a positive operator with unit trace. This means that $\mathcal{P}$ is a positive and trace-preserving map, namely

$$\rho \geq 0 \Longrightarrow \mathcal{P}\rho \geq 0, \qquad \mathrm{tr}\{\mathcal{P}\rho\} = \mathrm{tr}\rho. \tag{21}$$

Our class of projection operators is assumed to consist of maps of the following general form:

$$\mathcal{P} = I_S \otimes \Lambda. \tag{22}$$

Here, I_S denotes the unit map acting on operators on $\mathcal{H}_S$, and Λ is a linear map that transforms operators on $\mathcal{H}_E$ into operators on $\mathcal{H}_E$. A projection super-operator of

this form leaves the system S unchanged and acts nontrivially only on the variables of the environment E. As a consequence of the positivity of $\mathcal{P}$ and of condition (22) the map Λ must be N_S-positive, where N_S is the dimension of $\mathcal{H}_S$. In the following we use the stronger condition that Λ is completely positive, because completely positive maps allow for a simple mathematical characterization (see Sect. 3.2).

Let us discuss the physical implications of these conditions. According to (5) and (22) the map Λ must itself be a projection, namely $\Lambda^2 = \Lambda$. Moreover, since $\mathcal{P}$ is trace-preserving, the map Λ must also be trace-preserving. Hence, we find that Λ represents a completely positive and trace-preserving map (CPT map, or quantum channel) which operates on the variables of the environment and has the property of a projection. The action of $\mathcal{P}$ may also be interpreted as that of a generalized quantum measurement which is carried out on the environment. A further physically reasonable consequence of the positivity of Λ and of (22) is that $\mathcal{P}$ maps product states to product states, and, more generally, separable (classically correlated) states to separable states. This means that the application of $\mathcal{P}$ does not create entanglement between the system and its environment.

Using (22) and the fact that Λ is trace-preserving we get

$$\rho_S \equiv \mathrm{tr}_E \rho = \mathrm{tr}_E\{\mathcal{P}\rho\}. \tag{23}$$

This relation connects the density matrix of the reduced system with the projection of a given state ρ of the total system. It states that, in order to determine ρ_S, we do not really need the full density matrix ρ, but only its projection $\mathcal{P}\rho$. Thus, $\mathcal{P}\rho$ contains the full information needed to reconstruct the reduced system's state.

3.2 Representation Theorem

What is the explicit structure of the projection super-operators satisfying the basic conditions formulated above? This question is answered by a representation theorem [24] which states that $\mathcal{P}$ fulfils the condition of Sect. 3.1 if and only if it can be written in the form

$$\mathcal{P}\rho = \sum_i \mathrm{tr}_E\{A_i \rho\} \otimes B_i, \tag{24}$$

where $\{A_i\}$ and $\{B_i\}$ are two sets of linear independent Hermitian operators satisfying the relations

$$\mathrm{tr}_E\{B_i A_j\} = \delta_{ij}, \tag{25}$$

$$\sum_i (\mathrm{tr}_E B_i) A_i = I_E, \tag{26}$$

$$\sum_i A_i^T \otimes B_i \geq 0. \tag{27}$$

Condition (25) guarantees that $\mathcal{P}$ is a projection super-operator, (26) ensures that $\mathcal{P}$ is trace-preserving, and (27) is equivalent to the condition of complete positivity (T denotes the transposition).

The standard projection (4) that projects onto an uncorrelated tensor product state is obviously of the form of Eq. (24). In fact, if we take a single $A = I_E$ and a single $B = \rho_0$ the conditions (25), (26), and (27) are trivially satisfied and Eq. (24) obviously reduces to Eq. (4). Of course, a projection $\mathcal{P}\rho$ of the form of Eq. (24) does not in general represent a simple product state. We therefore call such $\mathcal{P}$ correlated projection super-operators. They project onto states that contain statistical correlations between the system S and its environment E. In the following we will consider the case that one can find a representation of the projection with positive operators:

$$A_i \geq 0, \qquad B_i \geq 0. \tag{28}$$

Equation (27) is then trivially satisfied. Without restriction we may assume that the B_i are normalized to unit trace,

$$\mathrm{tr}_E B_i = 1, \tag{29}$$

such that condition (26) reduces to the simple form

$$\sum_i A_i = I_E. \tag{30}$$

Under these conditions $\mathcal{P}$ projects any state ρ onto a state which can be written as a sum of tensor products of positive operators. In the theory of entanglement (see, e.g., the recent review [43]) such states are called separable or classically correlated. Using a projection super-operator of this form, one thus tries to approximate the total system's states by a classically correlated state. The general representation of (24) includes the case of projection super-operators that project onto inseparable, entangled quantum states. We will not pursue here this possibility further and restrict ourselves to positive A_i and B_i in the following.

3.3 Correlated Initial States

As mentioned already in Sect. 2.2 a homogeneous NZ or TCL equation of motion presupposes a tensor product initial state if one uses the standard projection super-operator (4). However, this is no longer true for the correlated projection defined by (24). In fact, the general condition for the absence of an inhomogeneous term in the NZ equation (12) or the TCL equation (15) is given by

$$\mathcal{P}\rho(0) = \rho(0). \tag{31}$$

According to (24) this condition is equivalent to the assumption that $\rho(0)$ takes the form

$$\rho(0) = \sum_i \rho_i(0) \otimes B_i, \tag{32}$$

where

$$\rho_i(0) = \mathrm{tr}_E\{A_i \rho(0)\} \geq 0. \tag{33}$$

The right-hand side of (32) represents in general a correlated initial state. Hence, a great advantage of the correlated projection super-operators is given by the fact that they allow the treatment of correlated initial states by means of a homogeneous NZ or TCL equation [24].

3.4 Conservation Laws

A crucial step in applications of the correlated projection operator technique is the construction of an appropriate projection super-operator $\mathcal{P}$. An important strategy for this construction is to take into account the known conserved quantities of the model under study.

Suppose C is a conserved observable. A good choice for the projection super-operator $\mathcal{P}$ will then be a projection that leaves invariant the expectation value of C, i.e., that satisfies the relation

$$\mathrm{tr}\{C\rho\} = \mathrm{tr}\{C(\mathcal{P}\rho)\}. \tag{34}$$

To bring this relation into a more convenient form we introduce the adjoint $\mathcal{P}^\dagger$ of the projection super-operator $\mathcal{P}$. The adjoint map is defined with the help of the Hilbert-Schmidt scalar product

$$(X, Y) = \mathrm{tr}\{X^\dagger Y\} \tag{35}$$

for the space of operators acting on the state space of the total system through the relation

$$(X, \mathcal{P}Y) = (\mathcal{P}^\dagger X, Y).$$

This allows us to write Eq. (34) in the form $\mathrm{tr}\{C\rho\} = \mathrm{tr}\{(\mathcal{P}^\dagger C)\rho\}$. Requiring this to hold for all states ρ we get the relation

$$\mathcal{P}^\dagger C = C. \tag{36}$$

The adjoint of the projection (24) is obtained by interchanging the role of A_i and B_i. Hence, condition (36) can be written explicitly as

$$\mathcal{P}^{\dagger}C = \sum_i \mathrm{tr}_E\{B_i C\} \otimes A_i = C. \tag{37}$$

This equation represents a condition for the projection super-operator $\mathcal{P}$ on the basis of a known conserved quantity of the underlying model. It ensures that the projection super-operator leaves invariant this quantity and that the effective description respects the corresponding conservation law.

4 Generalization of the Lindblad Equation

Once a projection super-operator has been chosen the projection $\mathcal{P}\rho(t)$ of the time-dependent total system's state $\rho(t)$ is, according to (24), uniquely determined by the dynamical variables

$$\rho_i(t) = \mathrm{tr}_E\{A_i \rho(t)\}. \tag{38}$$

To be specific we assume in the following that the index i takes on the values $i = 1, 2, \ldots, n$. Since we require that the A_i are positive, the $\rho_i(t)$ are positive operators. From Eq. (23) we find the connection to the reduced density matrix,

$$\rho_S(t) = \sum_i \rho_i(t), \tag{39}$$

and the normalization condition takes the form

$$\mathrm{tr}_S \rho_S(t) = \sum_i \mathrm{tr}_S \rho_i(t) = 1. \tag{40}$$

Hence, we see that the state of the reduced system is uniquely determined by a set of n (unnormalized) density operators $\rho_i(t)$.

Our formulation leads to a natural question, namely what is the analog of the Lindblad equation (18) in the case of a correlated projection super-operator? To answer this question we first observe that the time-evolution leads to a transformation of the form

$$\{\rho_i(0)\} \mapsto \{\rho_i(t)\}, \tag{41}$$

transforming any initial set of positive operators $\rho_i(0) \geq 0$ into another set of positive operators $\rho_i(t) \geq 0$ at time $t > 0$. This transformation can conveniently be described with the help of an auxiliary n-dimensional Hilbert space $\mathbf{C}^n$ and a fixed orthonormal basis $\{|i\rangle\}$ for this space. Then we can identify the collection of densities $\rho_i(t)$ with a density matrix $\varrho(t)$ on the extended space

$$\mathcal{H}_{\mathrm{ext}} = \mathcal{H}_S \otimes \mathbf{C}^n \tag{42}$$

through the relation

$$\varrho(t) = \sum_i \rho_i(t) \otimes |i\rangle\langle i|. \tag{43}$$

This density matrix can be regarded as a block diagonal matrix

$$\varrho(t) = \begin{pmatrix} \rho_1(t) & 0 & \cdots & 0 \\ 0 & \rho_2(t) & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \rho_n(t) \end{pmatrix} \tag{44}$$

with blocks $\rho_i(t)$ along the main diagonal. Moreover, the reduced density matrix $\rho_S(t)$ is obtained by the partial trace of $\varrho(t)$ taken over the auxiliary space.

In close analogy to (19) the dynamics may now be viewed as a transformation

$$V_t : \varrho(0) \mapsto \varrho(t), \tag{45}$$

that preserves the block diagonal structure. It is important to emphasize that V_t is *not* a quantum dynamical map in the usual sense because it is *not* an operation on the space of states of the reduced system, but rather a map on the extended state space. In fact, the transition from $\varrho(0)$ to the reduced density matrix $\rho_S(0) = \sum_i \rho_i(0)$ is connected with a loss of information on the initial correlations, such that from the mere knowledge of $\rho_S(0)$ the dynamical behavior cannot be reconstructed.

It may be shown that V_t can be extended to a completely positive map for operators on $\mathcal{H}_{\text{ext}}$. Hence, we can construct an embedding of the dynamical transformation into a Lindblad dynamics on the extended state space. This is achieved by the requirement that there exist a Lindblad generator $\mathcal{K}$ acting on operators of the extended state space which preserves the block diagonal structure:

$$\mathcal{K}\left(\sum_i \rho_i \otimes |i\rangle\langle i|\right) = \sum_i \mathcal{K}_i(\rho_1, \ldots, \rho_n) \otimes |i\rangle\langle i|, \tag{46}$$

such that the time-evolution can be represented in the form

$$\sum_i \rho_i(t) \otimes |i\rangle\langle i| = \mathrm{e}^{\mathcal{K}t}\left(\sum_i \rho_i(0) \otimes |i\rangle\langle i|\right). \tag{47}$$

One can show that a Lindblad generator $\mathcal{K}$ with this property exists if and only if the densities $\rho_i(t)$ obey the master equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_i(t) = -\mathrm{i}\left[H^i, \rho_i(t)\right] + \sum_{j\lambda}\left(R_\lambda^{ij}\rho_j(t)R_\lambda^{ij\dagger} - \frac{1}{2}\left\{R_\lambda^{ji\dagger}R_\lambda^{ji}, \rho_i(t)\right\}\right). \tag{48}$$

The H^i are Hermitian operators on $\mathcal{H}_S$, while the R_λ^{ij} may be arbitrary operators on $\mathcal{H}_S$. The details of the proof of this statement can be found in Ref. [24].

The equation of motion (48) represents the desired non-Markovian generalization of the Lindblad equation (18) for the case of a classically correlated projection super-operator. This equation has many physical applications. In fact, master equations of the form of Eq. (48) have been derived by several authors and applied to various models featuring pronounced non-Markovian effects [24–29].

5 Conclusions

We have reviewed the theoretical treatment of non-Markovian quantum dynamics within the framework of the projection operator techniques. It has been shown that an efficient description of strong non-Markovian effects is made possible through the construction of correlated projection super-operators $\mathcal{P}$. The central idea behind this construction is to take into account large system–environment correlations by an extension of the set of dynamical variables. In fact, employing a correlated projection super-operator instead of a product-state projection, one enlarges the set of dynamical variables from the reduced density matrix ρ_S to a collection of densities ρ_i describing system states that are correlated with certain environmental states.

General physical conditions for a large class of correlated projection super-operators have been formulated, demanding essentially that $\mathcal{P}$ can be expressed in terms of a projective quantum channel that operates on the environmental variables. These conditions lead to a representation theorem for correlated projection super-operators and to a non-Markovian generalization of the Lindblad equation that is capable of modeling long memory times and large initial correlations, while preserving the physical conditions of positivity and normalization.

The method developed here has many applications to physically relevant models featuring non-Markovian dynamics. The investigated class of projections does not exhaust all possibilities. Future investigations should include the formulation of further classes of correlated projections, the study of time-dependent generators, as well as the application of correlated maps that project onto nonseparable, entangled quantum states.

Acknowledgment I would like to thank Francesco Petruccione, Daniel Burgarth, Jan Fischer, Jochen Gemmer, Mathias Michel, and Bassano Vacchini for many fruitful discussions.

References

1. H.P. Breuer, F. Petruccione: *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2007) 125, 129, 130
2. F. Haake, R. Reibold: Phys. Rev. A **32**, 2462 (1985) 125
3. B.L. Hu, J.P. Paz, Y. Zhang: Phys. Rev. D **45**, 2843 (1992) 125
4. H. Grabert, P. Schramm, G.-L. Ingold: Phys. Rep. **168**, 115 (1988) 125

5. P. Štelmachovič, V. Bužek: Phys. Rev. A **64**, 062106 (2001); Phys. Rev. A **67**, 029902(E) (2003) 125
6. J. Schliemann, A. Khaetskii, D. Loss: J. Phys.: Condens. Matter **15**, R1809 (2003) 125
7. J. Gemmer, M. Michel: Europhys. Lett. **73**, 1 (2006) 125
8. J. Gemmer, M. Michel, G. Mahler: *Quantum Thermodynamics*, Lecture Notes in Physics, vol. 657 (Springer, Berlin, Heidelberg, New York, 2004) 125
9. M. Michel, G. Mahler, J. Gemmer: Phys. Rev. Lett. **95**, 180602 (2005) 125
10. S. Nakajima: Progr. Theor. Phys. **20**, 948 (1958) 125
11. R. Zwanzig: J. Chem. Phys. **33**, 1338 (1960) 125
12. F. Haake: *Statistical Treatment of Open Systems*, Springer Tracts in Modern Physics, vol. 66 (Springer, Berlin, 1973) 125
13. R. Kubo, M. Toda, N. Hashitsume: *Statistical Physics II. Nonequilibrium Statistical Mechanics* (Springer, Berlin, 1991) 125
14. F. Shibata, Y. Takahashi, N. Hashitsume: J. Stat. Phys. **17**, 171 (1977) 126
15. S. Chaturvedi, F. Shibata: Z. Phys. B **35**, 297 (1979) 126
16. F. Shibata, T. Arimitsu: J. Phys. Soc. Jap. **49**, 891 (1980) 126
17. C. Uchiyama, F. Shibata: Phys. Rev. E **60**, 2636 (1999) 126
18. A. Royer: Phys. Lett. A **315**, 335 (2003) 126
19. H.P. Breuer: Phys. Rev. A **70**, 012106 (2004) 126
20. V. Gorini, A. Kossakowski, E.C.G. Sudarshan: J. Math. Phys. **17**, 821 (1976) 126, 131
21. G. Lindblad: Commun. Math. Phys. **48**, 119 (1976) 126, 131
22. H. Spohn: Rev. Mod. Phys. **52**, 569 (1980) 126
23. H.P. Breuer, D. Burgarth, F. Petruccione: Phys. Rev. B **70**, 045323 (2004) 126, 131
24. H.P. Breuer: Phys. Rev. A **75**, 022103 (2007) 126, 132, 133, 135, 138
25. H.P. Breuer, J. Gemmer, M. Michel, Phys. Rev. E **73**, 016139 (2006) 126, 131, 132, 138
26. A.A. Budini: Phys. Rev. A **74**, 053815 (2006) 126, 138
27. A.A. Budini: Phys. Rev. E **72**, 056106 (2005) 126, 138
28. M. Esposito, P. Gaspard: Phys. Rev. E **68**, 066112 (2003) 126, 138
29. M. Esposito, P. Gaspard, Phys. Rev. E **68**, 066113 (2003) 126, 138
30. R. Kubo: J. Math. Phys. **4**, 174 (1963) 129
31. A. Royer: Phys. Rev. A **6**, 1741 (1972) 129
32. N.G. van Kampen: Physica **74**, 215 (1974) 129
33. N.G. van Kampen: Physica **74**, 239 (1974) 129
34. A. Royer: Aspects of Open Quantum Dynamics. In: *Irreversible Quantum Dynamics*, Lecture Notes in Physics, vol. 622, ed by F. Benatti, R. Floreanini (Springer, Berlin, Heidelberg, New York, 2003) pp. 47–63 130
35. S. Tasaki et al.: Ann. Phys. **322**, 631 (2007) 130
36. K. Yuasa et al.: Ann. Phys. **322**, 657 (2007) 130
37. B. Vacchini: Phys. Rev. Lett. **84**, 1374 (2000) 131
38. B. Vacchini: J. Math. Phys. **42**, 4291 (2001) 131
39. K. Hornberger: Phys. Rev. Lett. **97**, 060601 (2006) 131
40. V. Romero-Rochin, I. Oppenheim: Physica A **155**, 52 (1989) 132
41. V. Romero-Rochin, A. Orsky, I. Oppenheim: Physica A **156**, 244 (1989) 132
42. V. Gorini, M. Verri, A. Frigerio: Physica A **161**, 357 (1989) 132
43. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki: *Quantum Entanglement*, arXiv:quant-ph/0702225 134

Testing Quantum Mechanics in High-Energy Physics

Beatrix C. Hiesmayr

In this set of lectures we show that particle physics can also contribute to fundamental questions about quantum mechanics (QM) and even shine new light in the fine workings of quantum physics and this at scales of energies which are not available for usual quantum systems. In particular the massive meson–antimeson systems are specially suitable as they offer a unique laboratory to test various aspects of particle physics ($\mathcal{CP}$ violation, $\mathcal{CPT}$ violation, etc.) as well as to test the foundations of QM (local realistic theories versus QM, Bell inequalities, decoherence effects, quantum marking and erasure concepts, Bohr's complementary relation, etc.).

Our focus lies upon

- Bell inequalities: Can one find setups for testing local realistic theories versus QM in experiments? What has a symmetry violation in particle physics to do with nonlocality? (Sect. 2)
- Quantum Erasers: "*Erasing the Past and Impacting the Future*," possible with K–mesons? Are there new eraser options? Can they be realized in near future experiments? (Sect. 3)
- How can decoherence models look like for mesons? How do mesons "lose" certain quantum features, e.g., entanglement? Can this be measured in experiments? (Sect. 4)

1 Short Manual to Neutral Kaons

In his 1935 paper [1] Erwin Schrödinger introduced the notion of entanglement (German: "Verschränkung") in reaction to the Einstein, Podolsky, and Rosen (EPR) paper [2] to stress "the basic essence" of quantum mechanics (QM). Since then many experiments in different fields of physics have been performed. Most of these experiments deal with ordinary matter. However, there are many more different

B.C. Hiesmayr (✉)
Faculty of Physics, University Vienna, Austria
e-mail: Beatrix.Hiesmayr@univie.ac.at

Hiesmayr, B.C.: *Testing Quantum Mechanics in High-Energy Physics*. Lect. Notes Phys. **787**, 141–185 (2010)
DOI 10.1007/978-3-642-02871-7_6 © Springer-Verlag Berlin Heidelberg 2010

kinds of particles which can be found in our universe. And interestingly, we can find among them as well systems, e.g., meson–antimeson systems, which show the peculiar EPR-correlations.

A meson consists of a quark and an anti-quark. For example there is the neutral kaon system ($K^0(\bar{s}d)$, $\bar{K}^0(s\bar{d})$) or the neutral B-meson system ($B^0(\bar{b}d)$, $\bar{B}^0(b\bar{d})$). The quarks s/b are called *strange* and *beauty*, respectively. These systems are entangled in their flavor quantum number, i.e., strangeness or beauty. Both systems can be handled within the same theoretical formalism; however, in this work we focus on kaons, and B-mesons are only mentioned when interesting (this will be clear later).

We start by giving a short historical introduction followed by introducing the neutral kaons as viewed nowadays. Then we proceed to Bell inequalities or if local realistic theories can be refuted also for neutral kaons. Then we discuss different types of "kaonic" quantum eraser where we find a new type not offered by other quantum systems and last but not least we discuss different decoherence models and their experimental tests.

1.1 The Discovery of Kaons

What are K-mesons or simply kaons? There have been a lot of puzzles about these particles before a good description was found. Let us shortly review a part of the story, see, e.g., Ref. [3].

In the year 1947 Rochester and Buther studied cosmic-ray showers with a cloud-chamber, and observed a V-shaped track which they interpreted as the decay in flight of a new heavy neutral particle. The momenta of the charged decay products could be measured by the curvature of the tracks in the applied magnetic field, and this knowledge, combined with the measurements of the opening angle of the V and of the ionization density of the tracks, yielded an estimated mass of about a thousand times the electron mass. Nobody at the time could have imagined what "strange" behavior this particle would show and what interesting implementations for theory as well for experiments it would provide!

Repetition of the experiment showed that these particles, together with a lot of other new particles, appeared quite often in cloud-chamber photographs. This indicated that they were produced with a frequency of a few percent relative to pions, the particles predicted by Yukawa in 1935, which had been discovered at about the same time. To be produced as often as this, these particles need to interact with matter almost as strongly as pions do. It was then surprising that they lived long enough to travel a measurable distance in the cloud-chamber; the paradox became more pronounced when it was found that the decay products were none other than pions.

If these particles were coupled so strongly to nucleons, one would expect a decay time of about $10^{-22}s$! There were also other "strange" particles found with this peculiar contrast in the behavior between production and decay.

The solution to the paradox was found in the so-called "strangeness" scheme by Gell-Mann [4], Nishijima and Nakano [5] in 1953, whereby a new quantity called "strangeness" is introduced through the definition

$$S = 2\left(Q - I_3 - \frac{B}{2}\right). \tag{1}$$

The baryon number B is a quantity assigned for example to neutrons and protons which is assumed to be additively conserved to account for the stability of atomic nuclei. The law of baryon conservation rests on very strong empirical evidence: the proton lifetime. Q is the electric charge and is also presumed to be exactly conserved. I_3 is the third component of the isospin, a concept which was introduced to describe the so-called charge-independence of nuclear forces for neutrons and protons. Nuclear and pion–nucleon interactions are known to be invariant under such isospin transformations to an accuracy of a few percent and all known deviations can be ascribed to the action of electro-magnetic interactions which are not isospin invariant. Gell-Mann, Nishijima, and Nakano argued that since strange particles evidently interact strongly with nucleons and pions, the simplest way to assure that interactions involving virtual strange particles do not disturb this well-known approximate symmetry of nuclear interactions is to extend the concept of isospin to apply also to strange particles. Thus every strange particle must be characterized by a well-defined value of I, respectively I_3. For pions and nucleons the quantity on the r.h.s. of (1), i.e., the strangeness number S, vanishes, so these are by definition non-strange particles.

It is clear from (1) that a particle for which S is non-zero cannot transform into any system of non-strange particles through interactions which conserve B, Q, and I_3. As has already been stated, B- and Q-conservation apply absolutely, as far as physicists are presently aware, to all interactions. About 1953 Gell-Mann, Nishijima, and Nakano suggested that I_3 is not conserved in all interactions responsible for strange particles. Assuming that I_3 and therefore S is conserved, then interactions are of a strength similar to nuclear forces, in such a way that they could account for the relatively large cross sections for the production of strange particles. By taking the I_3-nonconserving interactions to be much weaker (in fact similar in strength to the previously known decay interactions responsible for β-decay or π-decay) there was no difficulty in simultaneously explaining their slow decay. Every aspect of the strangeness scheme has been brilliantly confirmed and there is no doubt in its essential correctness. It has led to the nowadays accepted Standard Model (SM) of particle physics. We shall not, however, pursue this subject any further except in relation to the neutral kaon.

1.2 Kaons as Qubits and Quantum States

K-mesons are bound states of quarks and anti-quarks ($q\bar{q}$), where the quark ($q = u, d, s$) can be the up, down, or strange quark, and can be summarized in the following table:

K-meson	Quarks	S	$\mathbf{I_3}$
K^+	$u\bar{s}$	$+1$	$+1/2$
K^-	$\bar{u}s$	-1	$-1/2$
K^0	$d\bar{s}$	$+1$	$-1/2$
$\bar{K}^0$	$\bar{d}s$	-1	$+1/2$
K^0 particle	$\bar{K}^0$ antiparticle	S strangeness	I isospin

Not just for particle physicists the neutral kaon system is unique; these strange mesons are also fantastic quantum systems, we could even say they are selected by Nature to demonstrate fundamental quantum principles such as

- superposition principle,
- oscillation and decay property,
- quasi-spin property.

Their mass is about 497 MeV and they are pseudoscalars $J^P = 0^-$. They interact via strong interactions which are S conserving and weak interactions which are S violating, see Eq. (1). It is due to the weak interactions that the kaons change their oscillation, $K^0 \longleftrightarrow \bar{K}^0$. Both mesons K^0 and $\bar{K}^0$ have transitions to common states, therefore they mix, that means they *oscillate* between K^0 and $\bar{K}^0$ before decaying. For example if we have at time $t = 0$ a K^0 then it can transform virtually into two pions and then recombine into a $\bar{K}^0$ by the following Feynman graph.

Quantum mechanically we can describe the kaons in the following way. Kaons are characterized by their *strangeness* quantum number $+1, -1$:

$$S|K^0\rangle = +|K^0\rangle\,, \qquad S|\bar{K}^0\rangle = -|\bar{K}^0\rangle\,, \tag{2}$$

and the combined operation $\mathcal{CP}$ ($\mathcal{C}$, charge conjugation; $\mathcal{P}$, parity) gives

$$\mathcal{CP}\,|K^0\rangle = -|\bar{K}^0\rangle\,, \qquad \mathcal{CP}\,|\bar{K}^0\rangle = -|K^0\rangle\,. \tag{3}$$

It is straightforward to construct the $\mathcal{CP}$ eigenstates

$$|K_1^0\rangle = \frac{1}{\sqrt{2}}\big\{|K^0\rangle - |\bar{K}^0\rangle\big\}\,, \qquad |K_2^0\rangle = \frac{1}{\sqrt{2}}\big\{|K^0\rangle + |\bar{K}^0\rangle\big\}\,, \tag{4}$$

a quantum number conserved in strong interactions

$$\mathcal{CP}\,|K_1^0\rangle = +|K_1^0\rangle\,, \qquad \mathcal{CP}\,|K_2^0\rangle = -|K_2^0\rangle\,. \tag{5}$$

However, due to weak interactions $\mathcal{CP}$ symmetry is *violated* and the kaons decay in physical states, the short- and long-lived states, $|K_S\rangle$, $|K_L\rangle$, which differ slightly

in mass, $\Delta m = m_L - m_S = 3.49 \times 10^{-6}$ eV, but immensely in their lifetimes and decay modes:

$$|K_S\rangle = \frac{1}{N}\left\{p\,|K^0\rangle - q\,|\bar{K}^0\rangle\right\}, \qquad |K_L\rangle = \frac{1}{N}\left\{p\,|K^0\rangle + q\,|\bar{K}^0\rangle\right\}. \tag{6}$$

The weights $p = 1+\varepsilon$, $q = 1-\varepsilon$, with $N^2 = |p|^2 + |q|^2$ contain the complex $\mathcal{CP}$ *violating parameter* ε with $|\varepsilon| \approx 10^{-3}$. $\mathcal{CPT}$ *invariance* is assumed (T, time reversal). The short-lived K-meson decays dominantly into $K_S \longrightarrow 2\pi$ with a width or lifetime $\Gamma_S^{-1} \sim \tau_S = 0.89 \times 10^{-10}$ s and the long-lived K-meson decays dominantly into $K_L \longrightarrow 3\pi$ with $\Gamma_L^{-1} \sim \tau_L = 5.17 \times 10^{-8}$ s. However, due to $\mathcal{CP}$ violation we observe a small amount $K_L \longrightarrow 2\pi$, for which in 1980 Fitch and Cronin got the Nobel Prize, see Fig. 1 (more details about $\mathcal{CP}$ violation follow in Sect. 1.4).

In this description the superpositions (4) and (6)—or quite generally any vector in the two-dimensional complex Hilbert space of kaons—represent kaonic qubit states in analogy to the qubit states in quantum information. We will call such general superpositions of the strangeness eigenstates ***quasi-spin*** *states.*

1.3 Strangeness Oscillation

Before we discuss $\mathcal{CP}$ violation, let us consider a single kaon evolving in time. K_S, K_L are eigenstates of a non-Hermitian "effective mass" Hamiltonian

$$H \;=\; M - \frac{\mathrm{i}}{2}\,\Gamma \tag{7}$$

satisfying

$$H\,|K_{S,L}\rangle \;=\; \lambda_{S,L}\,|K_{S,L}\rangle \qquad \text{with} \qquad \lambda_{S,L} \;=\; m_{S,L} - \frac{\mathrm{i}}{2}\,\Gamma_{S,L}\,. \tag{8}$$

Since the decaying states evolve—according to the Wigner–Weisskopf approximation—exponentially in time

$$|K_{S,L}(t)\rangle \;=\; e^{-\mathrm{i}\lambda_{S,L}t}|K_{S,L}\rangle\,, \tag{9}$$

the subsequent time evolution for K^0 and $\bar{K}^0$ is given by

$$|K^0(t)\rangle = g_+(t)|K^0\rangle + \frac{q}{p}g_-(t)|\bar{K}^0\rangle\,, \tag{10}$$

$$|\bar{K}^0(t)\rangle = \frac{p}{q}g_-(t)|K^0\rangle + g_+(t)|\bar{K}^0\rangle, \tag{11}$$

with

$$g_\pm(t) \;=\; \frac{1}{2}\left[\pm e^{-\mathrm{i}\lambda_S t} + e^{-\mathrm{i}\lambda_L t}\right]. \tag{12}$$

Supposing that a K^0 beam is produced at $t = 0$, e.g., by the strong process $\pi^- + p \longrightarrow K^0 + \Lambda^0$, then the probability of finding a K^0 at time t in the beam is calculated to be

$$\begin{aligned} P(K^0, t; |K^0|) &= \left|\langle K^0|K^0(t)\rangle\right|^2 = \left|\langle K^0| \frac{N}{2p}\left\{e^{-i\lambda_S t}\,|K_S\rangle + e^{-i\lambda_L t}\,|K_L\rangle\right\}\right|^2 \\ &= |\frac{1}{2}\,e^{-i\lambda_S t} + \frac{1}{2}\,e^{-i\lambda_L t}|^2 \\ &= \frac{1}{4}\left\{e^{-\Gamma_S t} + e^{-\Gamma_L t} + 2\mathrm{Re}\{e^{-i\Delta m t}\}e^{-\Gamma t}\right\} \\ &= \frac{1}{4}\left\{e^{-\Gamma_S t} + e^{-\Gamma_L t} + 2\cos(\Delta m t)e^{-\Gamma t}\right\} \end{aligned} \tag{13}$$

or to find a $\bar{K}^0$ at time t:

$$\begin{aligned} P(\bar{K}^0, t; |K^0|) &= \left|\langle \bar{K}^0|K^0(t)\rangle\right|^2 = \left|\langle \bar{K}^0| \frac{N}{2p}\left\{e^{-i\lambda_S t}\,|K_S\rangle + e^{-i\lambda_L t}\,|K_L\rangle\right\}\right|^2 \\ &= |\frac{q}{2p}\,e^{-i\lambda_S t} - \frac{q}{2p}\,e^{-i\lambda_L t}|^2 \\ &= \frac{|q|^2}{4|p|^2}\left\{e^{-\Gamma_S t} + e^{-\Gamma_L t} - 2\cos(\Delta m t)e^{-\Gamma t}\right\}, \end{aligned} \tag{14}$$

with $\Delta m = m_L - m_S$ and $\Gamma = \frac{1}{2}(\Gamma_L + \Gamma_S)$.

The K^0 beam oscillates with frequency $\Delta m/2\pi$, where $\Delta m\,\tau_S \approx 0.47$. The oscillation is clearly visible at times of the order of a few τ_S, before all K_Ss have died out leaving only the K_Ls in the beam. So in a beam which contains only K^0 mesons at the beginning $t = 0$ there will occur $\bar{K}^0$ far from the production source through its presence in the K_L meson and this is called *strangeness oscillation*.

1.4 $\mathcal{CP}$ Violation

Now let us consider the charge asymmetry term defined by

$$\begin{aligned} \delta(t) &= \frac{P(K^0, t; |K^0|) - P(\bar{K}^0, t; |K^0|)}{P(K^0, t; |K^0|) + P(\bar{K}^0, t; |K^0|)} \\ &= \frac{\dfrac{|p|^2 - |q|^2}{N^2}\,(e^{-\Gamma_S t} + e^{-\Gamma_L t}) + 2\cos(\Delta m t)e^{-\Gamma t}}{(e^{-\Gamma_S t} + e^{-\Gamma_L t}) + \dfrac{|p|^2 - |q|^2}{N^2}\,2\cos(\Delta m t)e^{-\Gamma t}} \\ &= \frac{\cos(\Delta m t) + \delta\cosh(\Delta\Gamma/2\,t)}{\cosh(\Delta\Gamma/2\,t) + \delta\cos(\Delta m t)}, \end{aligned} \tag{15}$$

with $\delta = \frac{|p|^2-|q|^2}{|p|^2+|q|^2}$ and $\Delta\Gamma = \Gamma_L - \Gamma_S$. Expanding in the small $\mathcal{CP}$ violating parameter δ we obtain

$$\delta(t) = \frac{\cos(\Delta m t)}{\cosh(\Delta\Gamma/2\, t)} + \delta\,(1 - \frac{\cos^2(\Delta m t)}{\cosh^2(\Delta\Gamma/2\, t)}) + O(\delta^2). \tag{16}$$

For times $t \approx \tau_S$ one observes the strangeness oscillation, while for times $t \gg \tau_S$ one is sensitive to the small $\mathcal{CP}$ violating parameter δ. For large times t only the long-lived component K_L survives.

Experimentally one considers the total decay constants of the long-lived state into the $l3$ channel (l stands for lepton, here electron or muon)

$$\delta_l = \frac{\Gamma(K_L \to \pi^- l^+ \nu_l) - \Gamma(K_L \to \pi^+ l^- \bar{\nu}_l)}{\Gamma(K_L \to \pi^- l^+ \nu_l) + \Gamma(K_L \to \pi^+ l^- \bar{\nu}_l)}, \tag{17}$$

which would be zero for $\mathcal{CP}$ invariance. One finds that this leptonic asymmetry is nonvanishing, namely $\delta_l \approx 10^{-3}$. In our formalism and with the suitable phase choices this quantity equals

$$\delta_l \;=\; \delta \;=\; \frac{|p|^2 - |q|^2}{|p|^2 + |q|^2} \;=\; \frac{2\mathrm{Re}\{\varepsilon\}}{1 + |\varepsilon|^2} \;\equiv\; \langle K_S | K_L \rangle. \tag{18}$$

The empirical fact that the charge asymmetry does not vanish has the following meaning for physics: First, it is now possible to define the electric charge in an absolute sense. Positive charge is the charge of the lepton more often produced in the semileptonic K_L decays. And second, through $\mathcal{CP}$ violation we have an absolute definition of left and right, thus there is a difference between the world and the anti-world.

1.5 Quasi-spin of Kaons and Analogy to Photons

In comparison with spin-$\frac{1}{2}$ particles or with photons having the polarization directions V (vertical) and H (horizontal), it is very instructive to characterize the kaons by a *quasi-spin* (for details see Ref. [6]). We can regard the two states $|K^0\rangle$ and $|\bar{K}^0\rangle$ as the quasi-spin states up $|\uparrow\rangle$ and down $|\downarrow\rangle$ and can express the operators acting in this quasi-spin space by Pauli matrices. So we identify the strangeness operator S with the Pauli matrix σ_3, the $\mathcal{CP}$ operator with $(-\sigma_1)$ and for describing $\mathcal{CP}$ violation we also need σ_2. In fact, the Hamiltonian (7) then has the form

$$H \;=\; a\mathbf{1} + \mathbf{b}\cdot\boldsymbol{\sigma}\,, \tag{19}$$

with

$$b_1 = b\cos\alpha, \quad b_2 = b\sin\alpha, \quad b_3 = 0\,,$$
$$a = \frac{1}{2}(\lambda_L + \lambda_S), \quad b = \frac{1}{2}(\lambda_L - \lambda_S)\,, \tag{20}$$

and the angle α is related to the $\mathcal{CP}$ violating parameter ε by

$$e^{\mathrm{i}\alpha} = \frac{1-\varepsilon}{1+\varepsilon}\,. \tag{21}$$

The kaonic—quasi-spin—photonic analogy is summarized in the following table:

Neutral kaon	**Quasi–spin**	**Photon**
$\|K^0\rangle$	$\|\uparrow\rangle_z$	$\|V\rangle$
$\|\bar{K}^0\rangle$	$\|\downarrow\rangle_z$	$\|H\rangle$
$\|K_1^0\rangle$	$\|\nwarrow\rangle_x$	$\|-45^0\rangle = \frac{1}{\sqrt{2}}(\|V\rangle - \|H\rangle)$
$\|K_2^0\rangle$	$\|\nearrow\rangle_x$	$\|+45^0\rangle = \frac{1}{\sqrt{2}}(\|V\rangle + \|H\rangle)$
$\|K_S\rangle$	$\|\rightarrow\rangle_y$	$\|L\rangle = \frac{1}{\sqrt{2}}(\|V\rangle - \mathrm{i}\|H\rangle)$
$\|K_L\rangle$	$\|\leftarrow\rangle_y$	$\|R\rangle = \frac{1}{\sqrt{2}}(\|V\rangle + \mathrm{i}\|H\rangle)$

Note that neglecting the small $\mathcal{CP}$ violation $|K^0_{1,2}\rangle$ are equivalent to the mass eigenstates $|K_{S,L}\rangle$. Thus the last analogy just emphasizes that in the kaon case we have three important bases which in contrast to photons and spin-$\frac{1}{2}$ particles are not orthogonal.

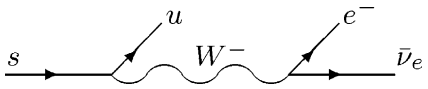
A good *optical analogy* to the phenomenon of strangeness oscillation can be achieved by using the physical effect of birefringence in optical fibers which lead to the rotation of polarization directions. Thus H (horizontal) polarized light is rotated after some distance into V (vertical) polarized light, and so on. On the other hand, the decay of kaons can be simulated by polarization-dependent losses in optical fibres, where one state has lower losses than its orthogonal state, see Ref. [7].

The description of kaons as qubits reveals close analogies to photons but also deep physical differences. Kaons oscillate, they are massive, they decay and can be characterized by symmetries like $\mathcal{CP}$. Even though some kaon features, like oscillation and decay, can be mimicked by photon experiments, they are inherently different since they are intrinsic properties of the kaon given by Nature. This is a crucial point when discussing Bell inequalities or other quantum features.

1.6 Two Measurement Procedures

For neutral kaons there exist two physical alternative bases; accordingly we have two observables for the kaons, namely the projectors on the two bases. The first basis is the strangeness eigenstate basis $\{|K^0\rangle, |\bar{K}^0\rangle\}$; it can be measured by inserting a piece of ordinary matter along the kaon trajectory, which corresponds to an *active* measurement of strangeness. Due to strangeness conservation of the strong interactions the incoming state is projected either onto K^0 by $K^0 p \to K^+ n$ or onto $\bar{K}^0$ by $\bar{K}^0 p \to \Lambda\pi^+$, $\bar{K}^0 n \to \Lambda\pi^0$ or $\bar{K}^0 n \to K^- p$. Here nucleonic matter plays the same role as a two channel analyzer for polarized photon beams.

Alternatively, the strangeness content of neutral kaons can be determined by observing their semileptonic decay modes. The strange quark s decays weakly as constituent of $\bar{K}^0$:



Due to their quark content the kaon $K^0(\bar{s}d)$ and the anti-kaon $\bar{K}^0(s\bar{d})$ have the following definite decays:

$$\begin{aligned} K^0(d\bar{s}) &\longrightarrow \pi^-(d\bar{u})\, l^+ \nu_l \qquad \text{where} \qquad \bar{s} \longrightarrow \bar{u}\, l^+ \nu_l \\ \bar{K}^0(\bar{d}s) &\longrightarrow \pi^+(\bar{d}u)\, l^- \bar{\nu}_l \qquad \text{where} \qquad s \longrightarrow u\, l^- \bar{\nu}_l\,, \end{aligned} \tag{22}$$

with l either muon or electron, $l = \mu, e$. When studying the leptonic charge asymmetry (compare with Eqs. (15) and (17))

$$\delta_l = \frac{\Gamma(K_L \to \pi^- l^+ \nu_l) - \Gamma(K_L \to \pi^+ l^- \bar{\nu}_l)}{\Gamma(K_L \to \pi^- l^+ \nu_l) + \Gamma(K_L \to \pi^+ l^- \bar{\nu}_l)}\,, \tag{23}$$

we notice that l^+ and l^- tag K^0 and $\bar{K}^0$, respectively, in the K_L state, and the leptonic asymmetry (23) is expressed by the probabilities $|p|^2$ and $|q|^2$ of finding a K^0 and a $\bar{K}^0$, respectively, in the K_L state:

$$\delta = \frac{|p|^2 - |q|^2}{|p|^2 + |q|^2}\,. \tag{24}$$

Obviously, the experimenter has no control of the kaon decay, neither of the mode nor of the time. The experimenter can only sort at the end of the day all observed events in proper decay modes and time intervals. We call this procedure opposite to the *active* measurement described above a *passive* measurement procedure of strangeness.

The second basis $\{K_S, K_L\}$ consists of the short- and long-lived states having well-defined masses $m_{S(L)}$ and decay widths $\Gamma_{S(L)}$. We have seen that it is the appropriate basis to discuss the kaon propagation in free space, because these states

preserve their own identity in time (9). Due to the huge difference in the decay widths the K_Ss decay much faster than the K_Ls. Thus in order to observe if a propagating kaon is a K_S or K_L at an instant time t, one has to detect at which time it subsequently decays. Kaons which are observed to decay before $\simeq t + 4.8\,\tau_S$ have to be identified as K_Ss, while those surviving after this time are assumed to be K_Ls. The number $4.8\tau_S$ is obtained by setting the probability to observe a K_S equal to the probability that a K_L state does not decay, i.e., $e^{-\Gamma_S t} \stackrel{!}{=} 1 - e^{-\Gamma_L t} \longrightarrow t \approx 4.8\tau_S$. This means that we equalized the errors in measuring both lifetime states and misidentifications reduce only to a few parts in 10^{-3} (see also Refs. [8, 9]). Note that the experimenter does not care about the specific decay mode; he records only a decay event at a certain time. We call this procedure an *active* measurement of lifetime.

Since the neutral kaon system violates the $\mathcal{CP}$ symmetry (recall Section 1.4) the mass eigenstates are not strictly orthogonal, $\langle K_S|K_L\rangle \neq 0$. However, neglecting $\mathcal{CP}$ violation—remember it is of the order of 10^{-3}—the K_Ss are identified by a 2π final state and K_Ls by a 3π final state. We call this procedure a *passive* measurement of lifetime, since the kaon decay times and decay channels used in the measurement are entirely determined by the quantum nature of kaons and cannot be in any way influenced by the experimenter.

Summarizing, we have the following two measurement procedures for the strangeness and the mass-eigenstate basis:

Strangeness basis $\langle K^0|\bar{K}^0\rangle = 0$

Active measurements (strong interactions)	*Passive* measurements (semileptonic decays)
$K^0 + p \longrightarrow K^+ + n$	$\bar{K}^0(\bar{d}s) \longrightarrow \pi^+(\bar{d}u)\ l^-\ \bar{\nu}_l$
$\bar{K}^0 + p \longrightarrow \Lambda + \pi^+$	$K^0(d\bar{s}) \longrightarrow \pi^-(d\bar{u})\ l^+\ \nu_l$
$\bar{K}^0 + n \longrightarrow K^- + p,\ \Lambda + \pi^0$	

Mass eigenstate basis $\langle K_S|K_L\rangle = \frac{2\mathrm{Re}\{\varepsilon\}}{1+|\varepsilon|^2} \approx 3 \times 10^{-3}$

Active measurements (free propagation)	*Passive* measurements (2 or 3π decay modes)
$\rightarrow$ any decay mode observed before	$\rightarrow 2\pi$s are identified as K_Ss
$4.8\tau_S$ are identified as K_S's	$\rightarrow 3\pi$s are identified as K_Ls
Misidentification: few parts in 10^{-3}!	Misidentification: few parts in 10^{-3}!

1.7 Entangled in Strangeness

Interestingly, also for strange mesons entangled states can be obtained, in analogy to the entangled spin up and down pairs, or H and V polarized photon pairs. Such states are produced by e^+e^--colliders through the reaction $e^+e^- \to \Phi \to K^0\bar{K}^0$, in particular at DA$\Phi$NE in Frascati, or they are produced in $p\bar{p}$–collisions, like, e.g., at LEAR at CERN. There, a $K^0\bar{K}^0$ pair is created in a $J^{PC} = 1^{--}$ quantum state and thus antisymmetric under C and P, and is described at the time $t = 0$ by the entangled state:

$$\begin{aligned} |\psi(t=0)\rangle &= \frac{1}{\sqrt{2}}\left\{|K^0\rangle_l \otimes |\bar{K}^0\rangle_r - |\bar{K}^0\rangle_l \otimes |K^0\rangle_r\right\} , \\ &= \frac{N_{SL}}{\sqrt{2}}\left\{|K_S\rangle_l \otimes |K_L\rangle_r - |K_L\rangle_l \otimes |K_S\rangle_r\right\} , \end{aligned} \tag{25}$$

with $N_{SL} = \frac{N^2}{2pq}$, in complete analogy to the entangled photon case

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}}\left\{|V\rangle_l \otimes |H\rangle_r - |H\rangle_l \otimes |V\rangle_r\right\} , \\ &= \frac{1}{\sqrt{2}}\left\{|L\rangle_l \otimes |R\rangle_r - |R\rangle_l \otimes |L\rangle_r\right\} . \end{aligned} \tag{26}$$

The neutral kaons fly apart and are detected on the left (l) and right (r) hand side of the source. Of course, during their propagation the $K^0\bar{K}^0$ pairs oscillate and the K_S, K_L states decay. This is an important difference to the case of photons which are stable.

Let us measure at time t_l a K^0 meson on the left hand side and at time t_r a K^0 or a $\bar{K}^0$ on the right hand side then we find an EPR–Bell correlation (**E**instein–**P**odolsky–**R**osen) analogously to the entangled photon case with polarization V–V or V–H. Assuming for simplicity stable kaons ($\Gamma_S = \Gamma_L = 0$) then we get the following result for the quantum probabilities:

$$\begin{aligned} P(K^0, t_l; K^0, t_r) = P(\bar{K}^0, t_l; \bar{K}^0, t_r) &= \frac{1}{4}\left\{1 - \cos(\Delta m(t_l - t_r))\right\} , \\ P(K^0, t_l; \bar{K}^0, t_r) = P(\bar{K}^0, t_l; K^0, t_r) &= \frac{1}{4}\left\{1 + \cos(\Delta m(t_l - t_r))\right\} , \end{aligned} \tag{27}$$

which is analogous to the probabilities of finding simultaneously two entangled photons along two chosen directions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$:

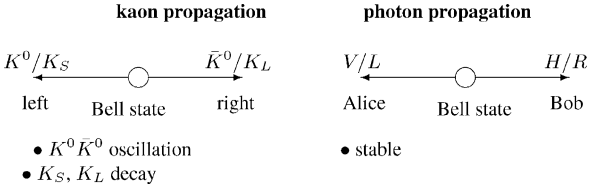
$$\begin{aligned} P(\boldsymbol{\alpha}, V; \boldsymbol{\beta}, V) = P(\boldsymbol{\alpha}, H; \boldsymbol{\beta}, H) &= \frac{1}{4}\left\{1 - \cos 2(\alpha - \beta)\right\} , \\ P(\boldsymbol{\alpha}, V; \boldsymbol{\beta}, H) = P(\boldsymbol{\alpha}, H; \boldsymbol{\beta}, V) &= \frac{1}{4}\left\{1 + \cos 2(\alpha - \beta)\right\} . \end{aligned} \tag{28}$$

Thus we observe a *perfect analogy* between times $\Delta m(t_l - t_r)$ and angles $2(\alpha - \beta)$.

Alternatively, we can also fix the time and vary the quasi-spin of the kaon, which corresponds to a rotation in quasi-spin space analogously to the rotation of polarization of the photon:

$$|k\rangle = a|K^0\rangle + b|\bar{K}^0\rangle \quad \longleftrightarrow \quad |\alpha, \phi; V\rangle = \cos\alpha|V\rangle + \sin\alpha\, e^{i\phi}|H\rangle\,. \quad (29)$$

Note that the weights a, b are not independent and not all kaonic superpositions are realized in Nature in contrast to photons. Depicting the kaonic–photonic analogy we have:



1.8 Kaons as Double Slits

The famous statement "the double slit contains the only mystery" of Richard Feynman is well known; his statement about kaons is not less to the point "If there is any place where we have a chance to test the main principles of quantum mechanics in the purest way—does the superposition of amplitudes work or doesn't it?— this is it." [10]. In this section we argue that single neutral kaons can be considered as double slits as well.

Bohr's complementarity principle or the closely related concept of duality in interferometric or double-slit like devices are at the heart of quantum mechanics. The qualitative well-known statement that "the observation of an interference pattern and the acquisition of which-way information are mutually exclusive" has only recently been rephrased to a quantitative statement [11, 12]:

$$\mathcal{P}^2(y) + \mathcal{V}_0^2(y) \leq 1\,, \quad (30)$$

where the equality is valid for pure quantum states and the inequality for mixed ones. $\mathcal{V}_0(y)$ is the fringe visibility which quantifies the sharpness or contrast of the interference pattern ("the wave-like property") and can depend on an external parameter y, whereas $\mathcal{P}(y)$ denotes the path predictability, i.e., the a priori knowledge one can have on the path taken by the interfering system ("the particle-like property"). It is defined by

$$\mathcal{P}(y) = |p_I(y) - p_{II}(y)|\,, \quad (31)$$

where $p_I(y)$ and $p_{II}(y)$ are the probabilities for taking each path ($p_I(y) + p_{II}(y) = 1$). It is often too idealized to assume that the predictability

and visibility are independent on an external parameter. For example, consider a usual double slit experiment, then the intensity is generally given by

$$I(y) \;=\; F(y)\left(1+\mathcal{V}_0(y)\cos(\phi(y)\right), \tag{32}$$

where $F(y)$ is specific for each setup and $\phi(y)$ is the phase-difference between the two paths. The variable y characterizes the detector position in this case, thus visibility and predictability are y-dependent.

In Ref. [13] physical situations were investigated for which the expressions of $\mathcal{V}_0(y)$, $\mathcal{P}(y)$ and $\phi(y)$ can be calculated analytically. This included interference patterns of various types of double slit experiments (y is linked to position), but also oscillations due to particle mixing (y is linked to time), e.g., by the kaon system, and also Mott scattering experiments of identical particles or nuclei (y is linked to a scattering angle). All these two-state systems belonging to distinct fields of physics can then be treated via the generalized complementarity relation in a unified way. Even for specific thermodynamical systems Bohr's complementarity can manifest itself, see Ref. [14]. Here we investigate briefly the neutral kaon case; for more details and applications to other systems see Ref. [13].

The time evolution of an initial K^0 state is given by Eq. (10) (in the following $\mathcal{CP}$ violation effects can safely be neglected):

$$|K^0(t)\rangle = \frac{1}{\sqrt{2}} e^{-\mathrm{i}m_L t-\frac{\Gamma_L}{2}t}\left\{e^{\mathrm{i}\Delta m t+\frac{\Delta\Gamma}{2}t}\,|K_S\rangle+|K_L\rangle\right\}, \tag{33}$$

where we denoted $\Delta\Gamma = \Gamma_L - \Gamma_S < 0$. We are only interested in kaons which survive up to a certain time t, thus we consider the following normalized state:

$$|K^0(t)\rangle \cong \frac{1}{\sqrt{2\cosh(\frac{\Delta\Gamma}{2}t)}}\, e^{-\frac{\Delta\Gamma}{4}t}\left\{e^{\mathrm{i}\Delta m t+\frac{\Delta\Gamma}{2}t}\,|K_S\rangle+|K_L\rangle\right\}. \tag{34}$$

State (34) can be interpreted as follows. The two mass eigenstates $|K_S\rangle$, $|K_L\rangle$ represent the two slits. At time $t = 0$ both slits have the same width and as time evolves one slit decreases as compared to the other; however, in addition the whole double slit shrinks due to the decay property of the kaons. This analogy gets more obvious if we consider for an initial K^0 the probabilities for finding after a certain time t a K^0 or a $\bar{K}^0$ state, i.e., the strangeness oscillation:

$$\begin{aligned} P(K^0,t) &= \left|\langle K^0|K^0(t)\rangle\right|^2 = \frac{1}{2}\left\{1+\frac{\cos(\Delta m t)}{\cosh(\frac{\Delta\Gamma}{2}t)}\right\},\\ P(\bar{K}^0,t) &= \left|\langle \bar{K}^0|K^0(t)\rangle\right|^2 = \frac{1}{2}\left\{1-\frac{\cos(\Delta m t)}{\cosh(\frac{\Delta\Gamma}{2}t)}\right\}. \end{aligned} \tag{35}$$

We observe that the oscillating phase is given by $\phi(t) = \Delta m\, t$ and the time-dependent visibility by

$$\mathcal{V}_0(t) = \frac{1}{\cosh(\frac{\Delta\Gamma}{2}t)}\,, \tag{36}$$

which is maximal at $t = 0$. The "which width" information, analogously to the "which way" information in usual double slit experiments, can be directly calculated from Eq. (34) and gives the following predictability:

$$\mathcal{P}(t) = |P(K_S,t) - P(K_L,t)| = \left| \frac{e^{\frac{\Delta\Gamma}{2}t} - e^{-\frac{\Delta\Gamma}{2}t}}{2\cosh(\frac{\Delta\Gamma}{2}t)} \right| = \left| \tanh\left(\frac{\Delta\Gamma}{2}t\right) \right| . \tag{37}$$

The larger the time t is, the more probable the propagation of the K_L component is, because the K_S component has died out; the predictability converges to its upper bound 1.

Both expressions, the predictability (37) and the visibility (36), satisfy the complementary relation (30) for all times t with the equality sign:

$$\mathcal{P}^2(t) + \mathcal{V}_0^2(t) \;=\; \tanh^2\left(\frac{\Delta\Gamma}{2}t\right) + \frac{1}{\cosh^2(\frac{\Delta\Gamma}{2}t)} \;=\; 1\,. \tag{38}$$

For time $t = 0$ there is full interference, the visibility is $\mathcal{V}_0(t=0) = 1$, and we have no information about the lifetimes or widths, $\mathcal{P}(t=0) = 0$. This corresponds to the usual double slit scenario. For large times, i.e., $t \gg 1/\Gamma_S$, the kaon is most probable in a long-lived state K_L and no interference is observed, we have information on the "which width." For times between the two extremes we obtain partially information on "which width" and on the interference contrast due to the natural instability of the kaons. However, the full information on the system is contained in Eq. (30) or rather in Eq. (38) and is for pure systems always maximal.

The complementarity principle was phrased by Niels Bohr in an attempt to express the most fundamental difference between classical and quantum physics. According to this principle, and in sharp contrast to classical physics, in quantum physics we cannot capture all aspects of reality simultaneously, the information content is always limited. Neutral kaons encapsulate indeed this peculiar feature in the very same way as a particle traveling through a double slit. But kaons are double slits provided by Nature for free!

1.9 Basic Definitions: Summarized

The normalization and the relative phase of the eigenvectors of the non-Hermitian mass matrix H, (7), are chosen by

$$\begin{aligned} &\langle K_S|K_S\rangle = \langle K_L|K_L\rangle = 1, \\ &\langle K_S|K_L\rangle = \langle K_S|K_L\rangle^* \geq 0. \end{aligned} \tag{39}$$

CPT(T, time reversal) invariance is assumed, but not CP-invariance. The following definitions (phase choices) are made:

$$\boxed{\begin{array}{l} p = 1 + \varepsilon \\ q = 1 - \varepsilon \\ N^2 = |p|^2 + |q|^2 = 2(1 + |\varepsilon|^2) \end{array}}$$

where ε is called the (indirect) $\mathcal{CP}$ violating parameter in mixing.

- The strangeness eigenstates K^0, $\bar{K}^0$:

$$\begin{aligned} |K^0\rangle &= \frac{1}{\sqrt{2}}\{\ \ |K_1^0\rangle + |K_2^0\rangle\} = \frac{N}{2p}\{\ \ |K_S\rangle + |K_L\rangle\}, \\ |\bar{K}^0\rangle &= \frac{1}{\sqrt{2}}\{-|K_1^0\rangle + |K_2^0\rangle\} = \frac{N}{2q}\{-|K_S\rangle + |K_L\rangle\}. \end{aligned} \tag{40}$$

- The mass eigenstates K_S, K_L:

$$\begin{aligned} |K_S\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\{|K_1^0\rangle + \varepsilon|K_2^0\rangle\} = \frac{1}{N}\{p|K^0\rangle - q|\bar{K}^0\rangle\}, \\ |K_L\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\{|K_2^0\rangle + \varepsilon|K_1^0\rangle\} = \frac{1}{N}\{p|K^0\rangle + q|\bar{K}^0\rangle\}. \end{aligned} \tag{41}$$

- The $\mathcal{CP}$ eigenstates K_1^0, K_2^0:

$$\begin{aligned} |K_1^0\rangle &= \frac{1}{\sqrt{2}}\{|K^0\rangle - |\bar{K}^0\rangle\} = \frac{\sqrt{1+|\varepsilon|^2}}{1-\varepsilon^2}\{|K_S\rangle - \varepsilon|K_L\rangle\}, \\ |K_2^0\rangle &= \frac{1}{\sqrt{2}}\{|K^0\rangle + |\bar{K}^0\rangle\} = \frac{\sqrt{1+|\varepsilon|^2}}{1-\varepsilon^2}\{|K_L\rangle - \varepsilon|K_S\rangle\}. \end{aligned} \tag{42}$$

The experimental measured quantity, the leptonic charge asymmetry δ, (23), is formally derived by

$$\begin{aligned} \langle \bar{K}^0|\bar{K}^0\rangle = 1 &\stackrel{!}{=} \\ &\stackrel{!}{=} \frac{N^2}{4|q|^2}\{\langle K_S|K_S\rangle + \langle K_L|K_L\rangle - \langle K_S|K_L\rangle - \langle K_L|K_S\rangle\} \\ &= \frac{N^2}{4|q|^2}\{2 - 2\langle K_S|K_L\rangle\} \\ &\Longrightarrow \delta := \langle K_S|K_L\rangle = \frac{|p|^2 - |q|^2}{N^2} = \frac{2\mathrm{Re}\{\varepsilon\}}{1+|\varepsilon|^2}\,. \end{aligned} \tag{43}$$

2 Bell Inequalities in High-Energy Physics

After some short introduction to the extensively discussed Bell inequalities, we proceed with kaons. We first introduce the most general CHSH–Bell inequality and then discuss two cases. One sensitive to $\mathcal{CP}$ violation and one sensitive to strangeness. The last one would allow a direct test of local realism, however, to find a violation and to understand its nature we have to go deep into the matter.

2.1 Short Introduction to BIs

Bell inequalities are a compelling example of an essential difference between quantum mechanics and classical physics. Understanding this difference is vital in learning how a quantum world works compared to the classical one, or more practically how to use it for technical applications, such as quantum communication and the quantum computations.

When one speaks in the everyday world of something like the moon or a human being, we assume that the physical properties of that object have an existence independent of its observation. The study of quantum theory and their immense verification in experiments forces us to conclude that an unobserved particle does not possess physical properties which exist independently of observation. Rather, such physical properties arise as a consequence of measurements performed upon the system. For instance a neutral kaon does not possess definite properties of being a particle with strangeness $S = +1$ *and* being a short-lived kaon each of which can be revealed by performing the appropriate measurement. Rather, quantum mechanics gives a set of rules which specify, given the state vector or the density matrix, the probabilities for the possible measurement outcomes if we decide to measure its strangeness content or its mass eigenstate.

Many physicists rejected this view of Nature and for non-quantum physicists the above statement must sound crazy. Even the co-founders of quantum theory were not satisfied with the achievements of their own developments. The most famous paper is by Einstein, Podolsky and Rosen [2] from 1935 in which they — as they believed — demonstrated that quantum mechanics is not a complete theory of Nature. Their belief, nowadays called the *EPR reality criterion*, is:

> If, without in any way disturbing a system, we can predict with certainty (i.e. with the probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

They developed a gedanken experiment, nowadays known as the EPR-experiment. They considered two particles flying in opposite directions where on the left hand side the experimenter, usually called Alice, and on the right hand side the experimenter, called Bob, can perform measurements on their particle. Each of them can choose *actively* among alternative setups, i.e., by exerting their free will. Then later on they meet and compare their results. If the initial state is a maximally entangled one, e.g., the antisymmetric Bell state

$$
\begin{aligned}
|\psi^-\rangle &= \frac{1}{\sqrt{2}}\left\{|\Uparrow\rangle_l \otimes |\Downarrow\rangle_r - |\Downarrow\rangle_l \otimes |\Uparrow\rangle_r\right\} \\
&= \frac{1}{\sqrt{2}}\left\{|0\rangle_l \otimes |1\rangle_r - |1\rangle_l \otimes |0\rangle_r\right\} \\
&= \frac{1}{\sqrt{2}}\left\{|H\rangle_l \otimes |V\rangle_r - |V\rangle_l \otimes |H\rangle_r\right\} \\
&= \frac{1}{\sqrt{2}}\left\{|K^0\rangle_l \otimes |\bar{K}^0\rangle_r - |\bar{K}^0\rangle_l \otimes |K^0\rangle_r\right\} \\
&= \frac{1}{\sqrt{2}}\left\{|B^0\rangle_l \otimes |\bar{B}^0\rangle_r - |\bar{B}^0\rangle_l \otimes |B^0\rangle_r\right\} \\
&= \frac{1}{\sqrt{2}}\left\{|I\rangle_l \otimes |\Uparrow\rangle_r - |II\rangle_l \otimes |\Downarrow\rangle_r\right\},
\end{aligned}
\tag{44}
$$

they observe strong correlations. If Alice finds in a certain direction spin up, then if by chance Bob measures in the same direction he finds with 100% probability that his particle is in the spin down state. Here the first line describes two spin-$\frac{1}{2}$ particles, the second line is given in the general notation of qubits, $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$, and the third line describes entangled photons (H/V, horizontal/vertical polarized). The fourth and fifth lines give the entangled state for K-mesons and B-mesons, respectively. The last line describes a single neutron propagating through a two-way interferometer (path I and path II) where the different paths correspond to different spins. Thus the single neutron is entangled in its outer degrees of freedom (path) and its inner degrees of freedom (spin).

To explain these correlations the EPR people argued that quantum mechanics (QM) is incomplete in the sense they defined above. Each particle has to carry a hidden parameter in order to ensure that the other particle always gives the opposite result, if Alice and Bob by chance have measured in the same direction.

In 1964 John S. Bell proved the important theorem that a whole class of such local realistic hidden variable theories cannot reproduce all statistical predictions of QM, Refs. [15, 16]. He discovered that any local realistic theory forces the obtained two-particle correlation functions to satisfy an inequality, called Bell inequality (BI), whereas QM, in certain cases, violates explicitly this BI. The importance of Bell's theorem lies in the experimental feasibility to discriminate between QM and local realistic theories. Abner Shimony has denoted that as a decision of a philosophical question via an experiment.

Indeed, Nature had the last laugh on the EPR paradox. Many experiments have been carried out, mainly with photons, e.g., Refs. [17–22], but also with atoms, e.g., Ref. [19], and confirm the quantum mechanical predictions. Hasegawa et al. in Ref. [23] reported an experiment with single neutrons in an interferometric device which shows a violation of a Bell-like inequality. The entanglement is achieved not between two separate particles but between two degrees of freedom

of a single neutron, namely, between the path it takes in the interferometer and its spin component which is different for the two paths. The mathematical description of the entangled state is the same as for the previously mentioned systems, see (44). However, as there are no two spatially separated particles, it is contextuality rather than nonlocality that is tested. Neutrons are also suited to study geometrical phases as the famous Berry phase, for a connection between Berry phases and contextuality [24].

It is well known that these initially more philosophical considerations of Einstein, Podolsky, and Rosen and Bell's theorem triggered a totally new field with ingenious and fascinating experiments and even new applications that will change our daily life: for instance, quantum cryptography (see, e.g., Ref. [25]), quantum teleportation [26] and its experimental realizations, e.g., Refs. [27, 28], and Shor's algorithm [29]. The consequences are far reaching and lead to a new discipline called quantum information science. While prototypes in quantum cryptography already exist, the realization of a quantum computer is still a formidable experimental challenge.

Moreover, it has turned out that there is also a quite practical use for BIs:

- It has been proven that quantum cryptography protocols are only safe if a Bell inequality is violated (see, e.g., Ref. [25] and references therein).
- In the quantum communication complexity field Bell inequalities are necessary and sufficient conditions for quantum protocols to beat the classical ones.

2.2 Bell Inequalities in High-Energy Physics

Knowing that in high-energy physics entangled systems also exist, it is obvious to ask whether Bell inequalities can also be violated for such massive systems produced in accelerators?

It turned out that the question can not be answered in a simple way. To describe the challenging difference of mesons compared to other quantum systems, e.g., photons, is one important requirement. A general formalism that allows for the description of decaying systems within quantum mechanics has been developed in Ref. [30], applied to kaons in Ref. [31], and is reviewed in Sect. 2.5.

The derivation of the CHSH–Bell type inequality for neutral particles is analogous to the original proof of Clauser, Horne, Shimony and Holt in 1969, Ref. [32], which is an extension of Bell's original proof but under the more realistic assumption that due to experimental imperfections not all particles can be detected. It can be found in Ref. [6]. In contrast to the case of photons, for kaons each experimenter can choose two properties, i.e., the quasi-spin (a certain superposition of the particle and antiparticle state; Sect. 1.5) and the time the kaon propagates until the measurement. The most general Bell inequality of the CHSH-type is given by

$$S_{k_n,k_m,k_{n'},k_{m'}}(t_1,t_2,t_3,t_4) = \\ \left|E_{k_n,k_m}(t_1,t_2) - E_{k_n,k_{m'}}(t_1,t_3)\right| + \left|E_{k_{n'},k_m}(t_4,t_2) + E_{k_{n'},k_{m'}}(t_4,t_3)\right| \leq 2\,. \tag{45}$$

As in the usual photon setup, Alice and Bob can choose among two settings, i.e., Alice: $\{(k_n, t_1); (k_{n'}, t_4)\}$ and Bob: $(\{(k_m, t_2); (k_{m'}, t_3)\})$. The expectation value $E_{k_n,k_m}(t_1, t_2)$ denotes then that Alice chooses to measure the quasi-spin k_n at time t_1 on the kaon propagating to her side and Bob chooses to measure k_m at time t_2 on his kaon. We notice already that in the neutral kaon case we have more options than in the photon case; we can vary in the quasi-spin space or vary the detection times or both. Furthermore, due to the time evolution and the decay property the kaon system is more complex and involved.

2.3 Variation in the Quasi-spin: What Has $\mathcal{CP}$ Violation to Do with Nonlocality?

Let us first set all times equal to zero and choose the quasi-spin states $k_n = K_S, k_m = \bar{K}^0, k_{n'} = k_{m'} = K_1^0$ (definition of the mass-, strangeness-, $\mathcal{CP}$-eigenstates, see Sect. 1.2), then the inequality (45) turns into the Wigner-type inequality (first considered by F. Uchiyama [51]):

$$P(K_S, \bar{K}^0) \leq P(K_S, K_1^0) + P(K_1^0, \bar{K}^0), \tag{46}$$

where $P(K_S, \bar{K}^0)$, respectively, denotes the probability to measure on the left-hand side a short-lived kaon K_S and on the right-hand side an anti-kaon $\bar{K}^0$.

For the initial antisymmetric Bell state (44) it is straightforward to calculate the quantum mechanical probabilities. However, there is a non-physical phase in the definition of the $\mathcal{CP}$ states for the neutral kaon system which is set to zero in conventional physics. We are interested in the maximal violation of the Bell inequality, thus when we optimize over that unphysical phase inequality (46) can be turned into (see Ref. [50])

$$\delta \leq 0\,, \tag{47}$$

where δ is the $\mathcal{CP}$ violating parameter in mixing. This Bell inequality (47) is experimentally testable!

Experimentally, δ corresponds to the leptonic asymmetry of kaon decays which is measured to be $\delta \approx 10^{-3}$ (see also Sect. 1.6). This value is in clear contradiction to $\delta \leq 0$, the value required by the Bell inequality! In this sense the violation of a symmetry in particle physics is surprisingly related to nonlocality.

Note also that if we interchange $\bar{K}^0$ with K^0 in inequality (46), we obtain $\delta \geq 0$. Thus both BIs induce that the premises of local realistic theories require $\delta = 0$ and hence that the $\mathcal{CP}$ symmetry is conserved.

For the Bell inequality under investigation we assumed that all quasi-spin states are measured at time $t = 0$, i.e., at their creation. This is of course impossible; including the times it turns out that for time greater $10^{-3}\tau_S$ no violation can be obtained, see Ref [6].

Although the Bell inequality (46) sensitive to the small $\mathcal{CP}$ violating parameter δ is as loophole free as possible, the probabilities involved are not directly measurable, because it is experimentally impossible to distinguish between a short-lived state $|K_S\rangle$ and the $\mathcal{CP}$ plus state $|K_1^0\rangle$. Thus a direct experimental verification, i.e., measuring the probabilities, is not possible. One has to consider other Bell inequalities for direct tests.

2.4 Variation in the Detection Times: A Bell Inequality Sensitive to Strangeness

Now let us assume that the detection times are varied and the quasi-spin is chosen in each case to be the anti-kaon state $\bar{K}^0$, thus the CHSH–Bell type inequality turns into

$$S_{\bar{K}^0,\bar{K}^0,\bar{K}^0,\bar{K}^0}(t_1,t_2,t_3,t_4) = \left|E_{\bar{K}^0,\bar{K}^0}(t_1,t_2) - E_{\bar{K}^0,\bar{K}^0}(t_1,t_3)\right| + \left|E_{\bar{K}^0,\bar{K}^0}(t_4,t_2) + E_{\bar{K}^0,\bar{K}^0}(t_4,t_3)\right| \leq 2\,. \tag{48}$$

Starting from the initial antisymmetric maximally entangled Bell state (44) there cannot be found **any violation** for any choice of the four times!

To get an intuition of why this is the case we can consider the ratio of the oscillation to decay given by $x = \frac{\Delta m}{\Gamma} \approx \frac{2\Delta m}{\Gamma_S} \approx 1$. If Nature provided the neutral kaon system with a ratio that would have been greater than 2, the above Bell inequality would be violated (see also Ref. [33]). Differently stated the system oscillates too slow compared to the decay or vice versa. And of course physicists have no tools to change the natural constants of a physical system.

Of course any Bell test in an experiment requires a detailed study of loopholes. But the proof of the existence of correlations which are stronger than those explainable by any local realistic theory is even more limited for mesons. It is only conclusive **iff** the following two main drawbacks are fulfilled (consult also Ref. [33]):

(1) "Active" measurements are required for any test, i.e., opening the possibility to choose among alternative setups or in other words the experimenter needs to have the free choice on the specific question asked to the system.
(2) When describing the time evolution of an unstable quantum system the "information" of the decay products cannot be ignored.

Note that these requirements are **not** loopholes. The first one can only be fulfilled by the neutral kaon system, because the decay rates of the other meson systems are too short to insert a piece of matter; thus only neutral kaons can refute local realistic theories via Bell inequalities. Decay events cannot be considered, clearly they are spontaneous events, i.e., no one has control *when* and into *what kind* of decay products an incoming particle decays. The second drawback is solved by using an appropriate formalism for decaying systems, which was rigourously developed in

Ref. [30] and is reviewed in Sect. 2.5. The physics behind these drawbacks is discussed in detail in Ref. [33].

The main message is that the above expectation values are appropriable for a Bell test in a real experiment, e.g., at DAΦNE machine in Frascati.

Is there really no way to violate the above Bell inequality sensitive to strangeness (48) for a certain initial state and if there is, what would be the maximum value?

Developing the general formalism one learns that decay is a "kind of decoherence" (Ref. [30] and Sect. 2.5). One knows from decoherence studies that some states are more robust against the interaction with the environment, i.e., decoherence, than others. This obviously rises the question whether we can find an initial state which is more robust against the "decoherence" caused by the decay property and thus would enable us to violate the Bell inequality.

The surprising result was indeed YES (Ref. [31])!

It turns out that a non-maximally entangled state violates the Bell inequality maximal. This is surprising because for bipartite qubits, e.g., entangled photons, one knows, that the maximally entangled state[1] gives the maximal violation of the Bell inequality. Kaons are also two-state systems thus we would have expected that they "behave" in a similar manner. Note that for systems with more than two degrees of freedom the maximally entangled state does no longer automatically violate a certain Bell inequality maximally [34]; however, there the situation is also more complicated.

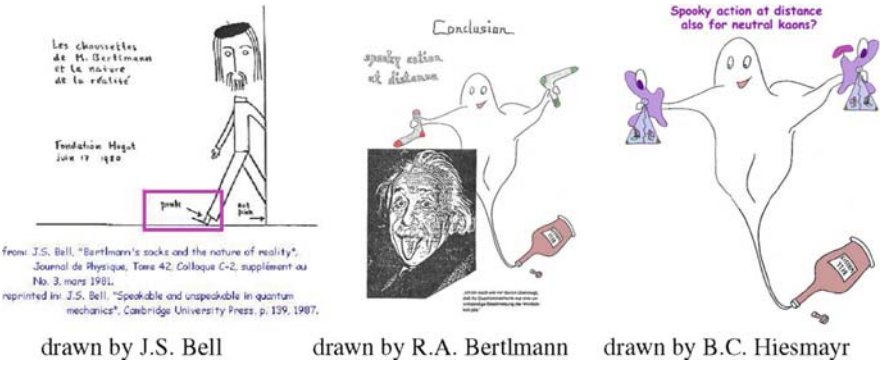


Fig. 1 First one is drawn by John S. Bell showing a young student with the habit to never wear two socks of the same colour. Then celebrating the 60th birthday of Bell this student, Reinhold. A. Bertlmann, draw a little "revenge," which now can be extended to neutral kaons. The Bell whisky drawn in the two last figures exists in reality and can be bought, e.g., in the U.K.

[1] A maximally entangled state is defined by being pure and the trace over one subsystem gives the totally mixed density matrix, the normed unity.

To understand how nonlocality, entanglement and purity for the kaonic quantum system are related we first use the general formalism for decaying systems and then analyze the time evolution of single kaons and bipartite kaons. This enables us to define a proper measure of entanglement and calculate the four expectation values of the Bell inequality sensitive to strangeness (48) and moreover find states for which this Bell inequality is violated. Given that these states can be produced in experiments we would have a conclusive test of refuting local realism for the K-meson system (except for some unavoidable loopholes), i.e., showing that at scales of these energies Einstein's *spooky action at distant* works.

2.5 Open Quantum Formalism of Decaying Systems

In order to discuss basic questions of QM we have to have an appropriate model to describe the time evolution of the meson system, i.e., oscillation ($K^0 \leftrightarrow \bar{K}^0$) and decay. In Ref. [30] it has been shown recently that a decaying system can be handled with the open quantum system formulation, i.e., a master equation of the Lindblad type (Refs. [35, 36] or consult, e.g., Ref. [37]). We now apply this formalism to kaons.

The neutral kaon system is a decaying two-state system due to the strangeness oscillation in time, $K^0 \leftrightarrow \bar{K}^0$, and is usually described via an effective Schrödinger equation which we may write in the Liouville–von Neumann form ($\hbar \equiv 1$):

$$\frac{d}{dt}\rho = -\mathrm{i}\, H_{\mathrm{eff}}\, \rho + \mathrm{i}\, \rho\, H_{\mathrm{eff}}^{\dagger}, \tag{49}$$

where ρ is a 2×2 density matrix and H_{eff} is non-Hermitian. Using the usual Wigner–Weisskopf-approximation the effective Hamilton can be decomposed as $H_{\mathrm{eff}} = H - \frac{\mathrm{i}}{2}\Gamma$ where the mass matrix H and the decay matrix Γ are both Hermitian and positive. To obtain this Hamiltonian the weak interaction Hamilton is treated as a perturbation and the interaction between the final states is neglected, in particular the decay states of the neutral kaons, e.g., pions, are assumed to be stable. The second approximation is to consider only the first pole contribution and this leads to the exponential time evolution of the two diagonal states of H_{eff}:

$$|K_{S/L}(t)\rangle = e^{-\mathrm{i}m_{S/L}t} e^{-\frac{\Gamma_{S/L}}{2}t} |K_{S/L}\rangle\,, \tag{50}$$

$m_{S/L}$ and $\Gamma_{S/L}$ are the masses or decay constants of the short or long-lived kaons.

Considering Eq. (50)—which describes the surviving components of the neutral kaon evolving in time—we notice that the state is not normalized for $t > 0$. Indeed, we are not including all information on the system under investigation. For $t > 0$ the neutral kaon is a superposition of the surviving and decaying components. In Ref. [30] it is shown that by at least doubling the original two-dimensional Hilbert space the decay can be incorporated into the dissipator of the enlarged space via a Lindblad operator. Let us start with the well-known master equation in the Lindblad

form (Refs. [35, 36] or consult, e.g., Ref. [37]):

$$\frac{d}{dt}\rho = -\mathrm{i}[\mathcal{H}, \rho] - \mathcal{D}[\rho]\,, \tag{51}$$

where the dissipator has the general form

$$\mathcal{D}[\rho] = \frac{1}{2}\sum_j (\mathcal{A}_j^\dagger \mathcal{A}_j \rho + \rho \mathcal{A}_j^\dagger \mathcal{A}_j - 2\mathcal{A}_j \rho \mathcal{A}_j^\dagger)\,. \tag{52}$$

In the case under consideration the density matrix is four-dimensional and lives on $\mathbf{H}_{\mathrm{tot}} = \mathbf{H}_s \otimes \mathbf{H}_f$ where s/f denotes "surviving" and "decaying" or "final," respectively, and has the following components:

$$\rho = \begin{pmatrix} \rho_{ss} & \rho_{sf} \\ \rho_{fs} & \rho_{ff} \end{pmatrix},$$

where ρ_{ij} with $i, j = s, f$ denote 2×2 matrices and $\rho_{sf} = \rho_{fs}^\dagger$. The Hamiltonian $\mathcal{H}$ is the Hamiltonian H from the effective Hamiltonian H_{eff} extended to the total Hilbert space $\mathbf{H}_{\mathrm{tot}}$:

$$\mathcal{H} = \begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix}.$$

To incorporate the decay ability we define one Lindblad generator, e.g. A_0, to be

$$\mathcal{A}_0 = \begin{pmatrix} 0 & 0 \\ B & 0 \end{pmatrix} \qquad \text{with} \quad B : \mathbf{H}_s \to \mathbf{H}_f\,,$$

where $B^\dagger B = \Gamma$ and Γ is the decay matrix of H_{eff}. All other Lindblad generators ($j > 0$) can only act on the surviving component of the density matrix, i.e.:

$$\mathcal{A}_j = \begin{pmatrix} A_j & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{with} \quad j \neq 0\,.$$

If the master equation for the density matrix on the total Hilbert space $\mathbf{H}_{\mathrm{tot}}$ is decomposed into the components of ρ it reads

$$\begin{aligned}
\dot{\rho}_{ss} &= -\mathrm{i}[H, \rho_{ss}] - \frac{1}{2}\{\underbrace{B^\dagger B}_{\Gamma}, \rho_{ss}\} - \tilde{D}[\rho_{ss}], \\
\dot{\rho}_{sf} &= -\mathrm{i}H\rho_{sf} - \frac{1}{2}\underbrace{B^\dagger B}_{\Gamma}\rho_{sf} - \frac{1}{2}\sum_j A_j^\dagger A_j \rho_{sf}\,, \\
\dot{\rho}_{ff} &= B\rho_{ss}B^\dagger\,,
\end{aligned}$$

where $\tilde{D}[\rho_{ss}] = \frac{1}{2}\sum_{j=1}(A_j^\dagger A_j \rho + \rho A_j^\dagger A_j - 2A_j \rho A_j^\dagger)$ describes any decoherence or dissipation which may occur (see also Sect. 4).

Indeed, we immediately see that this master equation describes the original effective Schrödinger equation (49) and thus the decay of neutral kaons (and in addition decoherence):

(1) By construction the time evolution of ρ_{ss} is independent of ρ_{sf}, ρ_{fs} and ρ_{ff}.
(2) Furthermore ρ_{sf}, ρ_{fs} completely decouples from ρ_{ss}. Consequently, these components ρ_{sf}, ρ_{fs} can without loss of generality be chosen to be zero; they are not physical and cannot be measured.
(3) With the initial condition $\rho_{ff}(0) = 0$ the time evolution is solely determined by ρ_{ss} and simply given by

$$\rho_{ff}(t) = B \int_0^t dt' \rho_{ss}(t') B^\dagger \,. \tag{53}$$

And indeed, these are the properties which we expect to describe a particle decay!

Summarizing, we showed that particle decay and decoherence/dissipation are related phenomena and that the Wigner Weisskopf approximation is Markovian and completely positive (for a definition see Sect. 4.1).

2.6 Time Evolution of Single Kaons

Without loss of generality the initial state can be chosen in the mass eigenstate basis $\{K_S, K_L\}$. The formal solution of Eq. (51) ($\Gamma = \frac{1}{2}(\Gamma_S + \Gamma_L)$ and the numbers $\rho_{SS} + \rho_{LL} = 1$) is (neglecting $\mathcal{CP}$ violation)

$$\rho(t) = \begin{pmatrix} e^{-\Gamma_S t}\rho_{SS} & e^{-\mathrm{i}\Delta m t - \Gamma t}\rho_{SL} & 0 & 0 \\ e^{\mathrm{i}\Delta m t - \Gamma t}\rho_{SL}^* & e^{-\Gamma_L t}\rho_{LL} & 0 & 0 \\ 0 & 0 & F_L\rho_{LL} & X^* \\ 0 & 0 & X & F_S\rho_{SS} \end{pmatrix}, \tag{54}$$

with $F_{S/L} = 1 - e^{-\Gamma_{S/L} t}$ and $X = \frac{\sqrt{\Gamma_S \Gamma_L}}{-\mathrm{i}\Delta m - \Gamma}(1 - e^{-\mathrm{i}\Delta m t - \Gamma t})\rho_{SL}$. Clearly, we have $\mathrm{Tr}\,\rho(t) = 1$ for all t and the decay is caused by the environment (treating the neutral kaon in QFT formalism, the decay would be caused by the QCD vacuum). The surviving part of the single kaon evolving in time is represented by the upper 2×2 block matrix ρ_{ss}, the lower one by the decaying part ρ_{ff}.

Only properties of the surviving components can be measured, e.g., by a piece of matter an incoming beam is forced to react with the matter via the strong interaction (which is strangeness conserving). If a reaction which can only be caused by a $\bar{K}^0$ is detected, one records a yes-event (Y). If no $\bar{K}^0$ is detected a no-event (N) is recorded (including a K^0 or a decay event). Then matter acts in the very same manner as an ordinary polarisator for photons. Note that an experimenter can *actively* choose the

initial state (up to experimental realization), the kind of detector (experimentally very limited) and where to place the detector, i.e., how much "decoherence" the system undergoes, whereas the kind of "decoherence" is given by Nature. Note that this "decoherence" is fundamentally different from that in other quantum systems which are stable. There the kind of decoherence depends on the environment, for kaons it is intrinsic to the system.

Consequently, an operator P projecting onto the states ρ_{ss} gives the two probabilities, for Y or N, that a certain state is detected at time t:

$$\begin{aligned}
\text{Prob(Y, t)} &= \text{Tr}\left(\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix} \rho(t)\right) = \text{Tr}(P\rho_{ss}(t)) \quad \text{and} \\
\text{Prob(N, t)} &= \text{Tr}\left(\begin{pmatrix} \mathbb{1} - P & 0 \\ 0 & \mathbb{1} \end{pmatrix} \rho(t)\right) \\
&= \text{Tr}((\mathbb{1} - P)\rho_{ss}(t)) + \text{Tr}(\rho_{ff}(t)) \\
&= 1 - \text{Tr}(P\rho_{ss}(t))\,.
\end{aligned}$$

Consequently, the expectation value becomes

$$E_P(t) = \text{Prob(Y, t)} - \text{Prob(N, t)} = 2\ \text{Tr(P}\rho_{\text{ss}}\text{(t))} - 1$$

and is **solely** determined by the surviving component ρ_{ss}!

We considered all possible projectors and the ρ_{ff} enters in the probabilities only via the trace, thus it is clear that the off-diagonal elements of ρ_{ff} are not relevant for any probability we may derive. This leaves a certain ambiguity in defining the decaying components and therefore purity and entanglement. We choose the off-diagonal elements of ρ_{ff} in (54) equal to zero because they give the lowest purity values.

Let us now consider the change of the properties of the state $\rho(t)$ with time by considering the purity defined by

$$\begin{aligned}
\text{Tr}\,\rho(t)^2 &= \text{Tr}(\rho_{ss}(t)^2) + \text{Tr}(\rho_{ff}(t)^2) \\
&= \text{Tr}\,\rho_{ss}(t)^2 + (\text{Tr}\,\rho_{ff}(t))^2 \\
&= \text{Tr}\,\rho_{ss}(t)^2 + (1 - \text{Tr}\,\rho_{ss}(t))^2 \\
&= \rho_{SS}^2(1 - 2e^{-\Gamma_S t} + 2e^{-2\Gamma_S t}) \\
&+\rho_{LL}^2(1 - 2e^{-\Gamma_L t} + 2e^{-2\Gamma_L t}) + 2|\rho_{SL}|^2 e^{-2\Gamma t}\,.
\end{aligned}$$

Note that the second equality sign is only true if the off-diagonal elements of ρ_{ff} vanish. Otherwise, we would add an additional in general time-dependent factor to the definition of the purity (for the formal integration of the order 10^{-2}). Again our definition of the purity is only depending on the surviving components. Starting with an arbitrary initial pure state we see that the decay ability of the system leads

to a decrease in purity for $t > 0$. For K_S or K_L the purity returns to 1 for $t \to \infty$ depending on the decay constants, see Fig. 2. After a time $t/\tau_{S/L} = \ln 2$ the minimal purity of 0.5 of a usual qubit system described by a 2×2 density matrix (trace state) is reached. For other superpositions the purity oscillates to a certain final purity which is not equal to 1. For an initial K^0 or $\bar{K}^0$ we reach the minimal purity of 0.375 at time $t/\tau_S = 401.881$, i.e., about 2/3 of the lifetime of the long-lived state. This is much lower than the purity of a qubit system. Indeed, this decaying system — where only two degrees of freedom can be measured — behaves, with regard to the purity properties, as a system with more degrees of freedom; neutral kaons are more like a double slit evolving in time, see Ref. [38] and Sect. 1.8. Clearly, we could renormalize the purity by choosing appropriate off-diagonal elements of ρ_{ff}.

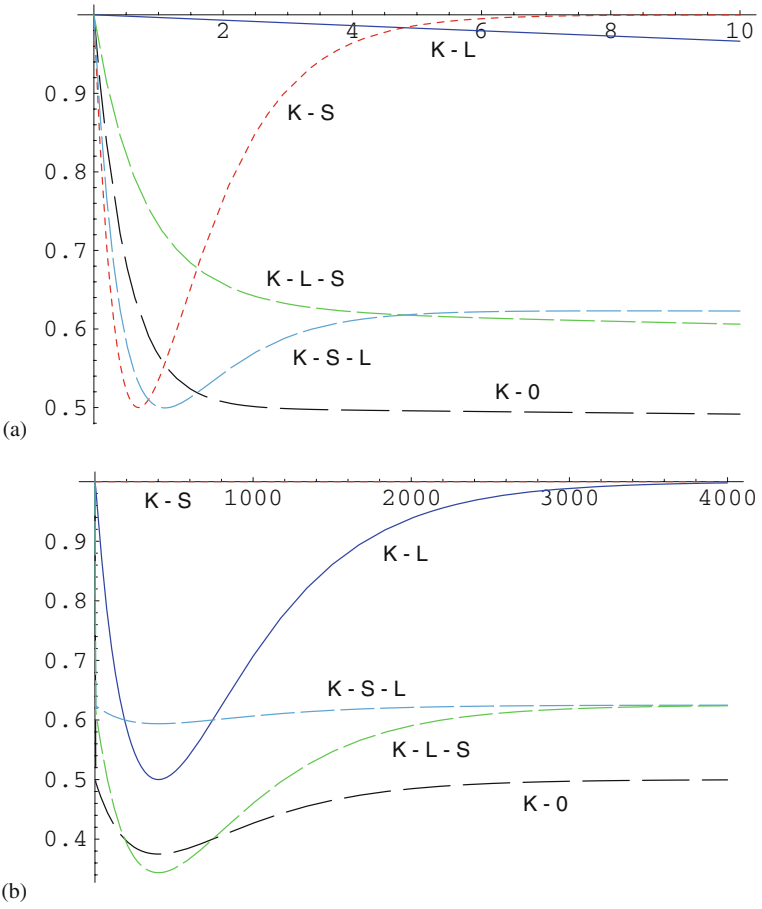


Fig. 2 (Color online) Here the purity $\mathrm{Tr}\,\rho(t)^2$ for single kaons (initially pure) for short (**a**) and longer (**b**) time scales is shown (units in $1/\Gamma_S$): K - S: K_S; K - L: K_L; K - 0: K^0 or $\bar{K}^0$; K - L - S: $1/2|K_S\rangle + \sqrt{3/4}|K_L\rangle$; K - S - L: $\sqrt{3/4}|K_S\rangle + 1/2|K_L\rangle$

The minimal purity which can be reached for this decaying system is 0.333068 and is obtained for an initially mixed state ($\rho_{SS} = 2/3; \rho_{SL} = 0; t/\tau_S = 0.694012$), and is thus greater than 0.25, the minimal value for a 4×4 density matrix (trace state). Note that in general it depends on the ratio between Γ_S/Γ_L and is therefore intrinsic to the described meson system.

2.7 The Time Evolution for Two Kaons

Any density matrix of a single kaon evolving in time, (54), can be decomposed in the following way:

$$\rho(t) = \sum_{nm} f_{nm}(t)\rho_{nm} \, |n\rangle\langle m| \, .$$

Clearly, for two kaons in a product state we have

$$\sigma(t) = \sum_{nmlk} f_{nm}(t) f_{lk}(t)\rho_{nm}\rho_{lk} \, |n\rangle\langle m| \otimes |l\rangle\langle k| \, ,$$

and, consequently, any two-kaon state is then given by

$$\sigma(t) = \sum_{nmlk} f_{nm}(t) f_{lk}(t)\sigma_{nmlk} \, |n\rangle\langle m| \otimes |l\rangle\langle k| \, , \tag{55}$$

where the time-dependent weights can be assumed to factorize. In order to do this one has to prove that the projectors commute with the generators of the time evolution under the trace (this was proven in a different formulation in Ref. [39]). We can even define a two-particle density matrix depending on the two different times representing the times when the two kaons are measured, i.e.,

$$\sigma(t_l, t_r) = \mathrm{diag}\{\sigma_{ssss}(t_l, t_r), \sigma_{ssff}(t_l, t_r), \sigma_{ffss}(t_l, t_r), \sigma_{ffff}(t_l, t_r)\} \, ,$$

where σ_{iijj} are 4×4 matrices.

We are now interested in the entanglement of such a density matrix. As a measure of entanglement we want to consider the entanglement of formation which is defined by

$$\mathcal{E}o\mathcal{F}(\rho) = \min_i \sum_i p_i S(\mathrm{Tr}_l(|\psi_i\rangle\langle\psi_i|)) \, ,$$

where S is the von Neumann entropy, the trace is taken over one subsystem (left or right) and ψ_i are the pure state decompositions of ρ. The von Neumann entropy is defined by

$$S(\rho) = -\,\mathrm{Tr}(\rho \ln \rho) \, .$$

A sufficient condition for entanglement is that under partial transpose (PT) the matrix has at least one negative eigenvalue. Only for bipartite two-level systems and for 2×3 systems, PT is also necessary for detecting all entangled states. For the density matrix under investigation PT acts in the following way:

$$PT[\sigma(t_l, t_r)] = \mathrm{diag}\{PT[\sigma_{ssss}(t_l, t_r)],\ PT[\sigma_{ssff}(t_l, t_r)],\\ PT[\sigma_{ffss}(t_l, t_r)],\ PT[\sigma_{ffff}(t_l, t_r)]\}\,.$$

The surviving–surviving block σ_{ssss} can lead to negative eigenvalues, i.e., it can be entangled, while the eigenvalues of the other blocks cannot become negative due to the vanishing off-diagonal elements; the eigenvalues remain unchanged under PT. Thus, whether the state under investigation is entangled depends only on σ_{ssss}. For 4×4 matrices entanglement of formation is an increasing function of the computable concurrence $\mathcal{C}$, found by Hill and Wootters [40].[2] Thus we can measure entanglement by the concurrence of σ_{ssss}.

To compute concurrence one defines the flipped matrix $\tilde{\sigma}_{ssss} = (\sigma_y \otimes \sigma_y)\sigma^*_{ssss}(\sigma_y \otimes \sigma_y)$ where σ_y is the y-Pauli matrix and the complex conjugation is taken in the $K_S K_L$ basis. The concurrence is then given by the formula $\mathcal{C} = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$ where the λ_is are the square roots of the eigenvalues, in decreasing order, of the matrix $\sigma_{ssss}\tilde{\sigma}_{ssss}$.

Let us now consider a general pure state at $t = 0$ (with $r_1^2 + r_2^2 + r_3^2 + r_4^2 = 1$):

$$|\psi(0)\rangle = r_1 e^{\mathrm{i}\phi_1}|K_S\rangle|K_S\rangle + r_2 e^{\mathrm{i}\phi_2}|K_S\rangle|K_L\rangle\\ + r_3 e^{\mathrm{i}\phi_3}|K_L\rangle|K_S\rangle + r_4 e^{\mathrm{i}\phi_4}|K_L\rangle|K_L\rangle\,. \tag{56}$$

Alice and Bob perform their measurements at certain times t_l, t_r, respectively. For a general initial pure state the concurrence is derived to be

$$\mathcal{C}(\sigma_{ssss}(t_l, t_r)) = 2\,|r_1 r_4 e^{\mathrm{i}\phi_1 + \mathrm{i}\phi_4} - r_2 r_3 e^{\mathrm{i}\phi_2 + \mathrm{i}\phi_3}|\, e^{-\Gamma(t_l + t_r)}\,.$$

It is simply the concurrence of the initial pure state multiplied by the time-dependent damping factor. For one time equal to zero the decrease in entanglement is lowest.

2.8 The States Violating the Bell Inequality Sensitive to Strangeness

For the two-particle state we obtain four probabilities:

[2] For higher dimensions no computable function of entanglement of formation is known.

$$
\begin{aligned}
\mathrm{Prob}(Y,t_l;Y,t_r) &= \mathrm{Tr}(P_l \otimes P_r \sigma_{ssss})\,,\\
\mathrm{Prob}(Y,t_l;N,t_r) &= \mathrm{Tr}((1-P_l)\otimes \mathbb{1}(\sigma_{ssss}+\sigma_{ssff}))\\
\mathrm{Prob}(N,t_l;Y,t_r) &= \mathrm{Tr}(\mathbb{1}\otimes(1-P_r)(\sigma_{ssss}+\sigma_{ffss}))\\
\mathrm{Prob}(N,t_l;N,t_r) &= 1+\mathrm{Tr}(P_l\otimes P_r\sigma_{ssss})\\
&\quad -\mathrm{Tr}((1-P_l)\otimes\mathbb{1}\sigma_{ssff})-(1-P_r)(\sigma_{ffss}),
\end{aligned}
$$

and the expectation value is given by

$$
\begin{aligned}
E_{P_l,P_r}(t_l,t_r) &= \mathrm{Prob}(Y,t_l;Y,t_r)+\mathrm{Prob}(N,t_l;N,t_r)-\mathrm{Prob}(Y,t_l;N,t_r)-\\
&= \mathrm{Prob}(N,t_l;Y,t_r)-1\\
&\quad +2(\mathrm{Prob}(Y,t_l;Y,t_r)+\mathrm{Prob}(N,t_l;N,t_r)),
\end{aligned}
\tag{57}
$$

where we used the fact that the sum of all probabilities gives one.

As projectors we choose $P_{r,l} = |\bar{K}^0\rangle\langle\bar{K}^0|$, and, after a cumbersome calculation, the expectation value is

$$
\begin{aligned}
E_{\bar{K}^0,\bar{K}^0}(t_l,t_r) = \;& 1+r_1^2\,e^{-\Gamma_S(t_l+t_r)}+r_2^2\,e^{-\Gamma_S t_l-\Gamma_L t_r}\\
&+r_3^2\,e^{-\Gamma_L t_l-\Gamma_S t_r}+r_4^2\,e^{-\Gamma_L(t_l+t_r)}\\
&-r_1^2\,(e^{-\Gamma_S t_l}+e^{-\Gamma_S t_r})-r_2^2\,(e^{-\Gamma_S t_l}+e^{-\Gamma_L t_r})\\
&-r_3^2\,(e^{-\Gamma_L t_l}+e^{-\Gamma_S t_r})-r_4^2\,(e^{-\Gamma_L t_l}+e^{-\Gamma_L t_r})\\
&+2\,r_1r_2\,(1-e^{-\Gamma_S t_l})\cos(\Delta m t_r+\phi_1-\phi_2)\,e^{-\Gamma t_r}\\
&+2\,r_1r_3\,\cos(\Delta m t_l+\phi_1-\phi_3)\,e^{-\Gamma t_l}\,(1-e^{-\Gamma_S t_r})\\
&+2\,r_2r_4\,\cos(\Delta m t_l+\phi_2-\phi_4)\,e^{-\Gamma t_l}\,(1-e^{-\Gamma_L t_r})\\
&+2\,r_3r_4\,(1-e^{-\Gamma_L t_l})\,\cos(\Delta m t_r+\phi_3-\phi_4)\,e^{-\Gamma t_r}\\
&+2\,r_1r_4\cos(\Delta m(t_l+t_r)+\phi_1-\phi_4)\,e^{-\Gamma(t_l+t_r)}\\
&+2\,r_2r_3\,\cos(\Delta m(t_l-t_r)+\phi_2-\phi_3)\,e^{-\Gamma(t_l+t_r)}\,.
\end{aligned}
\tag{58}
$$

We notice that for any initial state one always has damping functions from the decay property in this system different from other two-state systems and the expectation value converges for both times to infinity to $+1$. For the initial maximally entangled Bell states $|\phi^{\pm}\rangle = \frac{1}{2}\{|K_S K_S\rangle \pm |K_L K_L\rangle\}$ ($r_2 = r_3 = 0$) the oscillation goes with the sum of the times, different from the maximally entangled Bell states $\psi^{\pm}\frac{1}{2}\{|K_S K_L\rangle \pm |K_L K_S\rangle\}$ ($r_1 = r_4 = 0$) where the oscillation only depends on the difference of the times. Thus for $\phi^{\pm}$ a violation of the Bell inequality could occur earlier. Only the ϕ^- violates the Bell inequality slightly $S = 2.07$, but there are states violating it even more.

For all phases $\phi_i = 0$ we find the value

$$S = 2.1175$$

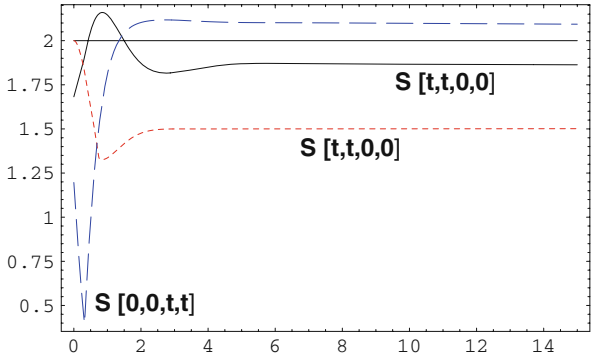


Fig. 3 In Fig. (**a**) the time-dependent S-function for the initial state $|\phi^+\rangle = \frac{1}{2}\{|K_S K_S\rangle + |K_L K_L\rangle\}$ (*dashed*), the state ξ defined in the text (*long dashed*) and the state χ also defined in the text (*solid*) (time in units of $[\frac{\Gamma_S}{\Delta m}]$) is shown. For the initial state ξ we obtain the maximum violation. For the initial state ξ the violation exists up to $1/\Gamma_L$

(state ξ ($r_1 = -0.8335; r_2 = r_3 = -0.2446; r_4 = 0.4308$): $t_1 = t_2 = 0; t_3 = t_4 = 5.77\tau_S$). If we also vary over the phases we obtain a slightly higher value

$$S = 2.1596$$

(state χ ($r_1 = -0.7823; r_2 = r_3 = 0.1460; r_4 = 0.5877; \phi_1 = -0.2751; \phi_2 = \phi_3 = -0.6784; \phi_4 = 0$): $t_1 = t_2 = 1.79\tau_S; t_3 = t_4 = 0$), see also Fig. 3. For the above cases the concurrence gives

$$\mathcal{C}(\xi) = 0.84\, e^{-\Gamma(t_l+t_r)} \quad \text{and} \quad \mathcal{C}(\chi) = 0.94\, e^{-\Gamma(t_l+t_r)}\,.$$

In Fig. 4(a)–(c) purity versus concurrence diagrams are drawn. For ϕ^+ we notice that the "decoherence" caused by the decay exceeds the purity-concurrence values of Werner states, which represent an upper limit for all possible decoherence modes in this picture given by a Lindblad equation for an initially maximally entangled qubit state, Ref. [48, 49]. An early decay of one kaon, Fig. 4(b), exceeds even the purity-concurrence value of maximally entangled mixed bipartite qubit states (MEMS) [41].

To sum up, the initial entanglement decreases with a sum of times, and it goes hand in hand with a decrease in purity first which can then for latter times increase again. For non-maximal entangled state the decrease of purity is much faster than for the maximally entangled states. This seems to help to violate the Bell–CHSH inequality though the ratio of oscillation to decay is low.

We show how to treat a single and bipartite decaying neutral kaon system in quantum mechanics and analyze the states via purity, entanglement and nonlocality. Only two degrees of freedom at a certain time can be measured reducing the set of observables and leaving some elements of the state undefined.

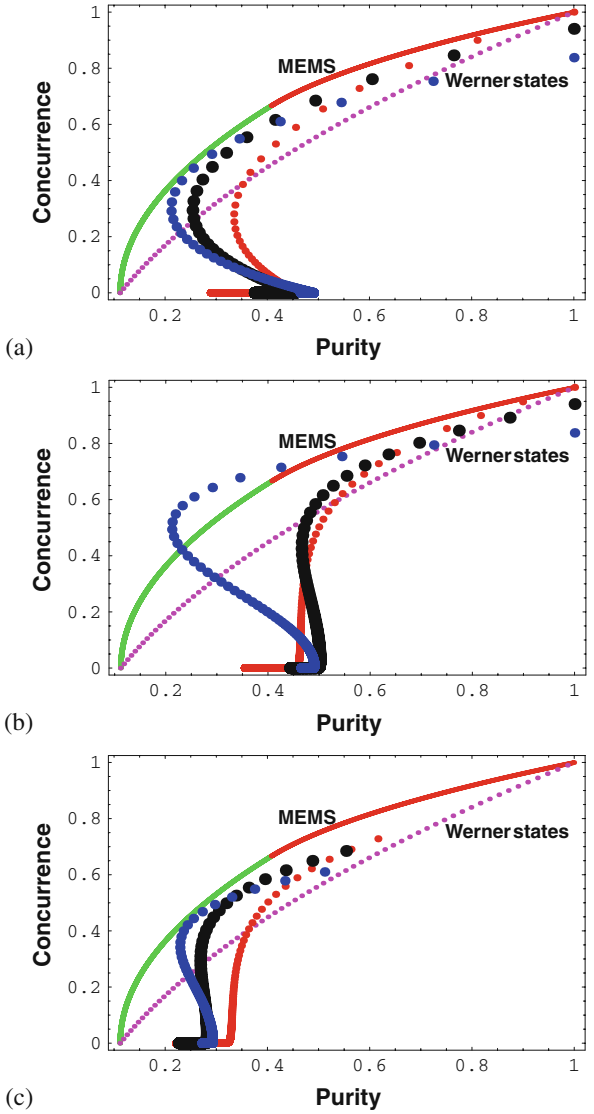


Fig. 4 (Colour online) In Fig. (**a**)–(**c**) a purity versus concurrence diagram is drawn (purity normalized $(dTr\rho^2 - 1)/(d - 1)$ with $d = 4$ for bipartite qubits and $d = 16$ for bipartite kaons). The limiting curve represents the maximally entangled mixed bipartite qubit states (MEMS) [41] and the nearly linear curve (*dashed, purple*) the Werner states for bipartite qubits. The dots are drawn for different initial states and the time proceeds from 0 to 100 with a step width of 0.05 (units as above). The smallest dots (*red*) for ϕ^+, next to smallest dots (*blue*) ξ and the biggest dots (*black*) are for χ. In Fig. (**b**) is shown the change in purity and concurrence for $t_l = t_r = t$, in (**c**) for $t_l = 0; t_r = t$ and in (**d**) for $t_l = 0.3; t_r = t$

In contrast to the case of photons, for the neutral kaon system nonlocaliy is a quite "dynamical" concept as correlations of states evolving up to different times are involved. For entangled photons there is no difference whether in principle the correlations are measured after one or several meters. With each measurement the experimenter chooses among two observables: the quasi-spin and the detection time. Consequently, considering Bell inequalities for mesons, (45), one can vary in the quasi-spin space or vary the detection times or both. If varying in the quasi-spins space and for simplicity choosing all times equal to zero, there is a connection between nonlocality and the violation of a symmetry in high-energy physics, i.e., the $\mathcal{CP}$ symmetry ($\mathcal{C}$=charge conjugation, $\mathcal{P}$=parity), Ref. [50]. Though the violation of this Bell inequality sensitive to $\mathcal{CP}$ violation is obtained for a single parameter δ, which is measured even with single kaons, a direct experimental verification, i.e., measuring the probabilities involved, is not possible.

Choosing a quasi-spin, i.e., observing on both sides an anti-kaon or not, and varying the times gives a Bell inequality sensitive to strangeness (48). Via inserting a piece of matter at a certain position from the source (corresponding to the detection time) this Bell inequality can in principle be realized experimentally. One obtains maximal violations for certain initial states—currently not available by experiments—for which the concurrence is not maximal in partial agreement with Ref. [34] that optimal Bell tests do not require maximally entangled states for systems with more than two degrees of freedom. A higher amount of entanglement does not necessarily imply an increase of a violation of the Bell inequality under investigation; in fact, the Bell inequality need not be violated at all due to the decay property.

Therefore, these results suggest that for the neutral kaon system nonlocality and entanglement are indeed some distinct quantum features which manifest themselves in a way different than for bipartite qubit or qutrit systems, and their relation is subtler than one naively expects.

There exist other ideas to find testable Bell inequalities, see, e.g., Refs. [42–47] and reference therein.

3 Kaonic Quantum Eraser

Two hundred years ago Thomas Young taught us that photons interfere. Nowadays, experiments with very massive particles, like the fullerenes, have demonstrated this fundamental feature of quantum mechanics [52] impressively. It seems that there is no physical reason why not even heavier particles should interfere except for technical ones. In Sect. 1.8 we have shown that the knowledge on the path through the double slit is the reason why interference is lost. The gedanken experiment of

Scully and Drühl in 1982 [53] surprised the physics community, if the knowledge on the path of the particle is erased, interference is brought back again.

Since that work many different types of quantum erasures have been analyzed and experiments were performed with atom interferometers [54] and entangled photons [55–60] where the quantum erasure in the so-called "delayed choice" mode captures best the essence and the most subtle aspects of the eraser phenomenon. In this case the meter, the quantum system which carries the mark on the path taken, is a system spatially separated from the interfering system which is generally called the object system. The decision *to erase or not* the mark of the meter system—and therefore *to observe or not* interference—can be taken long after the measurement on the object system has been completed. This was nicely phrased by Aharonov and Zubairy in their review article [61] as "erasing the past and impacting the future."

Here we want to present four different types of quantum erasure concepts for neutral kaons, Refs. [8, 62]. Two of them are analogous to performed erasure experiments with entangled photons, e.g. Refs. [55, 56], sketched in Figs. 5 and 7. In the first experiment with photons the erasure operation is carried out "actively," i.e., by exerting the free will of the experimenter, whereas in the latter experiment the erasure operation is carried out "partially actively," i.e., the mark of the meter system is erased or not by a well-known probabilistic law, e.g., by a beam splitter. However, different to photons the kaons can be measured by an *active* or a *passive* procedure (see Sect. 1.6).

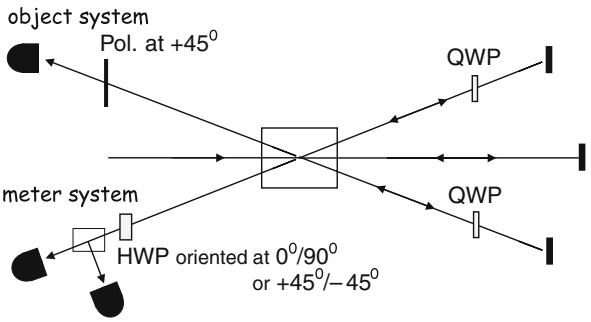


Fig. 5 The setup for an active eraser is sketched. A bump beam transverses twice a, e.g., type II crystal. The pairs produced in the first passage through the crystal cross two times a quarter-wave plate (QWP) which transforms an original horizontal polarized photon into a vertical one and vice versa. The pairs produced in the second passage through the crystal is directly directed to the measurement devices. The signal (object) photon is always measured after crossing a polarization analyzer aligned at $+45^\circ$. The idler (meter) photon crosses a half-wave plate (HWP) oriented at 0°, 90° (first setup) or $\pm 45^\circ$ (second setup) and is then analyzed by a polarization beam splitter. In the first setup—meter photon is measured in the H/V basis—one has full *which way* information, namely if the pair was produced at the first or second passage. In the second setup—meter photon is measured in the $+45^\circ/-45^\circ$ basis—the information on the first or second passage is erased: one observes fringes or antifringes

This offers new quantum erasure possibilities which can only be achieved with kaons. And moreover proves the very concept of a quantum eraser, namely sorting events to available information.

Discussion with the experimenters of the DAϕNE machine in Frascati on how to realize all four different possibilities described in the following are ongoing.

Active eraser with *active* measurements (S: *active*/*active*; T: *active*)

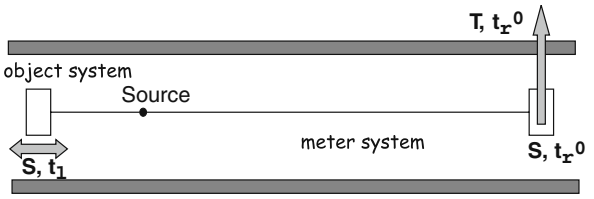


Fig. 6 The figure shows the two setups for an eraser with active marking and *active* measurements

3.1 Active Eraser with Active *Measurements*

Let us first discuss the photon analogy, e.g., the two experimental setups in Ref. [55]. In the first setup two interfering two-photon amplitudes are prepared by forcing a pump beam to cross twice the same nonlinear crystal. Idler and signal photons from the first down conversion are marked by rotating their polarization by 90° and then superposed to the idler (i) and signal (s) photons emerging from the second passage of the beam through the crystal. If type-II spontaneous parametric down conversion were used, we had the state [3]

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left\{ \underbrace{|V\rangle_i|H\rangle_s}_{\text{second passage}} - e^{\mathrm{i}\Delta\phi} \underbrace{|H\rangle_i|V\rangle_s}_{\text{first passage}} \right\}, \tag{59}$$

where the relative phase $\Delta\phi$ is under control by the experimenter (the symbol $\otimes$ for the tensor product of the states is dropped from now on). The signal photon, the object system, is always measured after crossing a polarization analyzer aligned at $+45°$, see Fig. 5. Due to entanglement, the vertical or horizontal idler polarization supplies full *which path* information for the signal (object) system, i.e., whether it was produced at the first or second passage. No interference can be observed in the signal–idler joint detections. To erase this information, the idler photon has to be detected in the $+45°/-45°$ basis.

In case of entangled kaons the state is described by (25). The analogy with (59) is quite obvious; however, kaons evolve in time, such that the state depends on the

[3] The authors of Ref. [55] used type-I crystals in their experiment.

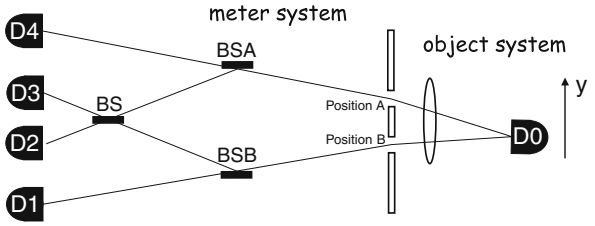


Fig. 7 The setup of a partially active eraser is sketched. An entangled photon pair can be produced either in region A or in region B. If the detectors $D1$ or $D2$ clicks, one knows the production region A or B, i.e., one has full *which path* information. Clicks of the detectors $D2$ or $D3$ erase the information; interference is observed. It is a partially active eraser, because the mark is erased by a probabilistic law; however, the experimenter has still partial control over the erasure and she/he can choose the ratio of transmittivity to reflectivity of the beam splitter BSA and BSB

two time measurements on the left-hand side, t_l, and on the right-hand side, t_r, or more precise on $\Delta t = t_l - t_r$, when normalized[4] to surviving kaon pairs:

$$\begin{aligned}|\phi(\Delta t)\rangle &= \frac{1}{\sqrt{1+e^{\Delta\Gamma\Delta t}}}\left\{|K_L\rangle_l|K_S\rangle_r - e^{\mathrm{i}\Delta m\Delta t}e^{\frac{1}{2}\Delta\Gamma\Delta t}|K_S\rangle_l|K_L\rangle_r\right\}\\ &= \frac{1}{2\sqrt{1+e^{\Delta\Gamma\Delta t}}}\left\{\left(1-e^{\mathrm{i}\Delta m\Delta t}e^{\frac{1}{2}\Delta\Gamma\Delta t}\right)\{|K^0\rangle_l|K^0\rangle_r - |\bar{K}^0\rangle_l|\bar{K}^0\rangle_r\}\right.\\ &\qquad\left.+\left(1+e^{\mathrm{i}\Delta m\Delta t}e^{\frac{1}{2}\Delta\Gamma\Delta t}\right)\{|K^0\rangle_l|\bar{K}^0\rangle_r - |\bar{K}^0\rangle_l|K^0\rangle_r\}\right\}.\end{aligned} \tag{60}$$

We notice that the phase $\Delta m\Delta t$ introduces automatically a time-dependent relative phase between the two amplitudes. The marking and erasure operations can be performed on entangled kaon pairs as in the optical case discussed above. The object kaon flying to the left hand side is measured always *actively* in the strangeness basis, see Fig. 6. As in the optical version the kaon flying to the right hand side, the meter kaon, is measured *actively* either in the strangeness basis by placing a piece of matter in the beam or in the "effective mass" basis by removing the piece of matter. Both measurements are *actively* performed. In the latter case we obtain information about the lifetime, namely *which width* the object kaon has, and clearly no interference in the joint detections can be observed.

3.2 Partially Passive Quantum Eraser with Active *Measurements*

Partially active eraser with *active* measurements (S: *active/active*; T: *active*)

[4] Thanks to this normalization, we work with bipartite two-level quantum systems like polarization entangled photons or entangled spin-1/2 particles. For an accurate description of the time evolution of kaons and its implementation consult Ref. [6].

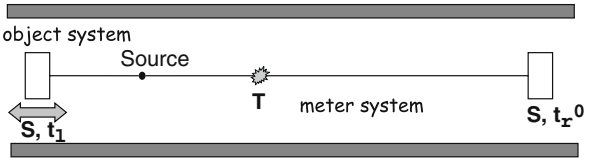


Fig. 8 The figure shows the setup for a partially passive quantum marking and *active* measurements on both sides

In Fig. 7 a setup realized in Ref. [56] is sketched where an entangled photon pair is produced either at position A or B. "Clicks" on detector $D1$ or $D4$ provide "which path" information. "Clicks" on detector $D2$ and $D3$ give no information about the position A or B; interference is observed in the joint events of the two photons, see Fig. 7.

For kaons a piece of matter is permanently inserted into both beams where the one for the meter system at the right hand side is fixed at time t_r^0, see Fig. 8. The experiment observes the region from the source to the piece of matter at the right hand side. In this way the kaon moving to the right—the meter system—takes the choice to show "which width" information by its decay during its free propagation until t_r^0 or not by being absorbed in the piece of matter. Again strangeness or lifetime is measured *actively*. The choice whether the "wave-like" property or the "particle-like" property is observed is naturally given by the instability of the kaons. It is "partially active," because the experimenter can choose at which fixed time t_r^0 the piece of matter is inserted. This is analogous to the optical case where the experimenter can choose the transmittivity of the two beam-splitters BSA and BSB in Fig. 7.

Furthermore, note that it is not necessary to identify K_S versus K_L for demonstrating the quantum marking and eraser principle.

3.3 Passive Eraser with "Passive" Measurements on the Meter

Passive eraser with *passive* measurements on the meter (S: *active*/*passive*; T: *passive*)

Next we consider the setup in Fig. 9. We take advantage—and this is specific for kaons—of the *passive* measurement. Again the strangeness content of the object system—kaon moving to the left hand side—is *actively* measured by inserting a

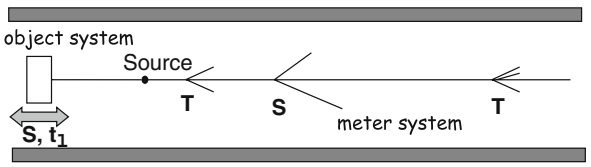


Fig. 9 The figure shows a setup of a quantum eraser which has no photon analog

piece of matter into the beam. In the beam of the meter no matter is placed in, the kaon moving to the right propagates freely in space. This corresponds to a *passive* measurement of either strangeness or lifetime on the meter by recording the different decay modes of neutral kaons. If a semileptonic decay mode is found, the strangeness content is measured. In the joint events interference is observed. If a two or three π decay is observed, the lifetime is observed and thus "which width" information of the object system is obtained; no interference is seen in the joint events. Clearly, we have a completely passive erasing operation on the meter; the experimenter has no control whether the lifetime mark is read out or not.

This experiment has no analog to any other considered two-level quantum system.

3.4 Passive Eraser with "Passive" Measurements

Passive eraser with *passive* measurements (S: *passive/passive*; T: *passive/passive*)

Finally we mention the setup in Fig. 10, where both kaons evolve freely in space and the experimenter observes *passively* their decay modes and times. The experimenter has no control over individual pairs neither on which of the two complementary observables at each kaon is measured nor when it is measured. This setup is totally symmetric, thus it is not clear which side plays the role of the meter and strictly speaking we cannot consider this experiment as a quantum eraser.

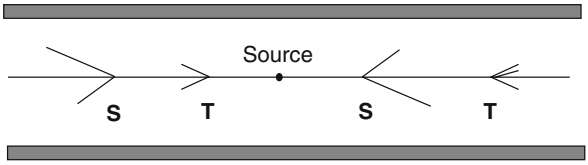


Fig. 10 For this type of quantum eraser, it is not clear which side plays the meter/object role as it is totally symmetric and involves only *passive* measurements. This clearly has no analog to photon experiments

3.5 Conclusions

We have discussed the possibilities offered by neutral kaon states, such as those copiously produced by ϕ-resonance decays at the DAΦNE machine, to investigate two fundamental issues of quantum mechanics: quantitative Bohr's complementarity and quantum eraser phenomena. In both cases, the use of neutral kaons allows for a clear conceptual simplification and to obtain the relevant formulae in a transparent and non-controversial way.

A key point is that neutral kaon propagation through the K_S and K_L components automatically parallels most of the effects of double slit devices. Thanks to this,

Bohr's complementarity principle can be quantitatively discussed in the most simple and transparent way. Similarly, the relevant aspects of quantum marking and the quantum eraser admit a more clear treatment with neutral kaons than with other physical systems. This is particularly true when the eraser is operated in the 'delayed choice' mode and contributes to clarify the eraser's working principle. Moreover, the possibility of performing passive measurements, a specific feature of neutral kaons not shared by other systems, has been shown to open new options for the quantum eraser. In short, we have seen that, once the appropriate neutral kaon states are provided as in the DAΦNE machine, most of the additional requirements to investigate fundamental aspects of quantum mechanics are automatically offered by Nature for free.

The CPLEAR experiment [71] did only part of the job (*active* strangeness–strange-ness measurements), but the KLOE 2 experiment could do the full program (and concrete designs are already under investigations)!

4 Decoherence and Measures of Entanglement

Another approach to study quantum features in high-energy physics is to study systems as open quantum systems, i.e., systems which interact with their environment (see, e.g., [37, 63, 64]). This turned out to be very fruitful, in particular when one discusses entanglement of meson–antimeson systems. First, because within a good developed theoretical framework one can develop different decoherence models with different properties, and second, these models and hypotheses can be put to test by experiments (and this is currently done/has been done). Here, any meson system is appropriate, but currently only for K-mesons and B-mesons enough data is available.

4.1 Short Introduction to Decoherence

The dynamics of closed quantum systems are covered by the Schrödinger equation. These are systems which do not suffer any unwanted interactions with the outside world—generally termed *environment* or *reservoir*—and are consequently described by unitary dynamics. However, in the real world it is often not possible to avoid interactions of the system of interest with other (classical or quantum) systems; such systems are called *open* quantum systems. In the last years the understanding of open quantum systems has enlarged impressively (see, e.g., Refs. [37, 63, 64]). The whole system $S + E$, where S is the quantum system under investigation and E is the environment, is assumed to be closed and thus the Schrödinger equation applies. Possible decoherence effects arise due to some interactions of the system with its environment. Sources for "standard" decoherence effects can be the experimental background or noise of the experimental setup. "Nonstandard" decoherence effects may result from fundamental "modifications"

of QM and may be traced back to the influence of quantum gravity [65]—quantum fluctuations in the space-time structure on the Planck mass scale—or due to dynamical state reductions [66, 67] and arise on a different energy scale. However, the general formalism for open quantum systems, i.e., a master equation derived under certain assumptions, allows to handle decoherence models without modeling the environment explicitly. The main goal is to develop different decoherence scenarios and put them to test experimentally.

The dynamics of an open system S is generally described by a map $\rho_S \longrightarrow \Lambda[\rho_S]$ from the state space of the quantum system onto itself. The map has to satisfy the semigroup law and has to be *completely* positive. Complete positivity means that for a positive map Λ all extensions $\Lambda_n = \Lambda \otimes \mathbb{1}_n$ defined on $\mathcal{H} \otimes \mathbb{C}^n$ for all n are also positive. This is a very important mathematical requirement. Physically Λ_n can be interpreted as an operation which acts locally on one subsystem and no action is taken on the other part of the system. Up to now this mathematical requirement has never been tested directly in experiments, though it is obviously a necessary constraint for all physical situations consisting of combined systems.

4.2 Decoherence Models Which can be Tested with Data from Accelerators

Before we develop a decoherence model in the open quantum formalism let us introduce a phenomenological parameter ζ which goes back to an idea of Schrödinger [1] and Furry [68] in the year 1935. It goes like follows.

Imagine that the source produces as an initial state the maximally entangled antisymmetric Bell state

$$\begin{aligned} |\psi^-(t)\rangle &= \frac{1}{2}\{|K_S\rangle_l \otimes |K_L\rangle_r - |K_L\rangle_l \otimes |K_S\rangle_r\} \\ &= \{|B_H\rangle_l \otimes |B_L\rangle_r - |B_L\rangle_l \otimes |B_H\rangle_r\}, \end{aligned} \tag{61}$$

where the first line describes kaons and the second line B-mesons (B_H,"heavy" mass eigenstate; B_L,"light" mass eigenstate). Then, while the two mesons propagate in different directions the wave function spontaneously factorize in 50% of the cases into the product state

$$|K_S\rangle_l \otimes |K_L\rangle_r \,; |B_H\rangle_l \otimes |B_L\rangle_r \,, \tag{62}$$

respectively, or in the other 50% of the cases into

$$|K_L\rangle_l \otimes |K_S\rangle_r \,; |B_L\rangle_l \otimes |B_H\rangle_r \,. \tag{63}$$

Do Experimental Data Rule Out Such Scenarios or Not?

To answer these questions let us extend the Schrödinger–Furry hypothesis by introducing an "effective" decoherence parameter by

$$
\begin{aligned}
&P^{\text{without decoherence}}(f_1 t_l,\, f_2 t_r) = \\
&\quad || \langle f_1|_l \langle f_2|_r \frac{1}{\sqrt{2}\,\mathrm{Det}S} \Big\{ |\tilde{k}_1(t_l)\rangle_l |\tilde{k}_2(t_r)\rangle_r - |\tilde{k}_2(t_l)\rangle_l |\tilde{k}_1(t_r)\rangle_r \Big\} ||^2 \\
&= \frac{1}{2\,|\mathrm{Det}S|^2} \Big\{ |\langle f_1|\tilde{k}_1(t_l)\rangle|^2 |\langle f_2|\tilde{k}_2(t_r)\rangle|^2 + |\langle f_1|\tilde{k}_2(t_l)\rangle|^2 |\langle f_2|\tilde{k}_1(t_r)\rangle|^2 \\
&\qquad\qquad - 2\mathrm{Re}\Big\{ \langle f_1|\tilde{k}_1(t_l)\rangle^* \langle f_2|\tilde{k}_2(t_r)\rangle^* \langle f_1|\tilde{k}_2(t_l)\rangle \langle f_2|\tilde{k}_1(t_r)\rangle \Big\} \Big\} \\
&\longrightarrow P_{\zeta_{\tilde{k}_1,\tilde{k}_2}}(f_1 t_l,\, f_2 t_r) = \\
&\quad || \langle f_1|_l \langle f_2|_r \frac{1}{\sqrt{2}\,\mathrm{Det}S} \Big\{ |\tilde{k}_1(t_l)\rangle_l |\tilde{k}_2(t_r)\rangle_r - |\tilde{k}_2(t_l)\rangle_l |\tilde{k}_1(t_r)\rangle_r \Big\} ||^2 \\
&= \frac{1}{2\,|\mathrm{Det}S|^2} \Big\{ |\langle f_1|\tilde{k}_1(t_l)\rangle|^2 |\langle f_2|\tilde{k}_2(t_r)\rangle|^2 + |\langle f_1|\tilde{k}_2(t_l)\rangle|^2 |\langle f_2|\tilde{k}_1(t_r)\rangle|^2 \\
&\qquad\qquad - 2\underbrace{(1-\zeta_{\tilde{k}_1,\tilde{k}_2})}_{\text{modification}} \mathrm{Re}\Big\{ \langle f_1|\tilde{k}_1(t_l)\rangle^* \langle f_2|\tilde{k}_2(t_r)\rangle^* \langle f_1|\tilde{k}_2(t_l)\rangle \langle f_2|\tilde{k}_1(t_r)\rangle \Big\} \Big\}\,,
\end{aligned}
\tag{64}
$$

where $f_{1/2}$ are the final states measured in the experiment and $\tilde{k}_{1/2}$ are, e.g., the mass eigenstates of the kaons or the B-mesons, respectively.

Here, the quantum mechanical probability is changed by multiplying the interference term by $1-\zeta$. If ζ vanishes then we have a scenario without decoherence; if $\zeta = 1$ the interference term vanishes, this is clearly the Schrödinger–Furry hypothesis. Thus ζ quantifies how good the quantum mechanical interference term is measured. This effective decoherence parameter—introduced by hand—interpolates continuously between two limits and represents a measure for the amount of decoherence which results in a loss of entanglement of the initial state.

However, we can write down the initial state in any basis we want (QM is basis-independent). Thus we could assume also a spontaneous factorization in a total different basis, e.g., in 50% of the cases into $|K^0\rangle_l \otimes |\bar{K}^0\rangle_r$ or in the other 50% of the cases into $|\bar{K}^0\rangle_l \otimes |K^0\rangle_r$ (the same for B-mesons). Clearly, this is a different decoherence scenario and corresponds to a different ζ.

Using the experimental data of the CPLEAR experiment [71] in 1998 at Cern gives (see Refs. [69, 70]):

$$\zeta_{K_S K_L} = 0.13 \pm 0.16\,, \tag{65}$$

$$\zeta_{K^0\bar{K}^0} = 0.4 \pm 0.7\,. \tag{66}$$

Clearly, the Schrödinger–Furry hypothesis in K_S, K_L basis is ruled out; however, the data does not exclude any decoherence. For the decoherence scenario assuming the interaction of the system with the environment in the $K^0\bar{K}^0$ basis both limits are possible, i.e., no decoherence, but as well Schrödinger–Furry spontaneous factorization of the wave function. Thus such a decoherence scenario cannot be ruled out by this experimental data set.

Recently (July 2006), Antonio DiDomenico of the KLOE collaboration [72] analyzed the data of the first observation of quantum interference in the process $\Phi \longrightarrow K_S K_L \longrightarrow \pi^+\pi^-\pi^+\pi^-$. They found for the decoherence parameter ζ

$$\zeta_{K_S K_L} = 0.018 \pm 0.040_{\text{stat}} \pm 0.007_{\text{sys}}\,, \tag{67}$$

$$\zeta_{K^0\bar{K}^0} = 0.01 \pm 0.21_{\text{stat}} \pm 0.04_{\text{sys}}\,. \tag{68}$$

These data represent a considerable improvement on those obtained from the CPLEAR experiment above. The results in the $K^0\bar{K}^0$ basis of the KLOE experiments benefit from the measurement in the other basis, the large cancelations between the interference term and the two terms that occur for the $\mathcal{CP}$ suppressed final state $\pi^+\pi^-$. It tests another experimental situation.

A similar scenario was considered for B-mesons by Bertlmann and Grimus [73]; they found

$$\zeta_{B_H B_L} = -0.06 \pm 0.1 \tag{69}$$

using the ARGUS and CLEO data where no time resolution was possible. Recently Go for the KEK collaboration used the data of the BELLE detector in Japan and found [74]

$$\zeta_{B_H B_L} = 0.029 \pm 0.057\,. \tag{70}$$

4.3 The Decoherence Model and its Connection to Measures of Entanglement

Let us now discuss in detail how a decoherence scenario could look like (see also Ref. [70]). We focus on kaons, since for B-mesons everything is analogous. We start with the Lindblad equation, (51) and (52), where we renormalize to surviving mesons for simplicity and choose the generators in the mass eigenstates, i.e.,

$$\begin{aligned} A_1 &= \sqrt{\lambda}\,|K_S\rangle\langle K_S|\,, \\ A_2 &= \sqrt{\lambda}\,|K_L\rangle\langle K_L|\,, \end{aligned} \tag{71}$$

where λ is the strength of interaction with the environment. Solving the differential equations we find that this model is simply connected to the above discussed scenario of a spontaneous factorization of the wave function in the mass eigenstate basis, i.e.,

$$e^{-\lambda \min\{t_l, t_r\}} = 1 - \zeta(\min\{t_l, t_r\})\,. \tag{72}$$

From the EPR-discussions we learn that as long as the two kaons are entangled, they have to be considered as one system and one system has only one time. If one meson is measured or decays, the other one is so to say "alone" and therefore does not longer undergo the same decoherence as before. This is the reason why only the time of the first measured kaon is of importance.

Currently, Gerald Richter of the High Energy Group in Vienna (HEPHY) is using the data of the KEK accelerator for B-mesons to test this decoherence model. With that he is also able to rule out other decoherence models.

Remarkably, for the model under discussion it turns out that the amount of decoherence is very simply connected to measures of entanglement, i.e., the *loss of concurrence* $1 - \mathcal{C}$ (for definition see page 168) and *loss of entanglement of formation* $1 - \mathcal{E}$ (for definition see page 167) is

$$\begin{aligned} 1 - \mathcal{C}\,(\rho(t)) &= \zeta(t) \;=\; 1 - e^{-\lambda t} \\ 1 - \mathcal{E}\,(\rho(t)) &\simeq \frac{1}{\ln 2}\,\zeta(t) \;\simeq\; \frac{\lambda}{\ln 2}\,t\,, \end{aligned}$$

where we expanded for small values of the decoherence parameter λ, or respectively $\zeta(t)$. The strength of the interaction of the system with the environment λ or $\zeta(t)$ are directly measurable in experiment. In this way the very basic mathematical and theoretical concepts are directly confronted with experiments.

Acknowledgement I want to thank Marcus Huber for carefully reading the manuscript and for many helpful comments.

References

1. E. Schrödinger: *Die gegenwärtige situation der quantenmechanik*, Naturw. **23**, 807, 823, 844 (1935) 141, 179
2. A. Einstein, B. Podolski and N. Rosen: *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935) 141, 156
3. Q. Ho-Kim and P.X. Yem: *Elementary Particles and Their Interations* (Springer Verlag, 1998) 142
4. M. Gell-Mann: *Isotopic spin and new unstable particles*, Phys. Rev. **92**, 833 (1953) 143
5. T. Nakano and K. Nishijima, "*Charge independence for V-particles*", Prog. Theor. Phys., Osaka **10**, 581 (1953) 143
6. R.A. Bertlmann, and B.C. Hiesmayr: *Bell inequalities for entangled kaons and their unitary time evolution*, Phys. Rev. A **63**, 062112 (2001) 147, 158, 159, 175
7. N. Gisin and A. Go: *EPR Test with Photons and Kaons*, Am. J. Phys. **69**, 264 (2001) 148
8. A. Bramon, G. Garbarino and B.C. Hiesmayr: *Quantum marking and quantum erasure for neutral kaons*, Phys. Rev. Lett. **92**, 020405 (2004) 150, 173
9. A. Bramon, G. Garbarino and B.C. Hiesmayr: *Active and passive quantum eraser for neutral kaons*, Phys. Rev. A **68**, 062111 (2004) 150
10. R.P. Feynman, R.B. Leighton and M. Sands: *The Feynman Lectures on Physics*, Vol. 3, (Addison-Wesley, 1965), pp. 1–1, 11–20 152

11. D. Greenberger and A. Yasin: *Simultaneous wave and particle knowledge in a neutron interferometer*, Phys. Lett. A **128**, 391 (1988) 152
12. B.-G. Englert: *Fringe visibility and which-way information: An inequality*, Phys. Rev. Lett. **77**, 2154 (1996) 152
13. A. Bramon, G. Garbarino and B.C. Hiesmayr: *Quantitative complementarity in two-path interferometry*, Phys. Rev. A **69**, 022112 (2004) 153
14. B.C. Hiesmayr and V. Vedral: *Interferometric wave-particle duality for thermodynamical systems*, quant-ph/0501015 153
15. J.S. Bell: *On the Einstein-Podolsky-Rosen Paradox*, Physics 1, 195 (1964) 157
16. J.S. Bell: *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987) 157
17. A. Aspect and P. Grangier and G. Roger: *Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: a new violation of Bell's inequalities*, Phys. Rev. Lett. **49**, 91 (1982) 157
18. A. Aspect, J. Dalibard and G. Roger: *Experimental test of Bell's inequalities using time-varying analyzers*, Phys. Rev. Lett. **49**, 1804 (1982) 157
19. M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland: *Experimental violation of Bell's inequality with efficent detection*, Nature **409**, 791 (2001) 157
20. G. Weihs, T. Jennewein, C. Simon, H. Weinfurter and A. Zeilinger: *Violation of Bell's inequality under strict Einstein locality conditions*, Phys. Rev. Lett. **81**, 5039 (1998) 157
21. W. Tittel, J. Bendel, N. Gisin und H. Zbinden: *Violation of Bell inequalities by photons more than 10 km apart*, Phys. Rev. Lett. **81**, 3563 (1998) 157
22. M. Michler, H. Weinfurter and M. Zukowski: *Experiments towards falsification of noncontextual hidden variable theories*, Phys. Rev. Lett. **84**, 5457 (2000) 157
23. Y. Hasegawa, R. Loidl, G. Badurek, M. Baron and H. Rauch: *Violation of a Bell-like inequality in single-neutron interferometry*, Nature **425**, 45 (2003) 157
24. R.A. Bertlmann, K. Durstberger, Yuji Hasegawa and B.C. Hiesmayr: *Berry phase in entangled systems: an experiment with single neutrons*, Phys. Rev. A **69**, 032112 (2004) 158
25. N. Gisin, G. Ribordy, W. Tittel and H. Zbinden: *Quantum crytography*, Rev. Mod. Phys. **74**, 145 (2002) 158
26. C.H. Bennett, G. Brassard, C. Crepeau, R. Josza, A. Peres and W. Wootters: *Teleporting an unknown quantum state via dual classical and EPR Channels*, Phys. Rev. Lett. **70**, 1895 (1993) 158
27. R. Ursin, T. Jennewein, M. Aspelmeyer, R. Kaltenbaek, M. Lindenthal, P. Walther and A. Zeilinger: *Quantum teleportaion across the Danube*, Nature **430**, 849 (2004) 158
28. I. Marcikic, H. de Riedmatten, W. Tittel, H. Zbinden and N. Gisin: *Long distance quantum teleportation of qubits from photons at 1300 nm to photons at 1550 nm wavelength*, Nature **421**, 509–513 (2003) 158
29. P. Shor: *Algorithms for quamtum computation: Discrete logarithms and factoring*, Symposium on Foundations of Computer Science, held in Santa Fe, New Mexico (1994) 158
30. R.A. Bertlmann, W. Grimus and B.C. Hiesmayr: *Open-quantum-system formulation of particle decay*, Phys. Rev. A **73**, 054101 (2006) 158, 161, 162
31. B.C. Hiesmayr: *Nonlocality and entanglement in a strange system*, Eur. Phys. J. C **50**, 73 (2007) 158, 161
32. J.E. Clauser, M.A. Horne, A. Shimony, R.A. Holt: *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23**, 880 (1969) 158
33. R.A. Bertlmann, A. Bramon, G. Garbarino, and B.C. Hiesmayr: *Violation of a Bell inequality in particle physics experimentally verified?*, Phys. Lett. A **332**, 355 (2004) 160, 161
34. A. Acin, R. Gill and N. Gisin: *Optimal bell tests do not require maximally entangled states*, Phys. Rev. Lett. **95**, 210402 (2005) 161, 172
35. G. Lindblad: *On generatores of quantum dynamical semigroups*, Comm. Math. Phys. **48**, 119 (1976) 162, 163
36. V. Gorini, A. Kossakowski and E.C.G. Sudarshan: *Completely positive dynamical semigroups of N-level systems*, J. Math. Phys. **17**, 821 (1976) 162, 163

37. H.-P. Breuer and F. Petruccione: *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002) 162, 163, 178
38. A. Bramon, G. Garbarino and B.C. Hiesmayr: *Quantitative duality and neutral kaon interferometry in CPLEAR experiments*, Eur. J. Phys. C **32**, 377 (2004) 166
39. B.C. Hiesmayr: PhD-Thesis, University of Vienna, Vienna (2002) 167
40. S. Hill and W.K. Wootters: *Entanglement of a pair of quantum bits*, Phys. Rev. Lett. **78**, 5022 (1997) 168
41. S. Ishizaka and T. Hiroshima: *Maximally entangled mixed states in two qubits*, Phys. Rev. A **62**, 022310 (2000) 170, 171
42. A. Bramon, R. Escribano and G. Garbarino: *Bell's inequality tests: From photons to B-mesons*, J. Mod. Opt. **52**, 1681–1684 (2005) 172
43. A. Bramon, R. Escribano and G. Garbarino: *Bell's inequality tests with meson-antimeson pairs*, Found.Phys. **36**, 563–584 (2006) 172
44. A. Bramon and M. Nowakowski: *Bell inequalities for entangled pairs of neutral kaons*, Phys. Rev. Lett. **83**, 1 (1999) 172
45. A. Bramon and G. Garbarino: *Novel Bell's inequalities for entangled $K^0\bar{K}^0$ pairs*, Phys. Rev. Lett. **88**, 040403 (2002) 172
46. A. Bramon and G. Garbarino: *Test of local realism with entangled kaon pairs and without Inequalities*, Phys. Rev. Lett. **89**, 160401 (2002) 172
47. B. Ancochea, A. Bramon and M. Nowakowski: *Bell inequalities for $K^0\bar{K}^0$ pairs from Φ–resonance decays*, Phys. Rev. D **60**, 094008 (1999) 172
48. R.A. Bertlmann, K. Durstberger and Y. Hasegawa: *Decoherence modes of entangled qubits within neutron interferometry*, Phys. Rev. A **73**, 022111 (2006) 170
49. M. Ziman and V. Buzek: *Concurrence versus purity: Influence of local channels on Bell states of two qubits*, Phys. Rev. A **72**, 052325 (2005) 170
50. R.A. Bertlmann, W. Grimus, B.C. Hiesmayr: *Bell inequality and CP violation in the neutral kaon system*, Phys. Lett. A *289*, 21 (2001) 159, 172
51. F. Uchiyama: *Generalized Bell inequality in the kaon system*", Phys. Lett. A **231**, 295 (1997) 159
52. M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. Van der Zouw and A. Zeilinger: *Wave–particle duality of C*-60, Nature **401**, 680 (1999) 172
53. M.O. Scully and K. Drühl: *Quantum Eraser: A proposed photon correlation experiment concerning observation and "delayed choice" in quantum mechanics*, Phys. Rev. A **25**, 2208 (1982) 173
54. S. Dürr and G. Rempe: *Complementarity and quantum erasure in an atom interferometer*, Opt. Commun. **179**, 323 (2000) 173
55. T.J. Herzog, P.G. Kwiat, H. Weinfurter and A. Zeilinger: *Complementarity and the Quantum Eraser*, Phys. Rev. Lett **75**, 3034 (1995) 173, 174
56. Y.-H. Kim, R. Yu, S.P. Kuklik, Y. Shih and M.O. Scully: *Delayed "Choice" Quantum Eraser*, Phys. Rev. Lett. **84**, 1 (2000) 173, 176
57. T. Tsegaye, G. Björk, M. Atatüre, A.V. Sergienko, B.W.A. Saleh and M.C. Teich: *Complementarity and quantum erasure with entangled-photon states*, Phys. Rev. A **62**, 032106 (2000) 173
58. S.P. Walborn, M.O. Terra Cunha, S. Padua and C.H. Monken: *Double-slit quantum eraser*, Phys. Rev. A **65**, 033818 (2002) 173
59. A. Trifonov, G. Björk, J. Söderholm and T. Tsegaye: *Comprehensive experimental test of quantum erasure*, Eur. Phys. J. D **18**, 251 (2002) 173
60. H. Kim, J. Ko and T. Kim: *Quantum-eraser experiment with frequency-entangled photon pairs*, Phys. Rev. A **67**, 054102 (2003) 173
61. Y. Aharonov and M.S. Zubairy: *Time and the quantum: Erasing the past and impacting the future*, Science **307**, 875 (2005) 173
62. A. Bramon, G. Garbarino and B.C. Hiesmayr: *Active and passive quantum erasers for neutral kaons*, Phys. Rev. A **69**, 062111 (2004) 173
63. M.A. Nielson and I.L. Chuang: *Quantum Computation and Quantum Information* (Cambridge University Press, U.K., 2000) 178

64. D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu and H.D. Zeh: *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, 1996) 178
65. R. Penrose: *Quantum computation, entanglement and state reduction*, Phil. Trans. Roy. Soc. Lond. A **356**, 1927 (1998) 179
66. G.C. Ghirardi, A. Rimini and T. Weber: *Unified dynamics for microscopic and macroscopic systems*, Phys. Rev. D **34**, 470 (1986) 179
67. G.C. Ghirardi, A. Rimini and T. Weber: *The puzzling entanglement of Schrödinger's wave function*, Found. Phys. **18**, 1 (1988) 179
68. W.H. Furry: *Note on the quantum mechanical theory of measurement*, Phys. Rev. **49**, 393 (1994) 179
69. R.A. Bertlmann, W. Grimus and B.C. Hiesmayr: *Quantum mechanics, Furry's hypothesis and a measure of decoherence*", Phys. Rev. D **60**, 114032 (1999) 180
70. R.A. Bertlmann, K. Durstberger and B.C. Hiesmayr: *Decoherence of entangled Kaons and its connection to entanglement measures*, Phys. Rev. A **68**, 012111 (2003) 180, 181
71. A. Apstolakis et.al.: *An EPR experiment testing the non-separability of the* $K^0\bar{K}^0$*-wave-function*, Phys. Lett. B **422**, 339 (1998) 178, 180
72. KLOE Collaboration: *First observation of quantum interference in the process* $\phi \to K_S K_L \to \pi_+\pi_-\pi_+\pi_-$*: a test of quantum mechanics and CPT symmetry*, hep-ex/0607027 181
73. R.A. Bertlmann and W. Grimus: *Quantum-mechanical interference over macroscopic distances in* $B^0\bar{B}^0$ *system*, Phys. Lett. B **392**, 426 (1997) 181
74. A. Go and A. Bay for the BELLE collaboration: *Measurements of EPR-type flavour entanglement in Upsilon* $(4S) \to B^0\bar{B}^0$ *decays*, tquant-ph/0702267 181

Five Lectures on Optical Quantum Computing

Pieter Kok

A quantum computer is a machine that can perform certain calculations much faster than a classical computer by using the laws of quantum mechanics. Quantum computers do not exist yet, because it is extremely difficult to control quantum mechanical systems to the necessary degree. What is more, we do at this moment not know *which physical system* is the best suited for making a quantum computer (although we have some ideas). It is likely that a mature quantum information processing technology will use (among others) light, because photons are ideal carriers for quantum information. These notes are an expanded version of the five lectures I gave on the possibility of making a quantum computer using light, at the Summer School in Theoretical Physics in Durban, 14–24 January, 2007. There are quite a few proposals using light for quantum computing, and I can highlight only a few here. I will focus on photonic qubits, and leave out continuous variables completely.[1] I assume that the reader is familiar with basic quantum mechanics and introductory quantum computing.

1 Light and Quantum Information

Simply put, a quantum computer works by storing information in physical carriers, which then undergo a series of unitary (quantum) evolutions and measurements. The information carrier is usually taken to be a *qubit*, a quantum system that consists of two addressable quantum states. Furthermore, the qubit can be put in arbitrary superposition states. The unitary evolutions on the qubits that make up the computation can be decomposed in single-qubit operations and two-qubit operations. Both types of operations or *gates* are necessary if the quantum computer is to outperform any classical computer.

P. Kok (✉)
Quantum and Nano-Technology Group, Department of Materials, Oxford University, Oxford, UK
e-mail: P.kok@sheffield.ac.uk

[1] For a review on optical quantum computing with continuous variables, see Braunstein and Van Loock, *Rev. Mod. Phys.* **77**, 513 (2005).

Kok, P.: *Five Lectures on Optical Quantum Computing*. Lect. Notes Phys. **787**, 187–219 (2010)
DOI 10.1007/978-3-642-02871-7_7 © Springer-Verlag Berlin Heidelberg 2010

1.1 Photons as Qubits

We define the *computational basis states* of the qubit as some suitable set of states $|0\rangle$ and $|1\rangle$. An arbitrary single-qubit operation can take the form of a compound rotation parameterized by two angles θ and ϕ:

$$\begin{aligned} |0\rangle &\rightarrow \cos\theta\,|0\rangle + \mathrm{i}e^{\mathrm{i}\phi}\sin\theta|1\rangle, \\ |1\rangle &\rightarrow \mathrm{i}e^{\mathrm{i}\phi}\sin\theta\,|0\rangle + \cos\theta|1\rangle. \end{aligned} \tag{1}$$

This can be represented graphically in the *Bloch* or *Poincaré* sphere (see Fig. 1).

What type of light can be used as a qubit? The smallest excitation of the electromagnetic field is the *photon*. We cannot construct a standard wave function for the photon, but we can identify the different degrees of freedom that we can use as a qubit: a photon can have the choice between *two spatially separated beams* (or modes), or it can have two distinct *polarizations* [1]. These two representations are mathematically equivalent, as we will show below.

The emission and absorption of photons with momentum k is described mathematically using *creation* and *annihilation* operators:

$$\hat{a}(k)|n\rangle_k = \sqrt{n}|n-1\rangle_k \quad \text{and} \quad \hat{a}^\dagger(k)|n\rangle_k = \sqrt{n+1}|n+1\rangle_k\,. \tag{2}$$

It is straightforward to show that $\hat{n}(k) \equiv \hat{a}^\dagger(k)\hat{a}(k)$ is the number operator $\hat{n}(k)|n\rangle_k = n|n\rangle_k$. The canonical commutation relations between $\hat{a}$ and $\hat{a}^\dagger$ are given by

$$\begin{aligned} \left[\hat{a}(k), \hat{a}^\dagger(k')\right] &= \delta(k-k'), \\ \left[\hat{a}(k), \hat{a}(k')\right] &= \left[\hat{a}^\dagger(k), \hat{a}^\dagger(k')\right] = 0\,. \end{aligned} \tag{3}$$

For the purposes of these notes, we use subscripts to distinguish the creation and annihilation operators for different modes, rather than the functional dependence on k. In photon language, we can define the logical qubit states on two spatial modes a and b as:

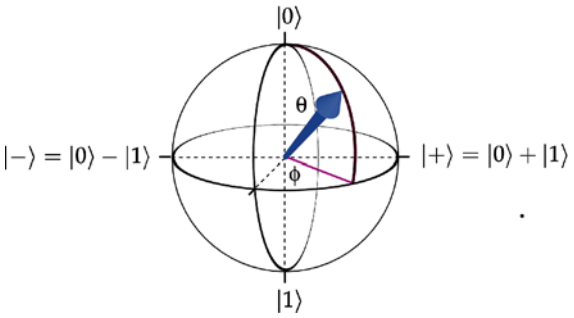


Fig. 1 Graphic representation of the Bloch sphere

$$|0\rangle_L = \hat{a}^\dagger|\circledcirc\rangle = |1,0\rangle_{ab} \qquad \text{and}$$
$$|1\rangle_L = \hat{b}^\dagger|\circledcirc\rangle = |0,1\rangle_{ab}, \tag{4}$$

where $|\circledcirc\rangle$ is the vacuum state and the 0 and 1 denote the photon numbers in the respective modes. The polarization qubits are defined as

$$|0\rangle_L = \hat{a}_H^\dagger|\circledcirc\rangle = |H\rangle \qquad \text{and}$$
$$|1\rangle_L = \hat{a}_V^\dagger|\circledcirc\rangle = |V\rangle. \tag{5}$$

Every state of the electromagnetic field can be written as a function of the creation operators acting on the vacuum state $|\circledcirc\rangle$. A change in the state can therefore also be described by a change in the creation operators (essentially, this is the difference between the Schrödinger and Heisenberg picture). In fact, it is often easier to work out how a physical operation changes the creation and annihilation operators than how it changes an arbitrary state. This is what we will do here. The single-qubit operations on single photons in terms of the creation and annihilation operators consist of the following transformations:

1. The **phase shift** changes the phase of the electromagnetic field in a given mode:

$$\hat{a}_{\text{out}}^\dagger = e^{\mathrm{i}\phi\hat{a}_{\text{in}}^\dagger\hat{a}_{\text{in}}}\,\hat{a}_{\text{in}}^\dagger\,e^{-\mathrm{i}\phi\hat{a}_{\text{in}}^\dagger\hat{a}_{\text{in}}} = e^{\mathrm{i}\phi}\hat{a}_{\text{in}}^\dagger\,, \tag{6}$$

   with the interaction Hamiltonian $H_\phi = \phi\,\hat{a}_{\text{in}}^\dagger\hat{a}_{\text{in}}$ ($\hbar = 1$). Physically, the phase shift can be implemented using a delay line or a transparent element with an index of refraction that is different from free space, or the optical fibre (or whatever medium the photons propagate through). In (6) we used the operator identity

$$e^{\alpha A}Be^{-\alpha A} = B + \alpha[A,B] + \frac{\alpha^2}{2!}[A,[A,B]] + \ldots, \tag{7}$$

   where A is Hermitian.
2. The **beam splitter** usually consists of a semi-reflective mirror: when light falls on this mirror, part will be reflected and part will be transmitted. Let the two incoming modes on either side of the beam splitter be denoted by $\hat{a}_{\text{in}}$ and $\hat{b}_{\text{in}}$, and the outgoing modes by $\hat{a}_{\text{out}}$ and $\hat{b}_{\text{out}}$. When we parameterize the probability amplitudes of these possibilities as $\cos\theta$ and $\sin\theta$, and the relative phase as φ, then the beam splitter yields an evolution in operator form

$$\hat{a}_{\text{out}}^\dagger = \cos\theta\,\hat{a}_{\text{in}}^\dagger + \mathrm{i}e^{-\mathrm{i}\varphi}\sin\theta\,\hat{b}_{\text{in}}^\dagger\,,$$
$$\hat{b}_{\text{out}}^\dagger = \mathrm{i}e^{\mathrm{i}\varphi}\sin\theta\,\hat{a}_{\text{in}}^\dagger + \cos\theta\,\hat{b}_{\text{in}}^\dagger\,. \tag{8}$$

   In terms of the Hamiltonian evolution, we have

$$\hat{a}_{\text{out}}^\dagger = e^{\mathrm{i}H_{\text{BS}}}\,\hat{a}_{\text{in}}^\dagger\,e^{-\mathrm{i}H_{\text{BS}}} \qquad \text{and} \qquad \hat{b}_{\text{out}}^\dagger = e^{\mathrm{i}H_{\text{BS}}}\,\hat{b}_{\text{in}}^\dagger\,e^{-\mathrm{i}H_{\text{BS}}}, \tag{9}$$

where the "interaction Hamiltonian" $H_{\rm BS}$ is given by

$$H_{\rm BS} = \theta e^{{\rm i}\varphi} \hat{a}_{\rm in}^\dagger \hat{b}_{\rm in} + \theta e^{-{\rm i}\varphi} \hat{a}_{\rm in} \hat{b}_{\rm in}^\dagger \,. \tag{10}$$

Mathematically, the two parameters θ and φ represent the angles of a rotation about two orthogonal axes in the Poincaré sphere. The physical beam splitter can be described by any choice of θ and φ, where θ is a measure of the transmittivity and φ gives the phase shift due to the coating of the mirror. An additional phase shift may be necessary to describe the workings of the physical object correctly.

This demonstrates that the beam splitter and the phase shift suffice to implement any single-qubit operation on a single photonic qubit. This case, where a single photon can be in two optical modes, is commonly called the *dual rail* representation, as opposed to the *single rail* representation where the qubit coincides with the occupation number of a single optical mode.

There are similar relations for transforming the polarization of a photon. Physically, the polarization is the spin degree of freedom of the photon. The photon is a spin-1 particle, but because it travels at the speed of light c, the longitudinal component is suppressed. We are left with two polarization states, which make an excellent qubit. The two important operations on polarization are:

1. The **polarization rotation** is physically implemented by quarter- and half-wave plates. We write $\hat{a}_{\rm in} \to \hat{a}_x$ and $\hat{b}_{\rm in} \to \hat{a}_y$ for some orthogonal set of coordinates x and y (i.e. $\langle x|y\rangle = 0$). The parameters θ and φ are now angles of rotation:

$$\begin{aligned} \hat{a}_{x'}^\dagger &= \cos\theta\, \hat{a}_x^\dagger + {\rm i} e^{-{\rm i}\varphi} \sin\theta\, \hat{a}_y^\dagger \,, \\ \hat{a}_{y'}^\dagger &= {\rm i} e^{{\rm i}\varphi} \sin\theta\, \hat{a}_x^\dagger + \cos\theta\, \hat{a}_y^\dagger \,. \end{aligned} \tag{11}$$

   This evolution has the same Hamiltonian as the beam splitter, and it formalizes the equivalence between polarization and two-mode logic.
2. The **polarizing beam splitter** (PBS) spatially separates modes with orthogonal polarization. If the PBS is cut to separate horizontal and vertical polarization, the transformation of the incoming modes ($a_{\rm in}$ and $b_{\rm in}$) yields the following outgoing modes ($a_{\rm out}$ and $b_{\rm out}$):

$$\begin{aligned} &\hat{a}_{{\rm in},H} \to \hat{a}_{{\rm out},H} \text{ and } \hat{a}_{{\rm in},V} \to \hat{b}_{{\rm out},V} \\ &\hat{b}_{{\rm in},H} \to \hat{b}_{{\rm out},H} \text{ and } \hat{b}_{{\rm in},V} \to \hat{a}_{{\rm out},V} \,. \end{aligned} \tag{12}$$

   Using quarter-wave plates and polarizers, we can also construct a PBS for different polarization directions (e.g. L and R), in which case we make the substitution $H \leftrightarrow L$, $V \leftrightarrow R$.

1.2 Interferometers

When there are many optical modes a_1 to a_N, we need a compact description if we are to apply beam splitters, phase shifters and such to these optical modes. Equations (8) and (11) can be written as a vector equation:

$$\begin{pmatrix} \hat{a}^\dagger_{\text{out}} \\ \hat{b}^\dagger_{\text{out}} \end{pmatrix} = \begin{pmatrix} \cos\theta & \mathrm{i}e^{-\mathrm{i}\varphi}\sin\theta \\ \mathrm{i}e^{\mathrm{i}\varphi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{a}^\dagger_{\text{in}} \\ \hat{b}^\dagger_{\text{in}} \end{pmatrix}. \tag{13}$$

In general, when we have many optical modes we can collect their corresponding operators in a vector, and if U is a unitary matrix, the multi-mode transformations become

$$\hat{\mathbf{a}}^\dagger_{\text{out}} = U\hat{\mathbf{a}}^\dagger_{\text{in}} \quad \text{or} \quad \hat{a}^\dagger_{j,\text{out}} = \sum_k U_{jk}\hat{a}^\dagger_{k,\text{in}}\,, \tag{14}$$

where $\hat{\mathbf{a}}_{\text{out}} \equiv (\hat{a}_1, \ldots, \hat{a}_N)$. A successive application of beam splitters and phase shifters is therefore equivalent to a series of unitary matrices associated with these elements. It turns out that any $N \times N$ unitary matrix can be decomposed in terms of 2×2 unitary matrices T_{jk} of the form[2] in Eq. (13) [2]. Therefore, any arbitrary *interferometer* (in which N optical modes interfere with each other) can be constructed from beam splitters, phase shifts, and polarization rotations. This is an extraordinarily powerful result, and we can use it to define a general interferometer as a unitary transformation U on N (spatial) modes, or an N-port.

We should note one very important thing, though: just because we can decompose U into a series of "single-qubit" operations defined above, it does not mean we can call this a quantum computer. Qubits should be well-defined physical systems that you can track through the computation. However, in an interferometer with n input photons it is possible (and inevitable) that some of them will end up in the same mode. Since photons are indistinguishable particles (or at least they should be in this model), we cannot track the quantum information they carry. Also, we still have not shown how to make two-qubit gates. This means that we have to work a bit harder to show we can make a quantum computer in this way.

Exercise 1. Prove the relations in (6) and (8).

Exercise 2. Show how to turn a qubit on two spatial modes into a polarization qubit.

Exercise 3. Write down the interaction Hamiltonian and unitary matrix for a mirror.

[2] To be precise, the $N \times N$ matrix is decomposed in terms of $T_{jk} \otimes \mathbb{I}_{N-2}$, where $\mathbb{I}_{N-2}$ is the $(N-2) \times (N-2)$ identity matrix.

2 Two-Qubit Gates and the KLM Scheme

While single-qubit operations on a photon are easy, two-qubit operations on two photons are very difficult. Consider the two-qubit gate that generates the following transformation:

$$|H,H\rangle_{ab} \quad \rightarrow \quad \frac{1}{\sqrt{2}}\left(|H,H\rangle_{cd}+|V,V\rangle_{cd}\right). \tag{15}$$

This is a perfectly sound quantum mechanical operation, and one that is often needed in a quantum computation. Let us see how we can implement this with photons and linear optical elements. In terms of the creation operators acting on the vacuum $|\circledcirc\rangle$, this transformation can be written as

$$\hat{a}_H^\dagger \hat{b}_H^\dagger |\circledcirc\rangle \quad \rightarrow \quad \frac{1}{\sqrt{2}}\left(\hat{c}_H^\dagger \hat{d}_H^\dagger + \hat{c}_V^\dagger \hat{d}_V^\dagger\right)|\circledcirc\rangle\,. \tag{16}$$

Let us substitute the operator transformations for $\hat{a}_H^\dagger$ and $\hat{b}_H^\dagger$:

$$\hat{a}_H^\dagger \hat{b}_H^\dagger = \left(\sum_j U_{j1}\hat{c}_j^\dagger\right)\left(\sum_k U_{k2}\hat{d}_k^\dagger\right) = \sum_{jk} U_{j1}U_{k2}\hat{c}_k^\dagger \hat{d}_j^\dagger\,. \tag{17}$$

By construction, this is a separable expression. However, the state we wish to create is entangled (inseparable)! So we can never get an entangling (two-qubit) gate this way. Therefore we arrive at the conclusion that single-photon inputs, N-ports and final read-out are not sufficient to make a quantum computer!

2.1 The KLM Approach

Clearly, we have to add something more. What about feed-forward? By making a measurement on part of the output of the N-port we may be able to reject or accept certain terms in a superposition, and effectively gain entanglement. This is the approach championed in the now famous "KLM" paper [3], after the authors Knill, Laflamme, and Milburn. First, they construct an N-port with suitable input states, which upon the correct detection signature gives a two-qubit gate. Since the detection is a true quantum mechanical process, the outcome is unknown beforehand, and the gate succeeds only a fraction of the time. The gate destroys the qubits (and hence the quantum information) when it fails. If this gate was used directly in a computation, the overall success probability of the computation would decrease exponentially with the number of two-qubit gates, so something else is needed.

Second, KLM show how a probabilistic optical gate can be applied to two qubits without destroying them. This relies on a method developed by Gottesman and Chuang, called the *teleportation trick*, and it allows us to use a previously created

entangled state to teleport gates into the quantum circuit [4]. I will describe this procedure in a different form later in these lecture notes. Before that, let us look at quantum gates in a bit more detail.

Three special single-qubit gates are the Pauli operators. In matrix notation (in the computational basis) these look like

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{18}$$

The X operator is a bit flip, and the Z operator is a phase flip. The Y operator is a combination of X and Z. Two very useful two-qubit gates are the controlled-Z, where a Z operation is applied to the second qubit if the first qubit is in state $|1\rangle$, and the controlled-NOT, where an X operation (a bit flip) is applied to the second qubit depending on the first. In matrix notation, these gates look like

$$U_{\mathrm{CZ}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad U_{\mathrm{CNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{19}$$

These two entangling gates have a very special property: when we apply these transformations to a tensor product of two Pauli matrices we get again two Pauli matrices:

$$U_{\mathrm{CZ}}^{\dagger}(P_1 \otimes P_2)U_{\mathrm{CZ}} = P_3 \otimes P_4\,, \tag{20}$$

and similarly for the CNOT gate. Operators with this property (of turning Pauli operators into Pauli operators) are members of the *Clifford* group. This is a very important symmetry in quantum information theory, as it forms the basis of quantum error correction.

Suppose we wish to apply the CZ gate to two qubits $|\phi_1\rangle$ and $|\phi_2\rangle$ (see Fig. 2). We can teleport these states to new qubit systems and then apply the CZ gate to the teleported qubits. This in itself achieves not much, but we can now commute the CZ gate through the corrective single-qubit Pauli gates to make the CZ part of the entanglement channel in teleportation. The fact that the CZ operation is part of the Clifford group now comes in handy: the commutation operation will *not* induce any new two-qubit gates.

Knill, Laflamme, and Milburn [3] used this trick to create two-qubit gates for single-photon qubits. The complication here was that the Bell measurement essential to teleportation cannot be carried out deterministically on single photons. To this end, KLM designed a teleportation protocol that uses $2n$ additional photons and succeeds with a success probability $n/(n+1)$. Since we require two teleportation events, the success probability of the two-qubit gate is $[n/(n+1)]^2$. Failure of the gate amounts to a measurement in the computational basis, which is easy to protect against with standard error correction (i.e. parity codes).

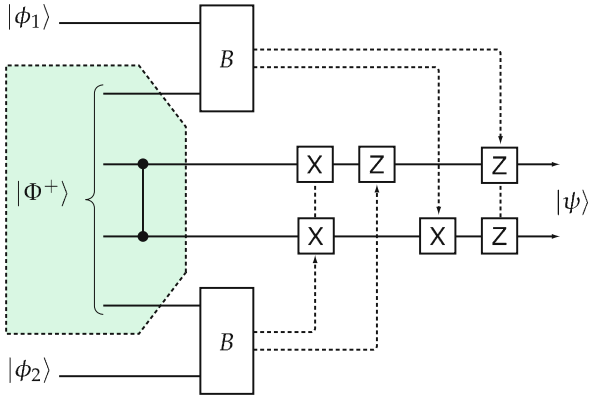


Fig. 2 The teleportation trick. The CZ operation is denoted by a *vertical line*, which connects to the two qubits with a *solid dot*. We teleport both qubits $|\phi_1\rangle$ and $|\phi_2\rangle$ (B denotes the Bell measurement), and apply the CZ to the output qubits. Then we commute the CZ from the *right* to the *left*, through the corrective Pauli operations of the teleportation. The CZ operation can then be performed *off-line*, together with the preparation of the entanglement channel for teleportation *(the shaded box)*

This may all seem a bit overwhelming, and the reader will be pleased to hear that several simplifications of this scheme have been proposed. In the next part of this lecture I will describe two very simple optical operations that can be used to create all the entanglement we need.

2.2 Two-Photon Interference

Quantum computing with photons and linear optical elements relies critically on two-photon interference, with or without polarization. In this section, I will first describe the quintessential two-photon Hong-Ou-Mandel effect. After that, I will extend it to the case of polarized photons.

The Hong-Ou-Mandel (HOM) effect [5] occurs when two *identical* photons (the same polarization, the same frequency, and the same spatio-temporal profile) each enter an input port of a 50:50 beam splitter (see Fig. 3). Mathematically, this can be condensed to the following: since the photons are identical, we can suppress all the spatio-temporal, frequency, and polarization information in the creation operator,

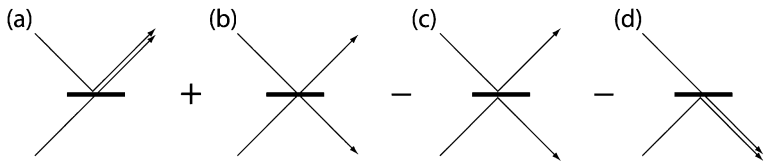


Fig. 3 The four amplitudes in the two-photon interference experiment by Hong, Ou, and Mandel. Components (**b**) and (**c**) always have opposite sign by virtue of unitarity of the beam splitter, and cancel

and write the input state as $|1,1\rangle_{ab} = \hat{a}^\dagger \hat{b}^\dagger |\circledcirc\rangle$ on the two input modes a and b. The 50:50 beam splitter is characterized by the transformation

$$\hat{a} \to \frac{\hat{c} + \hat{d}}{\sqrt{2}} \quad \text{and} \quad \hat{b} \to \frac{\hat{c} - \hat{d}}{\sqrt{2}}. \tag{21}$$

Classically, when the photons enter the beam splitter, each will independently choose whether it will exit in mode c or d. As a result, we expect the photons half of the time to come out in the same output (both in c or both in d), and half of the time they should come out in different output modes (one in mode c and one in mode d). However, quantum mechanically we get something different.

When we substitute the beam splitter transformation rules into the input state $|1,1\rangle_{ab}$, we obtain

$$\begin{aligned} |1,1\rangle_{ab} = \hat{a}^\dagger \hat{b}^\dagger |\circledcirc\rangle \to \frac{1}{2}\left(\hat{c}^\dagger + \hat{d}^\dagger\right)\left(\hat{c}^\dagger - \hat{d}^\dagger\right)|\circledcirc\rangle = \frac{1}{2}\left(\hat{c}^{\dagger 2} - \hat{d}^{\dagger 2}\right)|\circledcirc\rangle \\ = \frac{|2,0\rangle_{cd} - |0,2\rangle_{cd}}{\sqrt{2}}. \end{aligned} \tag{22}$$

We see that the $|1,1\rangle_{cd}$ term in the output modes of the beam splitter is suppressed. This is the HOM effect, and the absence of coincidence counts in such an interference experiment is called the HOM dip. When the input photons are distinguishable (for example if they have different frequencies, or if they arrive at different times at the beam splitter), the dip disappears, and we see the $|1,1\rangle_{cd}$ component in the superposition: the photons behave as classical particles. It is therefore extremely important in such experiments that the photons are truly indistinguishable. This is one of the hardest requirements to meet in linear optical quantum computing.

Another way to see how the HOM effect works is to write down all the different possibilities in which the photons can travel through the beam splitter (see Fig. 3). The output state of (b) and (c) are indistinguishable, so we do not know whether both photons were transmitted or reflected. Moreover, the beam splitter does not retain a memory how the photons interacted at its surface. Therefore, we have to sum the two possibilities coherently. Unitarity of the beam splitter ensures that the relative phase is -1, and the two processes cancel. We can run this experiment backwards as well, because a unitary evolution is reversible. The sources then become detectors, and vice versa. It is then easy to see that a coincidence count after a 50:50 beam splitter projects onto the state $|2,0\rangle - |0,2\rangle$ of the input modes.

The HOM effect is the corner stone of KLM-type optical quantum computing. When the qubit is a dual-rail single photon, every two-qubit gate is based on this effect. However, it may be sometimes more convenient to use polarization qubits. Can we construct a similar two-qubit interferometer? The answer is yes: assume that we have two photons impinging on a polarizing beam splitter. In Fig. 4 you can see that the action of this device looks very similar, except that there is no cancellation. When we erase polarization information in the output modes by 45° rotations and

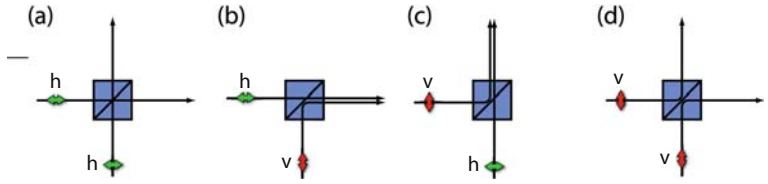


Fig. 4 The effect of a polarizing beam splitter on two input photons. The letters v and h denote vertical and horizontal polarization, respectively

perform single-photon detection, we can construct so-called *fusion gates*. These turn out to be extremely useful for optical quantum computing.

There are two types of fusion gates, aptly named type I and type II [6]. In type I, only one of the output ports is detected, while in type II both output ports are detected. Because of this detection, if we want to create entanglement we cannot start with single photons. The basic building block that is to be used with fusion gates is a Bell state, e.g. $|H,V\rangle + |V,H\rangle$. It is not easy to make these states on demand, but there are quite a few experimental efforts underway to create them with micro-pillar structures. I will first describe the precise workings of both fusion gates, and then I will show how they can be used to make large sets of entangled qubits.

The type-I fusion gate is a polarizing beam splitter cut for horizontal and vertical polarization, and one of the output modes (say, d) has a 45° polarization rotation, followed by photo-detection in the $\{H,V\}$ basis. Let us assume that the input modes a and b are entangled with some other modes, such that the most general input state can be written as $(f_1\hat{a}_H^\dagger + f_2\hat{a}_V^\dagger)(f_3\hat{b}_H^\dagger + f_4\hat{b}_V^\dagger)|\circledcirc\rangle$. Here, the f_k are arbitrary functions of creation operators on other optical modes. When we substitute the transformation of the polarizing beam splitter and the polarization rotation, we obtain the following operator:

$$\frac{f_1 f_3}{\sqrt{2}}\hat{c}_H^\dagger\left(\hat{d}_H^\dagger + \hat{d}_V^\dagger\right) + f_1 f_4 \hat{c}_H^\dagger \hat{c}_V^\dagger + \frac{f_2 f_3}{2}\left(\hat{d}_H^{\dagger 2} - \hat{d}_V^{\dagger 2}\right) + \frac{f_2 f_4}{\sqrt{2}}\hat{c}_V^\dagger\left(\hat{d}_H^\dagger - \hat{d}_V^\dagger\right).$$

After post-selecting the output state of all the modes (including the support of the f_k) on the detector outcome (d_H, d_V), we have

$$\begin{aligned}
(0,0) \quad &: \quad |\psi_{\text{out}}\rangle = f_1 f_2 \hat{c}_H^\dagger \hat{c}_V^\dagger |\circledcirc\rangle \\
(2,0) \text{ or } (0,2) : \quad &|\psi_{\text{out}}\rangle = \frac{f_2 f_3}{2}|\circledcirc\rangle \\
(1,0) \quad &: \quad |\psi_{\text{out}}\rangle = \frac{1}{\sqrt{2}}\left(f_1 f_3 \hat{c}_H^\dagger + f_2 f_4 \hat{c}_V^\dagger\right)|\circledcirc\rangle \\
(0,1) \quad &: \quad |\psi_{\text{out}}\rangle = \frac{1}{\sqrt{2}}\left(f_1 f_3 \hat{c}_H^\dagger - f_2 f_4 \hat{c}_V^\dagger\right)|\circledcirc\rangle . \qquad (23)
\end{aligned}$$

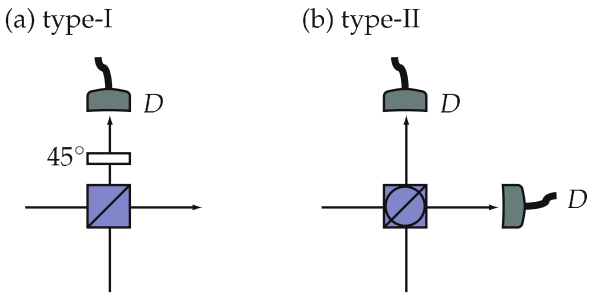


Fig. 5 Two types of fusion operators. (**a**) The type-I fusion operator employs a polarization beam splitter (PBS1) followed by the detection D of a single output mode in the 45° rotated polarization basis. This operation determines the parity of the input mode with probability 1/2. (**b**) The type-II fusion operator uses a diagonal polarization beam splitter (PBS2), detects both output modes, and projects the input state onto a maximally entangled Bell state with probability 1/2

In the case where we find a single photon in mode d (vertical or horizontal), it is easy to see that we create entanglement. In particular (and with abusive notation), suppose that $f_1 = \hat{a}_H^\dagger$, $f_2 = \hat{a}_V^\dagger$, $f_3 = \hat{b}_H^\dagger$, and $f_4 = \hat{b}_V^\dagger$. The arrival of a horizontal photon in mode d then signals the output state $(\hat{a}_H^\dagger \hat{b}_H^\dagger \hat{c}_H^\dagger + \hat{a}_V^\dagger \hat{b}_V^\dagger \hat{c}_V^\dagger)|\circledcirc\rangle$, that is, a three-photon GHZ state.

The type-II fusion operator in Fig. 5b works in a very similar way to the type-I fusion gate, except now the polarizing beam splitter is cut to diagonal polarization $|H\rangle \pm |V\rangle$, and both output ports are detected in the $\{H, V\}$ basis. When we again substitute the transformation of the polarizing beam splitter and the polarization rotation, we obtain the operator

$$\begin{aligned}(f_1 + f_2)(f_3 - f_4)\left(\hat{c}_H^{\dagger 2} - \hat{c}_V^{\dagger 2}\right) + (f_1 - f_2)(f_3 + f_4)\left(\hat{d}_H^{\dagger 2} - \hat{d}_V^{\dagger 2}\right) \\ + 2(f_1 f_3 + f_2 f_4)\hat{c}_H^\dagger \hat{d}_H^\dagger + 2(f_1 f_4 + f_2 f_3)\hat{c}_H^\dagger \hat{d}_V^\dagger + 2(f_1 f_4 + f_2 f_3)\hat{c}_V^\dagger \hat{d}_H^\dagger \\ + 2(f_1 f_3 + f_2 f_4)\hat{c}_V^\dagger \hat{d}_V^\dagger .\end{aligned}$$

Depending on the photon detection signature (c, d), we have the output state

$$\begin{aligned}(2H, 0) \text{ or } (2V, 0): &\quad |\psi_{\text{out}}\rangle = (f_1 + f_2)(f_3 - f_4)|\circledcirc\rangle \\ (0, 2H) \text{ or } (0, 2V): &\quad |\psi_{\text{out}}\rangle = (f_1 - f_2)(f_3 + f_4)|\circledcirc\rangle \\ (H, H) \text{ or } (V, V): &\quad |\psi_{\text{out}}\rangle = (f_1 f_3 + f_2 f_4)|\circledcirc\rangle \\ (H, V) \text{ or } (V, H): &\quad |\psi_{\text{out}}\rangle = (f_1 f_4 + f_2 f_3)|\circledcirc\rangle \,. \end{aligned} \tag{24}$$

Clearly, when we find one photon in each output port, the type-II fusion gate is an entangling gate. It can be interpreted as a *parity* measurement as follows: suppose that the functions f_k are not operators, but quantum state amplitudes of the two input modes a and b instead. The input state is then given by $f_1 f_3|H, H\rangle + f_2 f_4|V, V\rangle + f_1 f_4|H, V\rangle + f_2 f_3|V, H\rangle$. Clearly, finding two photons with *the same* polarization

in the output modes will project onto the *even parity* component of the input state, which finding two photons with *different* polarization will project onto the *odd parity* component.

The fusion gates are not your regular two-qubit gates, because you can not put a separable state in and get an entangled state out. In fact, there is no output in the type-II fusion gate at all. So how can we do quantum computing with this? The answer was given by Browne and Rudolph [6], and it involves a whole new approach to quantum computing. It is known as the "one-way model" of quantum computing, "cluster state quantum computing", of the more generic "measurement-based quantum computing". I will discuss the basic principles of this approach in the next section.

Exercise 4. Can you make *any* entanglement with single photons and N-ports?

Exercise 5. Show that the Hadamard, CZ and CNOT operators are members of the Clifford group.

Exercise 6. Calculate the success probability of the type-I fusion gate.

Exercise 7. Verify the relations (23) and (24). Convince yourself that the type-II fusion gate is a parity projection.

3 Cluster States

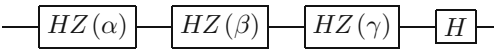
Cluster states were introduced by Raussendorf and Briegel [7], and form an alternative approach to quantum computing. The heart of this architecture is to create a large entangled state as a resource. The computation then proceeds as a series of (parallel) *single-qubit measurements*. Since all the entanglement is produced "off-line", this is a particularly powerful approach for single-photon quantum computing.

3.1 From Circuits to Clusters

Before we introduce cluster states, we must cast arbitrary single-qubit rotations into arbitrary rotations around the Z axis and Hadamard operations H. Using $H^2 = \mathbb{I}$ and $X = HZH$, any arbitrary rotation $\boldsymbol{\theta}$ can be decomposed into three Euler angles α, β, and γ:

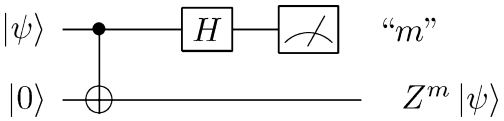
$$R(\boldsymbol{\theta}) = Z(\gamma)X(\beta)Z(\alpha) = H\,HZ(\gamma)\,HZ(\beta)\,HZ(\alpha)\,, \tag{25}$$

where $Z(\alpha) \equiv \exp(\mathrm{i}\alpha Z/2)$ and $X(\beta) \equiv \exp(\mathrm{i}\beta X/2)$. In circuit language, this becomes



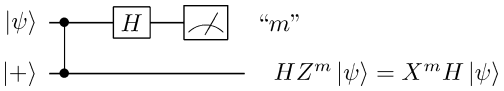
and I will now show how $HZ(\alpha)$ can be implemented via a single-qubit measurement.

Consider the following circuit diagram of single-qubit teleportation:

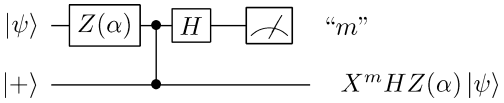


The measurement is in the computational basis, and the outcome "m" takes the value 0 or 1. Depending on this value, we apply a Pauli Z operation to the teleported qubit.

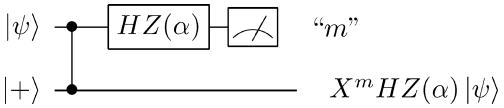
The next step is to translate the CNOT gate into the CZ gate. This procedure incurs two Hadamard gates, which are absorbed into the state of the ancilla ($|0\rangle \rightarrow |+\rangle$) and the teleported qubit.



A single-qubit rotation around the Z axis to the input qubit $|\psi\rangle$ can be written as:



The rotation around the z-axis commutes with the CZ operation (they are both diagonal in the computational basis), and we can write:

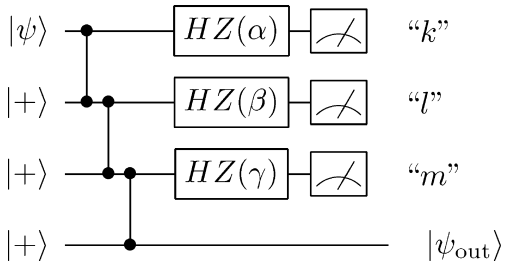


We can reinterpret this diagram as an entangled state $|\Psi\rangle = CZ|\psi, +\rangle$, followed by a single-qubit measurement on the first qubit that performs the single-qubit gate $HZ(\alpha)$ on the state $|\psi\rangle$. The precise measurement basis $A(\alpha)$ is determined by

$$\langle\psi|Z(-\alpha)H\ Z\ HZ(\alpha)|\psi\rangle = \langle\psi|Z(-\alpha)XZ(\alpha)|\psi\rangle = \langle\psi|A(\alpha)|\psi\rangle\,. \tag{26}$$

This corresponds to a measurement along an axis in the equatorial plane of the Bloch sphere.

In order to implement an arbitrary rotation $R(\boldsymbol{\theta})$, this procedure must be concatenated three times



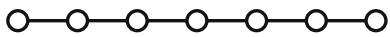
with $|\psi_{\text{out}}\rangle = (X^m HZ(\gamma))(X^l HZ(\beta))(X^k HZ(\alpha))|\psi\rangle$. However, the operators X^k, X^l and X^m depend on the measurement outcomes, and we should try to get rid of them by commuting them through the Pauli gates and Hadamards. We can again use the relations $ZX = -XZ$ to show that

$$Z(\beta)X = \sum_{n=0}^{\infty} \left(\frac{\mathrm{i}\beta}{2}\right)^n \frac{Z^n}{n!} X = \sum_{n=0}^{\infty} \left(\frac{-\mathrm{i}\beta}{2}\right)^n \frac{XZ^n}{n!} = XZ(-\beta)\,. \tag{27}$$

This therefore gives rise to an adjustment of the measurement bases depending on the previous measurement outcomes, and it results in a definite temporal direction in the computation (see Exercise 8 at the end of this lecture). Hence the name "one-way model" of quantum computing. (Remember that all the elements in the traditional circuit model are unitary operators, and therefore reversible.)

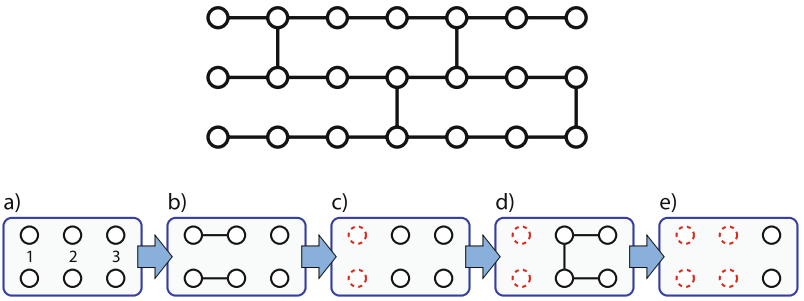
3.2 Universal Cluster States

Now that we have constructed arbitrary single-qubit operations, we do not need to start our circuit with the input state $|\psi\rangle$, but we can start with another $|+\rangle$ state and implement the first single-qubit rotation to obtain the required input state $|\psi\rangle$. We can then write the evolution of a single qubit graphically as a string of ancilla qubits in state $|+\rangle$ (circles), connected via CZ operations (edges): multi-qubit evolution is



then represented as a collection of such strings. The strings can be bridged vertically by edges, which in turn induce two-qubit operations: when all the nearest-neighbour connections are established and the qubits form an entangled grid, *any* quantum circuit can be realized if the cluster state is large enough. Such a state is called a *universal cluster state*.

To show that a vertical bridge induces a CZ gate, consider the following sequence of measurements and entangling operations: we start with two rows of three qubits without any entanglement (a). The two rows will form the two qubits we wish to apply the CZ gate to. We then entangle qubit 1 with qubit 2 in each row (b) and measure qubits 1 (c). After the first two measurements, the states of qubits 1 are



transferred to qubits 2. We can then apply the CZ gate to the two qubits, as well as the two CZ gates that connect qubits 2 with qubits 3 (d). Finally, we measure qubits 2 and transfer the quantum information to the output qubits 3 (e). The point here is that we can apply the CZ gate as if we are using it in the circuit model. However, because the measurements and the CZ operations in this sequence commute, we could have created all the entanglement at the start. Hence the vertical CZ gates are suitable as two-qubit gates in cluster state quantum computing.

An often heard objection to cluster state quantum computing is that it seems to be very wasteful with entanglement. Instead of having to create entanglement for every two-qubit gate in an N-qubit computation, we seem to need at least N entangling operations for *every* clock cycle! However, this is far too pessimistic. There is an enormous redundancy in a cluster state that is translated straight from the circuit model, as we did above. First of all, we can perform all single-qubit operations in the Clifford group *before* the computation starts: these operations also correspond to measurements, but their outcome does not affect any subsequent choice of measurement basis and can therefore be carried out at any stage. Also, these measurements will turn cluster states into smaller cluster states. As a result, we can calculate the effect of most Clifford operations and create a minimal cluster state that is in fact much smaller than what we found in the translation from the circuit model. Since most of the error correction in the quantum computation involves Clifford operations, this is a huge saving. Second, we do not have to create the complete cluster state for the computation all at once. We can create a cluster with a relatively shallow depth, and keep adding qubits to the right as we measure qubits on the left. That way, we make our cluster "just in time". The fusion gates can be used to create these cluster states with a moderate overhead per qubit, as I will show in the next section.

3.3 Making Cluster States with Fusion Gates

In order to show that we can make cluster states with fusion gates we need two things. First, I will give a slightly unconventional description of the CZ gate, which allows us to give a formal description of a cluster state. Second, I will rewrite the action of the fusion gate in bracket notation.

A cluster state is a collection of qubits initially in the $|+\rangle$ state, with CZ operations applied to a set of qubit pairs. Remember that the CZ gate is defined as a Z operation on the target[3] qubit if the state of the control qubit is $|1\rangle$. We can write this as

$$U_{\rm CZ}|+,\psi\rangle_{12} = |0\rangle_1|\psi\rangle_2 + |1\rangle_1 (Z_2|\psi\rangle_2)\,, \tag{28}$$

where we ignored the overall normalization factor $1/\sqrt{2}$. We can do this, because all the terms have the same absolute value of the amplitude, and we are interested only in the relative phases. Now let us see what happens when there are multiple qubits in the cluster state. Qubit 1 may then have edges with multiple qubits:

$$\mathcal{U}_{\rm CZ}|+,\psi\rangle_{1..N} = |0\rangle_1|\psi\rangle_{2..N} + |1\rangle_1 \left(Z_{j_1}\ldots Z_{j_k}|\psi\rangle_{2..N}\right), \tag{29}$$

where Z_j is the Pauli Z operator on qubit j, and $\mathcal{U}_{CZ}$ is a series of CZ gates on qubit 1 and its neighbours. To be accurate, we should write this as

$$\mathcal{U}_{\rm CZ}|+,\psi\rangle_{1..N} = |0\rangle_1|\psi\rangle_{2..N} + |1\rangle_1 \prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_{2..N}. \tag{30}$$

Here, we defined $\mathcal{Z}_j = \mathbb{I}_2 \otimes \ldots Z_j \otimes \mathbb{I}_{j+1}\ldots\otimes\mathbb{I}_N$, and the neighbourhood $n(1)$ is the set of qubits that are connected to qubit 1 via a CZ operation. Now suppose that we have two separate cluster states that we wish to fuse into one. We can write the separate states as

$$\left(|0\rangle_1|\psi\rangle_a + |1\rangle_1 \prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right) \otimes \left(|0\rangle_2|\phi\rangle_b + |1\rangle_2 \prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right). \tag{31}$$

So qubits 1 and 2 are connected to two different cluster states $|\psi\rangle$ and $|\phi\rangle$ on qubit sets a and b, respectively. The neighbourhoods $n(1)$ and $n(2)$ have their support in these respective qubit sets.

We want to apply the type-I fusion gate to qubits 1 and 2, but before we can do this, we should write the action of the fusion gate in a more convenient form. From the last two lines of (23) we see that a single photon in either d_H or d_V heralds success, so let us assume we detect one photon in d_H. How does that transform the input to the output? In the discussion leading up to (23) we assumed that the state was given by

$$\left(f_1 f_3\,\hat{a}_H^\dagger \hat{b}_H^\dagger + f_1 f_4\,\hat{a}_H^\dagger \hat{b}_V^\dagger + f_2 f_3\,\hat{a}_V^\dagger \hat{b}_H^\dagger + f_2 f_4\,\hat{a}_V^\dagger \hat{b}_V^\dagger\right)|\circledcirc\rangle. \tag{32}$$

[3] It so happens that the CZ gate is symmetric, so it does not really matter which qubit you call the control and which the target.

The fusion gate turns this into

$$\left(f_1 f_3 \hat{c}_H^\dagger + f_2 f_4 \hat{c}_V^\dagger\right) |\circledcirc\rangle, \tag{33}$$

and we can therefore deduce that only the $|H, H\rangle$ and $|V, V\rangle$ components survive. Moreover, the operator $\hat{c}_j^\dagger$ creates a photon with the same polarization as the ones that have just been detected: this can be written in bracket notation as

$$U_{\text{type I}}^{(H)} = |H\rangle\langle H, H| + |V\rangle\langle V, V|\,. \tag{34}$$

A similar expression can be deduced for the case where a vertically polarized photon is detected.

To show that the type-I fusion gate can connect two cluster states, let us write $U_{\text{type I}}$ in the computational basis: $U_{\text{type I}}^{(0)} = |0\rangle_3\,{}_{12}\langle 0, 0| + |1\rangle_3\,{}_{12}\langle 1, 1|$. The fusion gate is now applied to qubits 1 and 2 in Eq. (31). This gives

$$\begin{aligned} |\psi_{\text{out}}\rangle &= |0\rangle_3 |\psi, \phi\rangle_{ab} + |1\rangle_3 \prod_{j\in n(1)} \prod_{k\in n(2)} \mathcal{Z}_j \mathcal{Z}_k |\psi, \phi\rangle_{ab} \\ &\equiv |0\rangle_3 |\Psi\rangle_c + |1\rangle_3 \prod_{l\in n(1)\cup n(2)} \mathcal{Z}_l |\Psi\rangle_c, \end{aligned} \tag{35}$$

where we defined $|\psi, \phi\rangle \equiv |\Psi\rangle$ and c is the union of the two qubit sets a and b. Note that (35) is again of the form in (30), and is therefore another cluster state. If we had found measurement outcome 1 in the fusion gate, the same cluster state is created, up to a local Z operation. This shows that we can use type-I fusion gates to create cluster states. The same is true for type-II gates.

Of course, the fusion gates are probabilistic, and half of the time the gate fails. It turns out that the type-II gate is better behaved than the type-I when it fails, and we should therefore aim to create the cluster states with type-II gates. However, we cannot just take the cluster state we want to expand and a single Bell state and apply the type-II fusion, because in every successful gate we necessarily lose two photons through detection. We therefore need the type-I gate to create larger (but still small) cluster states, and use the type-II gate to add these to the cluster. How large should the mini-clusters be?

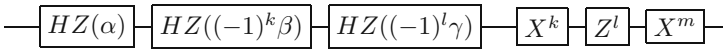
Suppose we have a (linear) cluster of size N, and we want to add a mini-cluster of size m. The success probability of the fusion gate is p, and upon failure we need to detect one extra qubit to return the large multi-qubit state to a cluster state. In order for the cluster to grow we need to obey the following bound:

$$p(N + m - 2) + (1 - p)(N - 2) > N \quad \text{or} \quad m > \frac{2}{p}. \tag{36}$$

Therefore, even if the success probability is very small (for instance because of detection inefficiencies), we can still choose the size of our mini-clusters m such

that we can efficiently grow large cluster states. However, the larger m is, the more the average cost of adding a qubit to the cluster, so we want p to be reasonably large.

Exercise 8. Show that the effective single-qubit operation corresponds to



Note that every measurement depends at most on the previous measurement outcome. How do the final three Pauli's affect the computation?

Exercise 9. Find the action of the type-I fusion operator conditioned on detecting a vertically polarized photon.

4 Quantum Computing with Matter Qubits and Photons

At this point we have pretty much all the ingredients that we need for linear optical quantum computing. One important component is still missing, however. Because we rely on post-selection and feed-forward in this quantum computer architecture, we need the ability to store the qubits (the single photons) from the time they are first entangled with the cluster state, to the time they are detected. This means we need an optical *quantum memory*.

4.1 Qubit Memories

Loosely speaking, let the *fault-tolerant threshold* be the maximum error beyond which no error correction can save the quantum computation. A quantum memory for linear optical quantum computing with single photons must then meet the following strict requirements:

1. The photon must couple *into* the memory with high enough probability to surpass the fault-tolerant threshold.
2. The photon must couple *out of* the memory with high enough probability to surpass the fault-tolerant threshold.
3. The mode shape of the output photon must be identical to that of the input photon in order to facilitate high-fidelity interferometry.

Moreover, the memory errors are cumulative, so if all three errors above are just below the fault-tolerant threshold, the total error will surely be above the threshold.

A typical quantum memory used in experiments is a fibre-optical delay line. However, due to losses in the fibre this is not a scalable solution (in a full-scale quantum computer, the memory time is likely to be several clock cycles long). Therefore, for optical quantum computing with single photons to become a viable technology,

we need some other system that can store the qubit value of the photon. Possibilities are atomic vapours, or systems with optical transitions strongly coupled to a cavity.

However, now we have just lost the advantage of the single photon as our qubit, namely its robustness against decoherence: the decoherence will now take place in the quantum memory. This means that we still have to create very robust matter qubits. So rather than trying to couple photons *into* the memory, we can engineer the capability of single-qubit operations into the memory and use the memories themselves as *matter qubits*. This removes requirement 1 above. In the next section, I will show that we can also remove requirement 2.

4.2 The Double-Heralding Protocol

Since it seems that we need some matter system as a quantum memory for optical quantum computing, we will explore this avenue further and start out with the assumption that the memory is actually our qubit. The qubit can generate a photon depending on its state, and if we can apply fusion-style gates on two such photons, we may be able to create cluster states in matter qubits. Let us consider a matter system with two energy levels $|\uparrow\rangle$ and $|\downarrow\rangle$ that make up the qubit states. An excited level $|e\rangle$ in the system couples only to the $|\uparrow\rangle$ level via an optical transition [8].

We entangle two qubits by first preparing two of these systems in separate cavities in the separable (unnormalised) state $(|\downarrow\rangle + |\uparrow\rangle)(|\downarrow\rangle + |\uparrow\rangle)$. Subsequently, we apply an optical π-pulse to each system, and wait for a single photon to be emitted. This yields the total state

$$|\downarrow\downarrow\rangle|0,0\rangle + |\downarrow\uparrow\rangle|0,1\rangle + |\uparrow\downarrow\rangle|1,0\rangle + |\uparrow\uparrow\rangle|1,1\rangle\,,$$

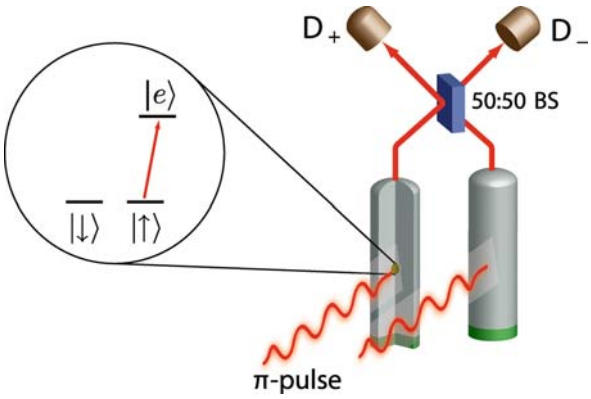


Fig. 6 Schematic of the double-heralding procedure. The two qubits are in separate physical systems and interact solely through their emitted photons. Path erasure of the photons generates the entanglement between the qubits

where $|0\rangle$ and $|1\rangle$ now denote the vacuum and a single photon in the freely propagating optical mode leaving the cavity, respectively (Fig. 6). When these two modes interact on a 50:50 beam splitter, the total state becomes (note the HOM effect)

$$|\downarrow\downarrow\rangle|0,0\rangle+\frac{1}{\sqrt{2}}\Big[(|\downarrow\uparrow\rangle+|\uparrow\downarrow\rangle)|0,1\rangle+(|\downarrow\uparrow\rangle-|\uparrow\downarrow\rangle|1,0\rangle+|\uparrow\uparrow\rangle(|2,0\rangle+|0,2\rangle)\Big].$$

Detecting both the outgoing modes of the beam splitter, each with a *realistic* detector (i.e. a detector with finite efficiency, and which cannot discriminate between optical states with one or more photons), gives the following state of the qubits (given just a single detector click in $D_{\pm}$):

$$\rho^{(\pm)}=f(\eta)|\Psi^{(\pm)}\rangle\langle\Psi^{(\pm)}|+[1-f(\eta)]|\uparrow\uparrow\rangle\langle\uparrow\uparrow|\,, \tag{37}$$

where $|\Psi^{(\pm)}\rangle=(|\downarrow\uparrow\rangle\pm|\uparrow\downarrow\rangle)/\sqrt{2}$ and $f(\eta)\leq 1$ is a function of the combined collection and detection efficiency, η.

The state in (37) is an incoherent mixture of a maximally entangled state and the separable state $|\uparrow\uparrow\rangle\langle\uparrow\uparrow|$. However, we can remove this separable part by first applying a bit flip operation $|\downarrow\rangle\leftrightarrow|\uparrow\rangle$ to both matter qubits. We subsequently apply a second π-pulse to each matter system. The separable part cannot generate photons. Thus, conditional on observing another single detector click, we obtain the final two-qubit pure state:

$$|\Psi^{(\pm)}\rangle=\frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle\pm|\uparrow\downarrow\rangle)\,. \tag{38}$$

The total success probability of this procedure is $\eta^2/2$. Note that we have removed requirement 1 of the quantum memory since we do not couple photons *into* the matter system, and we alleviated requirement 2 of quantum memories by allowing for a reduced success probability of the entangling operation. The remaining challenge is to make indistinguishable the photons originating from different qubits. Recently, a group in Paris managed to control two atoms in optical tweezers sufficiently well so that the photons they emit are indistinguishable enough to show two-photon quantum interference [9].

4.3 Creating Cluster States with Double-Heralding

The *double-heralding* entangling procedure described above is very similar to the type-II fusion gate, in that it effectively performs a projective parity measurement. Double heralding can therefore be used to create cluster states for universal quantum computing. Let us see how it works in detail.

We again use the formalism used in Lecture 3, where we write out the cluster state in terms of the conditional Z operations. It is straightforward to show that the action of the double-heralding procedure is given by

$$E_+ = |01\rangle\langle 01| + |10\rangle\langle 10| \quad \text{and} \quad E_- = |01\rangle\langle 01| - |10\rangle\langle 10|\,, \tag{39}$$

where we have identified $|\uparrow\rangle$ with $|0\rangle$ and $|\downarrow\rangle$ with $|1\rangle$, and the labels + and – of the operator E denote the detection signature. Suppose we have two cluster states that we wish to join using the double-heralding procedure. Before they are connected, their state can again be written as

$$\left(|0\rangle_1|\psi\rangle_a + |1\rangle_1 \prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right) \otimes \left(|0\rangle_2|\phi\rangle_b + |1\rangle_2 \prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right), \tag{40}$$

with the qubit neighbourhoods $n(j)$ defined as before. Applying the operator E_+ (i.e. a successful entangling operation) then yields the state

$$|0\rangle_1|1\rangle_2|\psi\rangle_a \left(\prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right) + |1\rangle_1|0\rangle_2 \left(\prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right)|\phi\rangle_b\,. \tag{41}$$

We need to show that this is again locally equivalent to a cluster state. To this end, apply a Hadamard operation H_2 to qubit 2 and a bit flip X_1 to qubit 1. Since both operators are part of the Clifford group, this will not destroy the cluster state:

$$\begin{aligned} |C\rangle = {} & |00\rangle_{12} \left(\prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right)|\phi\rangle_b + |01\rangle_{12} \left(\prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right)|\phi\rangle_b \\ & + |10\rangle_{12}|\psi\rangle_a \left(\prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right) - |11\rangle_{12}|\psi\rangle_a \left(\prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right). \end{aligned} \tag{42}$$

The question is: Is this another cluster state? In order to show that this is indeed the case, it is sufficient to show that we can transform it into a known form of a cluster state using local Clifford operations. So let us apply $\prod_j \mathcal{Z}_{j\in n(1)}$ to the qubits in set a. Since $\mathcal{Z}_j^2 = \mathbb{I}$, we have

$$\begin{aligned} |C'\rangle = {} & |00\rangle_{12}|\psi\rangle_a|\phi\rangle_b + |10\rangle_{12} \left(\prod_{j\in n(1)} \mathcal{Z}_j|\psi\rangle_a\right)\left(\prod_{j\in n(1)} \mathcal{Z}_j|\phi\rangle_b\right) \\ & + |01\rangle_{12}|\psi\rangle_a|\phi\rangle_b - |11\rangle_{12} \left(\prod_{j\in n(2)} \mathcal{Z}_j|\psi\rangle_a\right)\left(\prod_{j\in n(2)} \mathcal{Z}_j|\phi\rangle_b\right). \end{aligned} \tag{43}$$

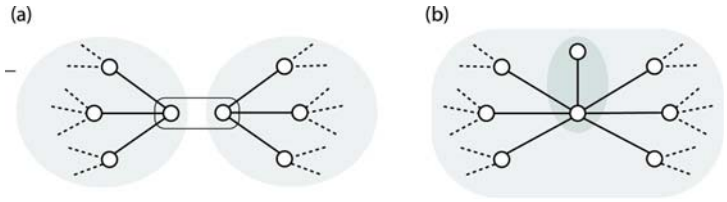


Fig. 7 Joining to cluster states. (**a**) The two separate clusters. (**b**) The double-heralding operation creates a redundantly encoded qubit (the qubits in the *dark shaded area*)

This can be written as

$$|C'\rangle = |0\rangle_1|\Psi\rangle_c\,(|0\rangle_2 + |1\rangle_2) + |1\rangle_1\left(\prod_{l\in n(1)} \mathcal{Z}_l|\Psi\rangle_c\right)(|0\rangle_2 - |1\rangle_2)$$
$$= |0\rangle_1|\Psi\rangle_c|+\rangle_2 + |1\rangle_1\left(\prod_{l\in n(1)} \mathcal{Z}_l\mathcal{Z}_2|\Psi\rangle_c|+\rangle_2\right). \tag{44}$$

It is clear that this is again of the form of a cluster state, since qubit 2 has experienced a Z operation depending on the state of qubit 1. We can in principle add qubit 2 to the set c and expand the neighbourhood $n(1)$. However, leaving it in this form reveals something interesting about the cluster. Qubit 2 is not entangled with any qubit other than qubit 1. We call this a leaf or a cherry in the cluster. More accurately, qubits 1 and 2 form a *redundantly encoded qubit*, useful for error correction (see Fig. 7b).

4.4 Complete Quantum Computer Architecture

Before I discuss the complete architecture of a quantum computer based on double heralding, let us explore some of the advantages and disadvantages of this approach.

The main advantage of the double-heralding protocol is that the resulting entanglement is completely independent of both the detector efficiency *and* the detector number-resolving capability. This is important, because it is extremely challenging to make photo-detectors with near perfect (>98%) efficiency while keeping unwanted dark counts low. Because of this insensitivity to photon collection efficiency, it is also not necessary for the qubit to be in the strong coupling regime of the interaction between the optical transition and the electromagnetic field. Another advantage is that the protocol is inherently distributed: it does not matter whether the qubits are 10 μm apart, or 10 km. This is extremely useful for quantum communication. But more importantly, it allows us to really isolate each individual qubit and get a good control over decoherence. In addition, a slowly varying (random) phase in one of the input modes of the beam splitter will give at most an unobservable global phase shift. Finally, the protocol requires only a relatively simple level structure. There are potentially many systems that can be used for this scheme, from trapped ions and atoms, to NV centres in diamond and Pauli blockade quantum dots.

There are two main disadvantages to double-heralding based quantum computing. First, the success probability of the entangling operation is bounded by one half, and with photon loss the probability becomes $\eta^2/2$, where η is the total photon collection efficiency. When the losses in the system are considerable, this makes the creation of cluster states a very costly affair (even though we maintain mathematical scalability at all times). Fortunately, there is a way to circumvent this problem and simultaneously keep the advantages of double heralding. It is called the *broker-client* model [10]. Instead of one qubit per site, we engineer two qubits with a high-fidelity, high efficiency (but non-scalable) two-qubit gate. An example of this is an NV centre in diamond, where the two qubits are the electron spin and the nuclear spin. The nuclear spin is long-lived, and can be used to store a qubit from a cluster state. The electron spin can then be entangled with other electron spins via double-heralding, and when this succeeds, the entanglement is transferred to the nuclear spin using the two-qubit gate. This way, we can build up large cluster states without suffering exploding overhead costs.

The second disadvantage of this scheme is that the qubits must be almost identical. If they are not, the photons are likely to carry some information about their origins, and the entangling procedure gives us only non-maximal entanglement. In terms of the HOM experiment, the cancellation of detection coincidences at the output modes is no longer complete. When the photo-detectors have good time resolution, we can counter this problem to some degree [11]: knowledge of the arrival times of the photons in the detectors will ensure that the resulting entangled state remains pure, and a sophisticated adaptive strategy of which qubits to entangle next allows for some variation in the qubits.

How do we put all this together? Figure 8 shows a quantum computer that operates on the double-heralding principle. It has five main components:

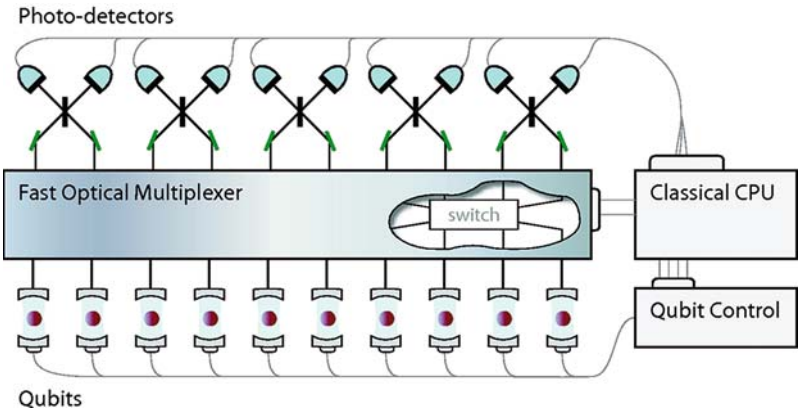


Fig. 8 Schematic of a quantum computer using double heralding. Maximum parallelizability is obtained using a fast optical multiplexer. A classical CPU is needed for the tracking of the measurement outcomes, the driving of the qubit control, and the setting of the multiplexer

1. The **Qubits** are kept in individual environments in order to keep decoherence to a minimum. In the broker-client model, there may be multiple qubits per site. The qubits must be nearly identical in order to create high-fidelity cluster states.
2. The **Qubit control** component is designed to address the individual qubits, applying both the π-pulses, the bit flips, and the single-qubit rotations needed for the qubit measurements. This may involve multiple lasers and/or microwave fields.
3. The **Optical multiplexer** is a router that directs the optical output modes of the qubits into the beam splitters. This way, we can apply the double-heralding procedure to two arbitrary qubits in the quantum computer. The 50:50 beam splitters that are drawn outside the multiplexer in Fig. 8 can be incorporated as well, so that we can in principle do a complete readout of all qubits in one clock cycle.
4. The **Photo-detectors** must have reasonably high detection efficiency and very low dark count rate. Good time resolution is also an advantage. There is no need for single-photon resolution.
5. The **Classical CPU** keeps track of the measurement outcomes, controls the switching of the multiplexer, and tells the qubit controller what to do. In addition, the CPU is used to program the quantum computer, and it interprets the final qubit readout.

Exercise 10. Calculate the effect of an unknown phase shift in one of the input modes of the beam splitter in the double-heralding protocol.

Exercise 11. Calculate the effect of partial which-path erasure.

Exercise 12. Verify the projective action of the double-heralding procedure.

Exercise 13. When we fail to add a micro-cluster to a cluster, how do we retrieve the cluster state?

5 Quantum Computing with Optical Nonlinearities

In the previous lecture we have seen how we can circumvent the need for quantum memories when we use material systems as qubits, together with a probabilistic entangling procedure. In this lecture, I show that we can obviate the need for quantum memories by choosing the right nonlinear interaction, which makes the entangling procedure (near) deterministic.

5.1 Kerr Nonlinearities

It has been known for a long time that we can make nonlinear optical gates using so-called Kerr nonlinearities. These consist of optically active materials that induce an effective photon–photon interaction. In particular, we consider the *cross-Kerr*

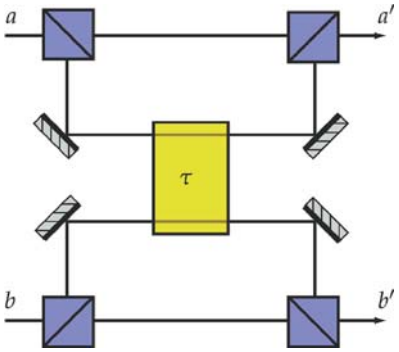


Fig. 9 A strong optical nonlinearity called a *cross-phase modulation* induces a phase shift on the vertically polarized part of mode b that depends on the number of vertically polarized photons in mode a. The setup is symmetric and creates a CZ gate

nonlinearity shown in Fig. 9. The central box has two input modes a and b, and the interaction Hamiltonian is of the form

$$H_K = \tau \hat{a}^\dagger \hat{a}\, \hat{b}^\dagger \hat{b}\,. \tag{45}$$

This leads to the following Bogoliubov transformations on the annihilation operators:

$$\hat{a} \to \hat{a}\, e^{\mathrm{i}\tau \hat{b}^\dagger \hat{b}}, \qquad \hat{b} \to \hat{b}\, e^{\mathrm{i}\tau \hat{a}^\dagger \hat{a}}, \tag{46}$$

in other words, the phase shift in mode a depends on the intensity of the field in mode b. It is straightforward to show that for $\tau = \pi$ and two polarized input photons in modes a and b, respectively, Fig. 9 represents a CZ gate.

Unfortunately, there are no real materials that have the properties that $\tau = \pi$ and are otherwise free of noise. The question thus arises: What can we do if $\tau = \theta \ll \pi$? The answer is that we can again construct a parity gate [12]. To this end, we make use of a reasonably bright coherent state that will carry the quantum correlations from one photon to the other. We consider the setup in Fig. 10. Let the two-qubit input state be $|\psi_{ab}\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$, and the coherent state is denoted by $|\alpha\rangle$. Assume that the interactions take place between $|1\rangle$ and $|\alpha\rangle$. The interaction is again a cross-Kerr nonlinearity, which produces a phase shift θ in the coherent state depending on the photon state in the signal mode. After the first interaction we obtain the three-mode optical state

$$|\psi_1\rangle = c_{00}|00\rangle|\alpha\rangle + c_{01}|01\rangle|\alpha\rangle + c_{10}|10\rangle|\alpha\, e^{\mathrm{i}\theta}\rangle + c_{11}|11\rangle|\alpha\, e^{\mathrm{i}\theta}\rangle\,, \tag{47}$$

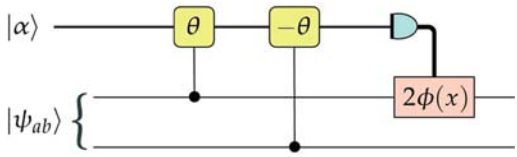


Fig. 10 A coherent state couples to two photonic qubits via the weak cross-Kerr nonlinearity indicated by θ. The measurement outcome determines a phase shift $2\phi(x)$ on one qubit

and after the second interaction we have

$$|\psi_2\rangle = c_{00}|00\rangle|\alpha\rangle + c_{01}|01\rangle|\alpha\, e^{-\mathrm{i}\theta}\rangle + c_{10}|10\rangle|\alpha\, e^{\mathrm{i}\theta}\rangle + c_{11}|11\rangle|\alpha\rangle\,. \tag{48}$$

We can separate this state into an even ($\{|00\rangle, |11\rangle\}$) and an odd ($\{|01\rangle, |10\rangle\}$) parity contribution.

The next step is to measure the $x = (\hat{a}+\hat{a}^\dagger)/\sqrt{2}$ quadrature of the coherent state. We see from (48) that such a measurement leaves the even parity subspace invariant, but not the odd subspace. To demonstrate a parity gate we calculate the projection of $|\psi_2\rangle$ onto the eigenstate of the measurement outcome $|x\rangle$:

$$\langle x|\psi_2\rangle = (c_{00}|00\rangle + c_{11}|11\rangle)\,\langle x|\alpha\rangle + c_{01}|01\rangle\langle x|\alpha\, e^{-\mathrm{i}\theta}\rangle + c_{10}|10\rangle\langle x|\alpha\, e^{\mathrm{i}\theta}\rangle\,. \tag{49}$$

Using (A4.12) on page 235 of Ref. [13]:

$$\langle x|\alpha\rangle = \frac{1}{\sqrt[4]{\pi}}\exp\left[-\frac{1}{2}\left(x-\sqrt{2}\alpha\right)^2 + \frac{1}{2}\alpha\left(\alpha-\alpha^*\right)\right], \tag{50}$$

and assuming that α is real,[4] we find that

$$\langle x|\alpha\rangle = \frac{1}{\sqrt[4]{\pi}}\exp\left[-\frac{1}{2}\left(x-\sqrt{2}\alpha\right)^2\right] \tag{51}$$

and

$$\langle x|\alpha e^{\mathrm{i}\theta}\rangle = \frac{1}{\sqrt[4]{\pi}}\exp\left[-\frac{1}{2}\left(x-\sqrt{2}\alpha\,\cos\theta\right)^2 + \mathrm{i}\alpha\,\sin\theta\,(\sqrt{2}x-\alpha\,\cos\theta)\right]. \tag{52}$$

The state after the measurement is therefore

$$|\psi'_{ab}\rangle = \langle x|\alpha\rangle\,(c_{00}|00\rangle + c_{11}|11\rangle) + |\langle x|\alpha\, e^{\mathrm{i}\theta}\rangle|\left(c_{01}\, e^{-\mathrm{i}\phi}|01\rangle + c_{10}\, e^{\mathrm{i}\phi}|10\rangle\right), \tag{53}$$

[4] In addition, we describe the coherent state in the co-rotating frame of reference, which allows us to suppress the free time evolution of the coherent state. In particular, this means that α is real for all times, and the nonlinear phase shift θ is included explicitly.

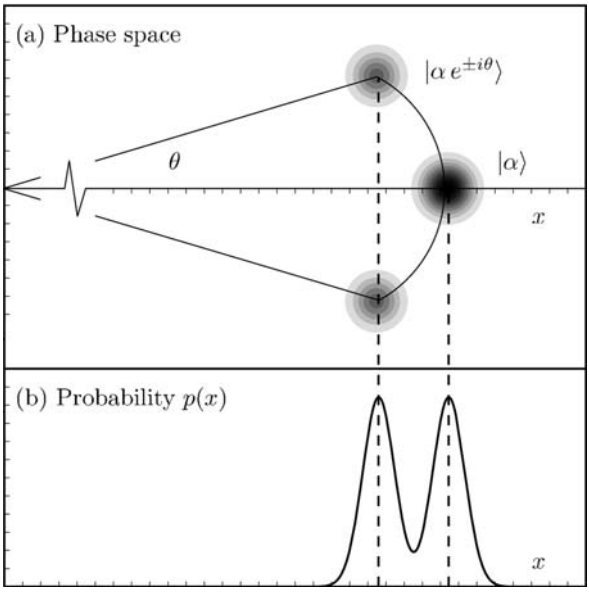


Fig. 11 (**a**) Phase space representation of the weak nonlinear parity gate. (**b**) The corresponding probability of the outcomes of an x-quadrature measurement. The overlap between the two Gaussian peaks must be made sufficiently large for the gate to work near deterministically

with

$$\phi(x) \equiv \alpha\, \sin\theta\, (\sqrt{2}x - \alpha\, \cos\theta)\,. \tag{54}$$

A phase space representation of the above procedure is given in Fig. 11. The relative phase $2\phi(x)$ can be corrected using regular phase shifts.

These two distributions peak at different values x_e and x_o for the even and odd subspace, respectively,

$$x_e = \sqrt{2}\alpha \qquad \text{and} \qquad x_o = \sqrt{2}\alpha\, \cos\theta\,. \tag{55}$$

The width of (the real part of) these distributions is of the order one. We can distinguish the two peaks (and correspondingly obtain a high fidelity) when $x_e - x_o$ is larger than twice the width of the distribution ($\theta \ll 1$):

$$x_e - x_o > 2 \quad \Leftrightarrow \quad \sqrt{2}\alpha\,(1 - \cos\theta) > 2 \quad \Leftrightarrow \quad \alpha\, \frac{\theta^2}{2\sqrt{2}} > 1\,. \tag{56}$$

Weak nonlinearities on the order of 10^{-5} can be achieved using electromagnetically induced transparencies. There are several tricks that can be used to increase the performance of this gate [14].

An x-quadrature measurement that can project the two-photon state onto either one of the parity subspaces is again a parity gate, and we have seen earlier how these projections are useful for quantum computing. Here, the parity projection is practically deterministic if the peak separation is big enough, which means that cluster state growth can be very efficient.

5.2 Zeno Gates

Another optical nonlinearity that may be used to construct a near-deterministic two-photon gate is two-photon absorption. This is the basis of the so-called Zeno gate by Franson, Jacobs, and Pittman [15], and is shown in Fig. 12. The gate works similarly to the strong Kerr gate shown in Fig. 9, but the detailed physics of the central (yellow) box differs.

Before I describe the Zeno gate, let us look at a possible experimental implementation of a beam splitter: typically, we think of a beam splitter as a semi-reflective mirror, but there are also other ways. Many people who study photons in optical fibres make the beam splitters with fibres as well, so how is this done? In general, *any* unitary two-mode transformation can be described by the matrix given in Eq. (8). If we take a length of fibre and splice it at both ends, we end up with two fibres that join for a certain length, and then separate again. Because the action of such a physical object behaves according to Eq. (8), we can model this as a beam splitter, where the transmission coefficient is now related to the length of the joined piece of fibre. The evolution is therefore something like this:

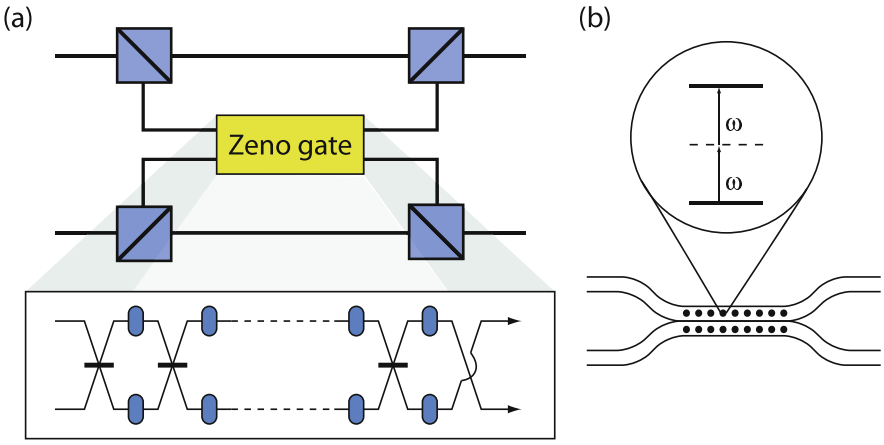


Fig. 12 (**a**) A CZ gate using the Zeno effect. The n beam splitters have transmittivity $1/n$, and are separated by dissipative two-photon absorbers. In the limit of $n \to \infty$, and perfect two-photon absorption, the Zeno gate implements a perfect CZ gate. (**b**) The beam splitters are implemented using a spliced fibre with two cores, each filled with atoms that absorb two-photon excitations

$$
\begin{aligned}
|00\rangle &\rightarrow |00\rangle, \\
|01\rangle &\rightarrow \cos\theta|01\rangle + \mathrm{i}\sin\theta|10\rangle, \\
|10\rangle &\rightarrow \mathrm{i}\sin\theta|01\rangle + \cos\theta|10\rangle, \\
|11\rangle &\rightarrow \cos 2\theta|11\rangle + \frac{\mathrm{i}\sin 2\theta}{\sqrt{2}}\left(|20\rangle + |02\rangle\right),
\end{aligned} \tag{57}
$$

where 0, 1, and 2 denote the photon occupation number of the mode, and we have chosen a convenient phase convention.[5]

The Zeno gate works like this spliced fibre, with a small but essential modification: consider two fibres with a hollow core that come together at some point, remain parallel for a certain distance, and then separate. Again, this is properly described by (8) and (57). Photons entering one fibre couple to the other fibre via the evanescent electromagnetic field, and they can tunnel from one core to the other. If the length is chosen correctly, we can make the photons come out in the other fibre. When two photons enter the device, one in each input fibre, we can also demonstrate the Hong-Ou-Mandel effect.[6]

To make the Zeno gate, we fill both cores with a linear array of atoms that have a strong two-photon absorption and negligible single-photon absorption. When only a single photon enters the device, the atoms have no effect on the dynamics, and the photon exits in some superposition of the output modes. On the other hand, when two photons enter the device, one in each input mode, they are prevented from building up the two-photon amplitude due to the absorption: after a very short distance, the two-photon input state evolves according to (57) with $\theta \ll 1$. When the photons encounter the first atom, the term in the superposition with two photons in one mode will transform into the vacuum because the atom absorbs the photons and dissipates the energy into the environment. This is effectively a measurement where we throw away the measurement outcome.

Since the length of free evolution in the fibre cores is so short (the atoms are placed closely together), the amplitude $\mathrm{i}\sin\theta$ is very small, and the probability of two-photon absorption is also tiny. Therefore, after the photons encounter the atoms, they are projected onto the $|11\rangle$ state with very high probability. En route to the next atom, they will evolve again, and this procedure repeats until the fibre cores separate. The atoms act as an almost continuous measurement, preventing the state from building up an appreciable absorption amplitude. This is commonly known as the *Zeno* effect. The photons will exit the interaction region in different modes, due to the suppression of the HOM effect.

Why does this work as a two-photon gate? To answer this, we look at the accumulated phases of the four possible input states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$. Clearly, the state $|00\rangle$ remains unchanged, because there are no photons at all. The length of the interaction is chosen such that a single-photon input ($|01\rangle$ or $|10\rangle$) is transmit-

[5] We can always include phase shifts in the fibres to make the interaction of this form.

[6] For the HOM effect to take place, the length of joined fibre must be half the length of the fibre in the case where single photons enter one core and exit the other.

ted perfectly into the other core. The accumulated phase for a transmitted photon is $e^{i\pi/2}$, according to (57). Finally, when two photons enter the device (i.e. the state $|11\rangle$), the beam splitter action is suppressed, and the photons are effectively reflected. The phase associated with perfect reflection is 1. By choosing suitable phase shifts in the output modes, this interaction can be turned into a CZ gate:

$$U_{\text{Zeno}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \rightarrow \text{phase shift} \quad U_{\text{CZ}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{58}$$

This is the physical intuition behind the Zeno gate. Let us derive this result slightly more formally.

We consider the ideal case where the two-photon absorption is perfect, and there is no single-photon absorption (or loss). Since the two-photon absorption is followed by spontaneous emission into the environment, the evolution is incoherent, and we can no longer use a pure state description of the situation. We therefore construct the positive operator-valued measures (POVMS) for the different measurement outcomes [16]. In general, an arbitrary input state ρ will evolve according to

$$\rho \;\rightarrow\; \tilde{\rho} = \mathcal{L}(\rho) \equiv \sum_{k=1,2} A_k \rho A_k^\dagger, \tag{59}$$

where the A_k are the Kraus operators (or *effects*) that define the effect of the measurement on the state. Each measurement outcome is represented by a specific A_k. Since we discard the measurement outcomes in our Zeno gate, we need to sum over all k. Note that here we are talking about the state ρ of one optical mode (or one fibre core). The Kraus operators satisfy the relation $\sum_k A_k^\dagger A_k = \mathbb{I}$, which ensures that $\tilde{\rho}$ is a proper density operator.

In this case we have two Kraus operators: one when there is no absorption, and one for two-photon absorption. When there is no absorption, nothing happens, and the corresponding Kraus operator is the *identity* operator on the relevant subspace (spanned by $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$). On the other hand, two-photon absorption can be formalized as changing the state $|2\rangle$ into $|0\rangle$. We therefore have

$$A_1 = |0\rangle\langle 0| + |1\rangle\langle 1| \quad \text{and} \quad A_2 = |0\rangle\langle 2|. \tag{60}$$

In order to evaluate the effect of the Zeno gate we need to apply the super-operator $\mathcal{L}(\rho)$ to both modes every time an atom is encountered. Clearly, $\mathcal{L}(\rho)$ is acting as the identity if there is at most one photon in the system, so $|01\rangle \rightarrow \mathrm{i}|10\rangle$ and $|10\rangle \rightarrow \mathrm{i}|01\rangle$. But what about the $|11\rangle$ term?

The density operator for the state $|11\rangle$ is given by $|11\rangle\langle 11|$, and the beam splitter evolution in (57) will give

$$\rho = \cos^2(2\theta)|11\rangle\langle 11| + \frac{1}{2}\sin^2(2\theta)\left(|20\rangle + |02\rangle\right)\left(\langle 20| + \langle 02|\right) + \frac{\mathrm{i}}{\sqrt{2}}\cos(2\theta)\sin(2\theta)\left[\left(|20\rangle + |02\rangle\right)\langle 11| - |11\rangle\left(\langle 20| + \langle 02|\right)\right]. \quad (61)$$

When we apply the super-operator $\mathcal{L}(\rho)$, we find that

$$\tilde{\rho} = \sum_k A_k \rho A_k^\dagger = \cos^2(2\theta)|11\rangle\langle 11| + \sin^2(2\theta)|00\rangle\langle 00|. \quad (62)$$

The term $|00\rangle\langle 00|$ is invariant under both the beam splitter evolution and the two-photon absorption, and does not change during the remainder of the gate. The $|11\rangle\langle 11|$ term will again undergo the evolution in (62). After the full length of the joined fibre (involving n atoms), the evolution is

$$|11\rangle\langle 11| \;\rightarrow\; \cos^{2n}(2\theta)|11\rangle\langle 11| + (1 - \cos^{2n}(2\theta))|00\rangle\langle 00|. \quad (63)$$

The case of the input state $|01\rangle$ is symmetric to the input state $|10\rangle$, so we need to discuss only one of them here. Since the Kraus operator in this subspace is the identity operator, the evolution is a series of n rotations over angle θ, which can be written as

$$\begin{aligned} |01\rangle &\rightarrow \cos(n\theta)|01\rangle + \mathrm{i}\sin(n\theta)|10\rangle, \\ |10\rangle &\rightarrow \mathrm{i}\sin(n\theta)|01\rangle + \cos(n\theta)|10\rangle. \end{aligned} \quad (64)$$

These are all the ingredients we need to analyse the ideal Zeno gate.

Remember that for the Zeno gate to work the single-photon input states must be swapped (perfect transmission), while total reflection must occur when there are two photons entering the device, one in each input mode. Therefore, we must choose $n\theta = \pi/2$, and this generates a phase shift i on the photon.

Using this choice of θ, the probability amplitude of the $|11\rangle$ term becomes $\cos^n(\pi/n)$. In order to have a proper Zeno effect, n must be very large. We can expand the cosine function to first order and take the limit of n to infinity:

$$\lim_{n\to\infty}\left(1 - \frac{\pi^2}{2n^2}\right)^n = 1. \quad (65)$$

Indeed, the two-photon absorption $[1 - \cos^{2n}(\pi/n)]$ is completely suppressed. Furthermore, the phase of the $|11\rangle$ term is unaffected.

We now have the situation where single photons are transmitted (and accumulate a phase i), while two input photons are both reflected (and do not experience a phase shift). A simple swap of the output modes will then result in $|01\rangle \leftrightarrow |10\rangle$, and the transformation becomes of the form of U_{Zeno} in (58).

So far, we analysed the Zeno gate in the ideal case of perfect two-photon absorption and no single-photon absorption using POVMS. When the situation is not ideal

(e.g. in the case of survival of the $|20\rangle$ and $|02\rangle$ terms and photon loss), the Kraus operators need to be modified and the calculation will become much harder. Alternatively, the problem can be formulated in Lindblad form or in terms of a master equation, which can then be solved using standard techniques [13].

Exercise 14. Prove (46).

Exercise 15. Verify (53) and (54).

Exercise 16. Calculate the fidelity of this gate for an equal-superposition input state $c_{00} = c_{01} = c_{10} = c_{11} = \frac{1}{2}$.

Exercise 17. Verify (57) and (58).

Exercise 18. Check that the Kraus operators in (60) obey the normalization condition, and verify (62).

Final Remarks

Single photons are very resilient to decoherence, and they travel at very high speed. This makes them the ideal carriers for quantum information. It is therefore likely that optical systems will play an important role in future quantum information technology. However, the lack of a direct interaction between photons means that some trickery must be used if you want them to carry out quantum computations. There have been several proposals for optical quantum computers, the best known of which is the Knill–Laflamme–Milburn scheme using only photons, linear optics, and photo-detectors. This scheme, and its improvements, needs quantum memories, because they rely critically on the feed-forward of measurement outcomes to modify subsequent interferometry. As a consequence, the advantage of photons as slow-decohering qubits is lost: decoherence now takes place in the memory.

In these lectures, I have argued that you can either choose to live with it and make the quantum memories your qubits (which of course means that you have to engineer high-quality qubits), or you can turn to nonlinear interactions to create deterministic gates (an equally daunting task). At this point it is not clear what the greater challenge is. Finally, all the results presented here rely to a greater or lesser extent on the ability to create identical single-photon wave packets.

Acknowledgements I thank Simon Benjamin and Dan Browne for valuable discussions, and Erik Gauger for carefully reading the manuscript. I also thank Francesco Petruccione and the University of KwaZulu-Natal in South Africa for inviting me to give these lectures. This work was done as part of the QIP IRC www.qipirc.org (GR/S82176/01).

References

1. P. Kok, W.J. Munro, K. Nemoto, T.C. Ralph, J.P. Dowling, and G.J. Milburn: Rev. Mod. Phys. **79**, 135 (2005) 188
2. M. Reck, A. Zeilinger, H.J. Bernstein, and P. Bertani: Phys. Rev. Lett. **73**, 58 (1994) 191

3. E. Knill, R. Laflamme, and G.J. Milburn: Nature **409**, 46 (2001) 192, 193
4. D. Gottesman and I.L. Chuang: Nature **402**, 390 (1999) 193
5. C.K. Hong, Z.Y. Ou, and L. Mandel: Phys. Rev. Lett. **57**, 2044 (1987) 194
6. D.E. Browne and T. Rudolph: Phys. Rev. Lett. **95**, 010501 (2005) 196, 198
7. R. Raussendorf and H.J. Briegel: Phys. Rev. Lett. **86**, 5188 (2001) 198
8. S.D. Barrett and P. Kok: Phys. Rev. A **71**, R060310 (2005) 205
9. J. Beugnon, M.P.A. Jones, J. Dingjan, B. Darquié, G. Messin, A. Browaeys, and P. Grangier: Nature **440**, 7085 (2007) 206
10. S.C. Benjamin, D.E. Browne, J. Fitzsimons, and J.J.L. Morton: New J. Phys. **8**, 141 (2006) 209
11. E.T. Campbell, J. Fitzsimons, S.C. Benjamin, and P. Kok: Phys. Rev. A **75**, 042303 (2007) 209
12. S.D. Barrett, P. Kok, K. Nemoto, R.G. Beausoleil, W.J. Munro, and T.P. Spiller: Phys. Rev. A **71**, R060302 (2005) 211
13. S.M. Barnett and P.M. Radmore: *Methods in Theoretical Quantum Optics* (Clarendon Press, Oxford, 1997) 212, 218
14. T.P. Spiller, K. Nemoto, S.L. Braunstein, W.J. Munro, P. van Loock, and G.J. Milburn: New J. Phys. **8**, 30 (2006) 213
15. J.D. Franson, B.C. Jacobs, and T.B. Pittman: Phys. Rev. A **70**, 062302 (2004) 214
16. K. Kraus: *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983) 216

Quantum Information and Relativity: An Introduction

Daniel R. Terno

Information and physics are closely and fascinatingly intertwined. Their relations become even more interesting when we leave a non-relativistic quantum mechanics for more exiting venues. My notes are planned as a guided tour for the first steps along that road, and at the same time as an illustration of the basic notions of quantum information theory in more exotic settings.

I start from a brief introduction to causality restrictions on the distributed measurements: the limitations that are imposed by final propagation velocity of the physical interactions. It is followed by the relativistic transformations of the states of massive particles and photons, from which we can deduce what happens to qubits which are realized as the discrete degrees of freedom. Building on this, I discuss the distinguishability of quantum signals, and briefly touch communication channels and the bipartite entanglement. Creation of the entanglement in scattering processes will be touched only briefly, while the fascinating experimental results on the entanglement in particle physics are discussed in the contribution by Beatrix Hiesmayr.

I do not follow a historical order or give all of the original references. A review [1] is used as the standard reference on quantum information and relativity. The results of the "usual" quantum information and the formalism of the measurement theory are usually given without any reference. In this case they can be found in at least one of the sources [2–6]. Finally, a word about units: $\hbar = c = 1$ are always assumed.

1 Causality and Distributed Measurements

Here I present the causality constraints on quantum measurements. For simplicity, measurements are considered to be point-like interventions. First recall the standard description of the measurement and the induced state transformation. Consider a system in the state ρ that is subject to measurement that is described by a positive

D.R. Terno (✉)
Centre for Quantum Computer Technology, Department of Physics
Macquarie University, Australia
e-mail: deterno@physics.mg.edu.au

Terno, D.R.: *Quantum Information and Relativity: An Introduction*. Lect. Notes Phys. **787**, 221–247 (2010)
DOI 10.1007/978-3-642-02871-7_8 © Springer-Verlag Berlin Heidelberg 2010

operator-valued measure (POVM) $\{E_\mu\}$. The probability of the outcome μ is

$$p_\mu = \mathrm{Tr}\, E_\mu \rho, \tag{1}$$

while the state transformation is given by some completely positive evolution:

$$\rho \to \rho'_\mu = \sum_m A_{\mu m} \rho A^\dagger_{\mu m} / p_\mu, \qquad \sum_m A^\dagger_{\mu m} A_{\mu m} = E_\mu. \tag{2}$$

If the outcome is left unknown, the update rule is

$$\rho \to \rho = \sum_{\mu m} A_{\mu m} \rho A^\dagger_{\mu m}. \tag{3}$$

Now consider a bipartite state ρ_{AB}. The operations of Alice and Bob are given by the operators $A_{\mu m}$ and $B_{\nu n}$, respectively. It is easy to see that if these operators commute,

$$[A_{\mu m}, B_{\nu n}] = 0, \tag{4}$$

then the observation statistics of Bob is independent of Alice's results and vice versa. Indeed, the probability that Bob gets a result ν, irrespective of what Alice found, is

$$p_\nu = \sum_\mu \mathrm{tr} \Big(\sum_{m,n} B_{\nu n}\, A_{\mu m}\, \rho\, A^\dagger_{\mu m}\, B^\dagger_{\nu n} \Big). \tag{5}$$

Now make use of (4) to exchange the positions of $A_{\mu m}$ and $B_{\nu n}$, and likewise those of $A^\dagger_{\mu m}$ and $B^\dagger_{\nu n}$, and then we move $A_{\mu m}$ from the first position to the last one in the product of operators in the traced parenthesis. Since the elements of a POVM satisfy $\sum_\mu E_\mu = \mathbb{1}$, (5) reduces to

$$p_\nu = \mathrm{tr} \Big(\sum_n B_{\nu n}\, \rho\, B^\dagger_{\nu n} \Big), \tag{6}$$

whence all the expressions involving Alice's operators $A_{\mu m}$ have totally disappeared. The statistics of Bob's result are not affected at all by what Alice may simultaneously do somewhere else. This proves that (4) indeed is a sufficient condition for no instantaneous information transfer. In particular, the local operations $A \otimes \mathbb{1}_B$ and $\mathbb{1}_A \otimes B$ are of this form.

Note that any classical communication between distant observers can be considered a kind of long range interaction. The propagation of signals is, of course, bounded by the velocity of light. As a result, there exists a partial time ordering of the various interventions in an experiment, which defines the notions earlier and later. The input parameters of an intervention are deterministic (or possi-

bly stochastic) functions of the parameters of earlier interventions, but not of the stochastic outcomes resulting from later or mutually spacelike interventions [1].

Even these apparently simple notions lead to non-trivial results. Consider a separable bipartite superoperator T,

$$T(\rho) = \sum_k M_k \rho M_k^\dagger, \qquad M_k = A_k \otimes B_k, \tag{7}$$

where the operators A_k represent operations of Alice and B_k those of Bob. Not all such superoperators can be implemented by local transformations and classical communication (LOCC) [7]. This is the foundation of the "non-locality without entanglement".

A classification of bipartite state transformations was introduced in [8]. It consists of the following categories. There are *localizable* operations that can be implemented locally by Alice and Bob, possibly with the help of prearranged ancillas, but without classical communication. Ideally, local operations are instantaneous, and the whole process can be viewed as performed at a definite time. A final classical output of such distributed intervention will be obtained at some point of the (joint) causal future of Alice's and Bob's interventions. For *semilocalizable* operations, the requirement of no communication is relaxed and one-way classical communication is possible. It is obvious that any tensor-product operation $T_\mathrm{A} \otimes T_\mathrm{B}$ is localizable, but it is not a necessary condition. For example the Bell measurements, which distinguishes between the four standard bipartite entangled qubit states,

$$|\Psi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|1\rangle \pm |1\rangle|0\rangle), \qquad |\Phi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|0\rangle \pm |1\rangle|1\rangle), \tag{8}$$

are localizable.

Other classes of bipartite operators are defined as follows: Bob performs a local operation T_B just before the global operation T. If no local operation of Alice can reveal any information about T_B, i.e., Bob cannot signal to Alice, the operation T is *semicausal*. If the operation is semicausal in both directions, it is *causal*. In many cases it is easier to prove causality than localizability (see Remark 3). There is a necessary and sufficient condition for the semicausality (and therefore, the causality) of operations [8].

These definitions of causal and localizable operators appear equivalent. It is easily proved that localizable operators are causal. It was shown that semicausal operators are always semilocalizable [9]. However, there are causal operations that are not localizable [8].

It is curious that while a complete Bell measurement is causal, the two-outcome incomplete Bell measurement is not. Indeed, consider a two-outcome PVM

$$E_1 = |\Phi^+\rangle\langle\Phi^+|, \qquad E_2 = \mathbb{1} - E_1. \tag{9}$$

If the initial state is $|01\rangle_{\rm AB}$, then the outcome that is associated with E_2 always occurs and Alice's reduced density matrix after the measurement is $\rho_{\rm A} = |0\rangle\langle 0|$. On the other hand, if before the joint measurement Bob performs a unitary operation that transforms the state into $|00\rangle_{\rm AB}$, then the two outcomes are equiprobable, the resulting states after the measurement are maximally entangled, and Alice's reduced density matrix is $\rho_{\rm A} = \frac{1}{2}\mathbb{1}$. A simple calculation shows that after this incomplete Bell measurement two input states $|00\rangle_{\rm AB}$ and $|01\rangle_{\rm AB}$ are distinguished by Alice with a probability of 0.75.

Here is another example of a semicausal and semilocalizable measurement which can be executed with one-way classical communication from Alice to Bob. Consider a PVM measurement, whose complete orthogonal projectors are

$$|0\rangle \otimes |0\rangle, \quad |0\rangle \otimes |1\rangle, \quad |1\rangle|+\rangle, \quad |1\rangle \otimes |-\rangle, \tag{10}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The Kraus matrices are

$$A_{\mu j} = E_\mu \delta_{j0}. \tag{11}$$

From the properties of complete orthogonal measurements [8], it follows that this operation cannot be performed without Alice talking to Bob. A protocol to realize this measurement is the following. Alice measures her qubit in the basis $\{|0\rangle, |1\rangle\}$, and tells her result to Bob. If Alice's outcome was $|0\rangle$, Bob measures his qubit in the basis $\{|0\rangle, |1\rangle\}$, and if it was $|1\rangle$, in the basis $\{|+\rangle, |-\rangle\}$.

If one allows for more complicated conditional state evolution [10], then more measurements are localizable. In particular, consider a *verification* measurement, i.e., the measurement yields a μ-th result with certainty, if the state prior to the classical interventions was given by $\rho = E_\mu$, but without making any specific demand on the resulting state ρ'_μ.

It is possible to realize a verification measurement by means of a shared entangled ancilla and Bell-type measurement by one of the parties [11]. Verification measurement of Eq. (10) can illustrate this construction. In addition to the state to be tested, Alice and Bob share a Bell state $|\Psi^-\rangle$. They do not have to coordinate their moves. Alice and Bob perform tasks independently and convey their results to a common center, where a final decision is made.

The procedure is based on the teleportation identity

$$|\Psi\rangle_1|\Psi^-\rangle_{23} = \tfrac{1}{2}\left(|\Psi^-\rangle_{12}|\Psi\rangle_3 + |\Psi^+\rangle_{12}|\tilde{\Psi}^{(z)}\rangle_3 + |\Phi^-\rangle_{12}|\tilde{\Psi}^{(x)}\rangle_3 + |\Phi^+\rangle_{12}|\tilde{\Psi}^{(y)}\rangle_3\right), \tag{12}$$

where $|\tilde{\Psi}^{(z)}\rangle$ means the state $|\Psi\rangle$ rotated by π around the z-axis, etc. The first step of this measurement corresponds to the first step of a teleportation of a state of the spin from B (Bob's site) to A (Alice's site). Bob and Alice do not perform the full teleportation (which requires a classical communication between them). Instead, Bob performs only the Bell measurement at his site which leads to one of the branches of the superposition in the rhs of Eq. (12).

The second step of the verification measurement is taken by Alice. Instead of completing the teleportation protocol, she measures the spin of her particle in the z direction. According to whether that spin is up or down, she measures the spin of her ancilla in the z or x direction, respectively. This completes the measurement and it only remains to combine the local outcomes to get the result of the nonlocal measurement [11]. This method can be extended to arbitrary Hilbert space dimensions.

Remarks

1. Measurements in quantum field theory are discussed in [1, 6].
2. An algebraic field theory approach to statistical independence and to related topics is presented in [12, 13].
3. To check the causality of an operation T whose outcomes are the states $\rho_\mu = T_\mu(\rho)/p_\mu$ with probabilities $p_\mu = \operatorname{tr} T_\mu(\rho)$, $\sum_\mu p_\mu = 1$ it is enough to consider the corresponding superoperator:

$$T'(\rho) := \sum_\mu T_\mu(\rho). \tag{13}$$

   Indeed, assume that Bob's action prior to the global operation lead to one of the two different states ρ_1 and ρ_2. Then the states $T'(\rho_1)$ and $T'(\rho_2)$ are distinguishable if and only if some of the pairs of states $T_\mu(\rho_1)/p_{\mu 1}$ and $T_\mu(\rho_2)/p_{\mu 2}$ are distinguishable. Such probabilistic distinguishability shows that the operation T is not semicausal.
4. Absence of the superluminal communication makes it possible to evade the theorems on the impossibility of a bit commitment. In particular the protocol RBC2 allows a bit commitment to be indefinitely maintained with unconditionally security against all classical attacks, and at least for some finite amount of time against quantum attacks [14, 15].
5. In these notes I am not going to deal with the relativistic localization POVM. Their properties (and difficulties in their construction) can be found in [1]. An exhaustive survey of the spatial localization of photons is given in [16]. Here we only note in passing that if $E(\mathcal{O})$ is an operator that corresponds to the detection of an event in a spacetime region $\mathcal{O}$, since they are not thought to be implemented by physical operations confined to that spacetime area, the condition $[E(\mathcal{O}_1), E(\mathcal{O}_2)] = 0$ is not required [17, 18].

2 Quantum Lorentz Transformations

There is no elementary particle that is called "qubit". Qubits are realized by particular degrees of freedom of more or less complicated systems. To decide how qubits transform (e.g., under Lorentz transformations) it may be necessary to consider again the entire system. In the following our qubit will be either a spin of a massive particle or a polarization of a photon. A quantum Lorentz transformation connects

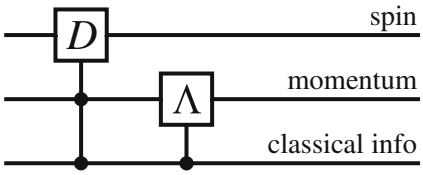


Fig. 1 Relativistic state transformation as a quantum circuit: the gate D which represents the matrix $D_{\xi\sigma}[W(\Lambda, p)]$ is controlled by both the classical information and the momentum p, which is itself subject to the classical information Λ

the description of a quantum state $|\Psi\rangle$ in two reference frames that are connected by a Lorentz transformation Λ (i.e., their coordinate axes are rotated with respect to each other and the frames have a fixed relative velocity). Then $|\Psi'\rangle = U(\Lambda)|\Psi\rangle$, and the unitary $U(\Lambda)$ is represented on Fig. 1. The purpose of this section is to explain the elements of this quantum circuit.

From the mathematical point of view the single-particle states belong to some irreducible representation of the Poincaré group. An introductory discussion of these representations and their relations with states and quantum fields may be found, e.g., in [19, 20]. Within each particular irreducible representation there are six commuting operators. The eigenvalues of two of them are invariants that label the representation by defining the mass m and the intrinsic spin j. The basis states are labeled by three components of the momentum $\mathbf{p}$ and the spin operator Σ_3. Hence a generic state is given by

$$|\Psi\rangle = \sum_{\sigma} \int d\mu(p)\psi_\sigma(p)|p, \sigma\rangle. \tag{14}$$

In this formula $d\mu(p)$ is the Lorentz-invariant measure,

$$d\mu(p) = \frac{1}{(2\pi)^3}\frac{d^3\mathbf{p}}{2E(\mathbf{p})}, \tag{15}$$

where the energy $E(\mathbf{p}) = p^0 = \sqrt{\mathbf{p}^2 + m^2}$. The improper momentum and spin eigenstates are δ-normalized,

$$\langle p, \sigma|q, \sigma'\rangle = (2\pi)^3(2E(\mathbf{p}))\delta^{(3)}(\mathbf{p} - \mathbf{q})\delta_{\sigma\sigma'}, \tag{16}$$

and are complete on the one-particle space, which is $\mathcal{H} = \mathbb{C}^{2j+1}\otimes L^2(\mathbb{R}^3, d\mu(p))$ for spin-j fields. To find the transformation law we have to be more concrete about the spin operator. The operator $\Sigma_3(p)$ is a function of the generators of the Poincaré group. One popular option is helicity, $\Sigma_3 = \mathbf{J}\cdot\mathbf{P}/|\mathbf{P}|$, where $\mathbf{J}$ is the angular momentum and $\mathbf{P}$ is the particle's 3-momentum. This quantity is well-defined for both massive and massless particles. For massive particles we use the z-component of the rest-frame (or Wigner spin), that we now describe in the next section.

2.1 Massive Particles

The construction involves picking a reference 4-momentum k, which for massive particles is taken to be $k_R = (m, \mathbf{0})$. The Wigner spin $\mathbf{S}(p)$ is defined to coincide with the non-relativistic spin $\mathbf{S}$ in particle's rest frame. The state of a particle at rest is labeled $|k_R, \sigma\rangle$:

$$\mathbf{S}^2|k_R, \sigma\rangle = j(j+1)|k_R, \sigma\rangle, \qquad \mathbf{S}_3|k_R, \sigma\rangle = \sigma|k_R, \sigma\rangle. \tag{17}$$

The spin states of arbitrary momenta are defined as follows. The standard rotation-free boost that brings k_R to an arbitrary momentum p, $p^\mu = L(p)^\mu{}_\nu k^\nu$ is given by

$$L(p) = \begin{pmatrix} \frac{E}{m} & \frac{p_1}{m} & \frac{p_2}{m} & \frac{p_3}{m} \\ \frac{p_1}{m} & 1 + \frac{p_1^2}{m(m+E)} & \frac{p_1 p_2}{m(m+E)} & \frac{p_1 p_3}{m(m+E)} \\ \frac{p_2}{m} & \frac{p_2 p_1}{m(m+E)} & 1 + \frac{p_2^2}{m(m+E)} & \frac{p_2 p_3}{m(m+E)} \\ \frac{p_3}{m} & \frac{p_3 p_1}{m(m+E)} & \frac{p_3 p_2}{m(m+E)} & 1 + \frac{p_3^2}{m(m+E)} \end{pmatrix}. \tag{18}$$

The Wigner spin $\mathbf{S}(p)$ and the one-particle basis states are defined by

$$|p, \sigma\rangle \equiv U[L(p)]|k_R, \sigma\rangle, \qquad \mathbf{S}_3(p)|p, \sigma\rangle = \sigma|p, \sigma\rangle. \tag{19}$$

In deriving the transformation rules we begin with the momentum eigenstates. Using the group representation property and (19) the transformation is written as

$$U(\Lambda) = U[L(\Lambda p)]U[L^{-1}(\Lambda p)\Lambda L(p)]U[L^{-1}(p)]. \tag{20}$$

The element of the Lorentz group

$$W(\Lambda, p) \equiv L^{-1}(\Lambda p)\Lambda L(p), \tag{21}$$

leaves k_R invariant, $k_R = Wk_R$. Hence it belongs to the stability subgroup (or Wigner little group) of k_R. For $k_R = (m, \mathbf{0})$ it is a rotation. Pressing on,

$$U(\Lambda)|p, \sigma\rangle = U[L(\Lambda p)]U[W(\Lambda, p)]|k_R, \sigma\rangle, \tag{22}$$

and as a result,

$$U(\Lambda)|p, \sigma\rangle = \sum_\xi D_{\xi\sigma}[W(\Lambda, p)]|\Lambda p, \xi\rangle, \tag{23}$$

where $D_{\xi\sigma}$ are the matrix elements of the representation of the Wigner rotation $W(\Lambda, p)$.

We consider only spin-$\frac{1}{2}$ particles, so $\sigma = \pm\frac{1}{2}$. Any 2×2 unitary matrix can be written as $\hat{D} = \exp(-\mathrm{i}\omega\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})$, where ω is a rotation angle and $\hat{\mathbf{n}}$ is a rotation axis that corresponds to $W(\Lambda, p)$.

The wave functions transform according to $\psi'_\xi(q) = \langle \xi, q|U(\Lambda)|\Psi\rangle$ so the same state in the Lorentz-transformed frame is

$$|\Psi'\rangle = U(\Lambda)|\Psi\rangle = \sum_{\sigma,\xi} \int_{-\infty}^{\infty} D_{\sigma\xi}[W(\Lambda, \Lambda^{-1}p)]\psi_\xi(\Lambda^{-1}p)|\sigma, p\rangle d\mu(p). \quad (24)$$

For pure rotation $\mathcal{R}$ the three-dimensional (more exactly, 3D block of 4D matrix; here and in the following we use the same letter for a 4D and 3D matrix for $\mathcal{R} \in$SO(3)) Wigner rotation matrix is the rotation itself,

$$W(\mathcal{R}, p) = \mathcal{R}, \qquad \forall p = (p^0, \mathbf{p}). \quad (25)$$

As a result, the action of Wigner spin operators on $\mathcal{H}_1$ is given by than the halves of Pauli matrices that are tensored with the identity of L^2.

2.2 Photons

The single-photon states are labeled by momentum $\mathbf{p}$ (the 4-momentum vector is null, $E = p^0 = |\mathbf{p}|$) and helicity $\sigma_{\mathbf{p}} = \pm 1$, so the state with a definite momentum is given by $\sum_{\sigma=\pm1} \alpha_\sigma |p, \sigma_{\mathbf{p}}\rangle$, where $|\alpha_+|^2 + |\alpha_-|^2 = 1$. Polarization states are also labeled by 3-vectors $\boldsymbol{\epsilon}^\sigma_{\mathbf{p}}$, $\mathbf{p} \cdot \boldsymbol{\epsilon}^\sigma_{\mathbf{p}} = 0$, that correspond to the two senses of polarization of classical electromagnetic waves. An alternative labeling of the same state, therefore, is $\sum_{\sigma=\pm1} \alpha_\sigma |p, \boldsymbol{\epsilon}^\sigma_{\mathbf{p}}\rangle$. Helicity is invariant under proper Lorentz transformation, but the basis states acquire phases.

The little group element $W(\Lambda, p) = L^{-1}(\Lambda p)\Lambda L(p)$ is defined with respect to the standard four-momentum $k_R = (1, 0, 0, 1)$. The standard Lorentz transformation is

$$L(p) = R(\hat{\mathbf{p}})B_z(u), \quad (26)$$

where $B_z(u)$ is a pure boost along the z-axis with a velocity u that takes k_R to the vector $(|\mathbf{p}|, 0, 0, |\mathbf{p}|)$ and $R(\hat{\mathbf{p}})$ is the standard rotation that carries the z-axis into the direction of the unit vector $\hat{\mathbf{p}}$. If $\hat{\mathbf{p}}$ has polar and azimuthal angles θ and ϕ, the standard rotation $R(\hat{\mathbf{p}})$ is accomplished by a rotation by θ around the y-axis, that is followed by a rotation by ϕ around the z-axis. Hence,

$$R(\hat{\mathbf{p}}) = \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \cos\phi\sin\theta \\ \cos\theta\sin\phi & \cos\phi & \sin\phi\sin\theta \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \quad (27)$$

(here only the non-trivial 3D block is shown).

An arbitrary little group element for a massless particle is decomposed according to

$$W(\Lambda, p) = S(\beta, \gamma)R_z(\xi), \tag{28}$$

where the elements $S(\beta, \gamma)$ form a subgroup that is isomorphic to the translations of the Euclidean plane and $R_z(\xi)$ is a rotation around the z-axis. We are interested only in the angle ξ, since β and γ do not correspond to the physical degrees of freedom. However, they are important for gauge transformations. Finally, the little group elements are represented by

$$D_{\sigma'\sigma} = \exp(i\xi\sigma)\delta_{\sigma'\sigma}. \tag{29}$$

It is worthwhile to derive more explicit expressions for ξ. I begin with rotations, $\Lambda = \mathcal{R}$. Since rotations form a subgroup of a Lorentz group, $R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})$ is a rotation that leaves $\hat{\mathbf{z}}$ invariant and thus is of the form $R_z(\omega)$ for some ω. A boost in (t, z) plane and a rotation around z-axis commute, $[R_z, B_z] = 0$, so

$$W(\mathcal{R}, p) = R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}}) = R_z(\xi). \tag{30}$$

Any rotation can be described by two angles that give a direction of the axis and the third angle that gives the amount of rotation around that axis. If $\mathcal{R}\mathbf{p} = \mathbf{q}$, we decompose the rotation matrix as

$$\mathcal{R} = R_{\hat{\mathbf{q}}}(\omega)R(\hat{\mathbf{q}})R^{-1}(\hat{\mathbf{p}}), \tag{31}$$

where $R_{\hat{\mathbf{q}}}(\omega)$ characterizes a rotation around $\hat{\mathbf{q}}$, and $R(\hat{\mathbf{q}})$ and $R(\hat{\mathbf{p}})$ are the standard rotations that carry the z-axis to $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$, respectively. Using (30) we find that $S = \mathbb{1}$ and the two rotations are of the same conjugacy class,

$$R_z(\xi) = R^{-1}(\mathcal{R}\hat{\mathbf{p}})R_{\mathcal{R}\hat{\mathbf{p}}}(\omega)R(\mathcal{R}\hat{\mathbf{p}}), \tag{32}$$

so we conclude that $\xi = \omega$.

A practical description of polarization states is given by spatial vectors that correspond to the classical polarization directions. Taking again k_R as the reference momentum, two basis vectors of linear polarization are $\boldsymbol{\epsilon}^1_{k_R} = (1, 0, 0)$ and $\boldsymbol{\epsilon}^2_{k_R} = (0, 1, 0)$, while to the right and left circular polarizations correspond $\boldsymbol{\epsilon}^{\pm}_{k_R} = (\boldsymbol{\epsilon}^1_{k_R} \pm \mathrm{i}\boldsymbol{\epsilon}^2_{k_R})/\sqrt{2}$.

Phases of the states obtained by the standard Lorentz transformations $L(p)$ are set to 1. Since the standard boost $B_z(u)$ leaves the four-vector $(0, \boldsymbol{\epsilon}^{\pm}_{k_R})$ invariant, we define a polarization basis for any $\mathbf{p}$ as

$$\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}} = \boldsymbol{\epsilon}^{\pm}_{\hat{\mathbf{p}}} \equiv R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^{\pm}_{k_R}, \tag{33}$$

while the transformation of polarization vectors under an arbitrary rotation $\mathcal{R}$ is given by the rotation itself. To see the agreement between transformations of spatial vectors and states, consider a generic state with a momentum $\mathbf{p}$. Its polarization is described by the polarization vector $\boldsymbol{\alpha}(\mathbf{p}) = \alpha_+ \boldsymbol{\epsilon}^+_{\mathbf{p}} + \alpha_- \boldsymbol{\epsilon}^-_{\mathbf{p}}$, or by the state vector $\alpha_+|\mathbf{p}, +\rangle + \alpha_-|\mathbf{p}, -\rangle$. Using (33) we see that the transformation of $\boldsymbol{\alpha}(\mathbf{p})$ is given by

$$\mathcal{R}\boldsymbol{\alpha}(\mathbf{p}) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega) R(\mathcal{R}\hat{\mathbf{p}}) R^{-1}(\hat{\mathbf{p}}) \boldsymbol{\alpha}(\mathbf{p}) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega) R(\mathcal{R}\hat{\mathbf{p}}) \boldsymbol{\alpha}(k_R) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega) \boldsymbol{\alpha}(\mathcal{R}\hat{\mathbf{p}}). \tag{34}$$

If $\mathbf{q} = \mathcal{R}\mathbf{p}$ the transformation results in $\alpha_+ e^{\mathrm{i}\omega} \boldsymbol{\epsilon}^+_{\mathbf{q}} + \alpha_- e^{-\mathrm{i}\omega} \boldsymbol{\epsilon}^-_{\mathbf{q}}$, and since $\omega = \xi$, it is equivalent to the state transformation

$$U(\mathcal{R})(\alpha_+|p, +\rangle + \alpha_-|p, -\rangle) = \alpha_+ e^{\mathrm{i}\xi}|q, +\rangle + \alpha_- e^{-\mathrm{i}\xi}|q, -\rangle. \tag{35}$$

For a general Lorentz transformations the triad $(\boldsymbol{\epsilon}^1_{\mathbf{p}}, \boldsymbol{\epsilon}^2_{\mathbf{p}}, \hat{\mathbf{p}})$ is rigidly rotated, but in a more complicated fashion. To obtain the phase for a general Lorentz transformation, we decompose the latter into two rotations and a standard boost B_z along the z-axis:

$$\Lambda = \mathcal{R}_2 B_z(u) \mathcal{R}_1. \tag{36}$$

It can be shown that B_z alone does not lead to a phase rotation. Therefore,

$$\xi = \omega_1 + \omega_2, \tag{37}$$

where both ω_1 and ω_2 are due to the rotations and are given by (31). Note that although $B_z(u)$ alone does not lead to a phase rotation, it can affect the value of ω_2, since it indirectly appears in the definition of $\mathcal{R}_2$.

Remarks

1. A comprehensive discussion of the Poincaré group in physics can be found in [21, 22]. Useful expressions for Wigner rotations and their applications for massive particles are given in [23–25].
2. In this transformation I do not assume any additional normalization factors. A condition of unitarity is $UU^\dagger = U^\dagger U = \mathbb{1}$, but there also other conventions in the literature.
3. A double infinity of the positive energy solutions of the Dirac equation (functions $u_p^{(1/2)}$ and $u_p^{(-1/2)}$) spans an improper basis of this space. There is a one-to-one correspondence between Wigner and Dirac wave functions. Basis vectors of Wigner and Dirac Hilbert spaces are in the one-to-one correspondence [22],

$$u_p^{(1/2)} \Leftrightarrow |\tfrac{1}{2}, p\rangle, \qquad u_p^{(-1/2)} \Leftrightarrow |\tfrac{1}{2}, p\rangle, \tag{38}$$

while the wave functions are related by

$$\Psi^{\alpha}(p) = \psi_{1/2}(p)u_p^{(1/2)\alpha} + \psi_{-1/2}(p)u_p^{(-1/2)\alpha} \tag{39}$$

$$2m\psi_{\sigma}(p) = \sum_{\alpha=1}^{4} \overline{u}_{\alpha p}^{(-\sigma)}\Psi^{\alpha}(p). \tag{40}$$

4. Another approach to the construction of the Wigner rotation $\hat{D}$ is based on the homomorphism between Lorentz group and $SL(2)$ [22].
5. When not restricted to a single-particle space the Wigner spin operator is given by

$$\mathbf{S} = \tfrac{1}{2}\sum_{\eta,\zeta} \boldsymbol{\sigma}_{\eta\zeta} \int d\mu(p)(\hat{a}_{\eta p}^{\dagger}\hat{a}_{\zeta p} + \hat{b}_{\eta p}^{\dagger}\hat{b}_{\sigma p}), \tag{41}$$

   where $\hat{a}_{\eta p}^{\dagger}$ creates a mode with a momentum p and spin η along the z-axis, etc. A comparison of different spin operators can be found in [26].
6. If one works with the 4-vectors, then in the helicity gauge the polarization vector is given by $\epsilon_p = (0, \boldsymbol{\epsilon}_{\mathbf{p}})$. A formal connection between helicity states and polarization vectors is made by first observing that three spin-1 basis states can be constructed from the components of a symmetric spinor of rank 2. Unitary transformations of this spinor that are induced by $\mathcal{R}$ are in one-to-one correspondence with transformations by $\mathcal{R}$ of certain linear combinations of a spatial vector. In particular, transformations of the helicity ± 1 states induced by rotations are equivalent to the rotations of $(\boldsymbol{\epsilon}_{ks}^{1} \pm \mathrm{i}\boldsymbol{\epsilon}_{ks}^{2})/\sqrt{2}$ (the z-axis is the initial quantization direction). While $p_{\mu}\epsilon_p^{\mu} = 0$ gauge condition is Lorentz-invariant, the spatial orthogonality is not. The role of gauge transformations in preserving the helicity gauge and some useful expressions for the phase that photons acquire can be found in [27–29].

3 Implications of Quantum Lorentz Transformations

3.1 Reduced Density Matrices

In a relativistic system whatever is outside the past light cone of the observer is unknown to him, but also cannot affect his system, therefore does not lead to decoherence (here, I assume that no particle emitted by from the outside the past cone penetrates into the future cone). Since different observers have different past light cones, by tracing out they exclude from their descriptions different parts of spacetime. Therefore any transformation law between them must tacitly assume that the part excluded by one observer is irrelevant to the system of another.

Another consequence of relativity is that there is a hierarchy of dynamical variables: *primary variables* have relativistic transformation laws that depend only on the Lorentz transformation matrix Λ that acts on the spacetime coordinates. For example, momentum components are primary variables. On the other hand,

secondary variables such as spin and polarization have transformation laws that depend not only on Λ, but also on the momentum of the particle. As a consequence, the reduced density matrix for secondary variables, which may be well defined in any coordinate system, has no transformation law relating its values in different Lorentz frames.

Moreover, an unambiguous definition of the reduced density matrix is possible only if the secondary degrees of freedom are unconstrained, and photons are the simplest example when this definition fails. In the absence of a general prescription, a case-by-case treatment is required. I describe a particular construction, valid with respect to a certain class of tests.

3.2 Massive Particles

For a massive qubit the usual definition of quantum entropy has no invariant meaning. The reason is that under a Lorentz boost, the spin undergoes a Wigner rotation, that as shown on Fig. 1 is controlled both by the classical data and the corresponding momentum. Even if the initial state is a direct product of a function of momentum and a function of spin, the transformed state is not a direct product. Spin and momentum become entangled.

Let us define a reduced density matrix,

$$\rho = \int d\mu(p)\psi(p)\psi^{\dagger}(p). \tag{42}$$

It gives statistical predictions for the results of measurements of spin components by an ideal apparatus which is not affected by the momentum of the particle. Note that I tacitly assumed that the relevant observable is the Wigner spin. The spin entropy is

$$S = -\operatorname{Tr}(\rho \log \rho) = -\sum_j \lambda_j \log \lambda_j, \tag{43}$$

where λ_j are the eigenvalues of ρ.

As usual, ignoring some degrees of freedom leaves the others in a mixed state. What is not obvious is that in the present case the amount of mixing depends on the Lorentz frame used by the observer. Indeed consider another observer (Bob) who moves with a constant velocity with respect to Alice who prepared that state. In the Lorentz frame where Bob is at rest, the state is given by (24).

As an example, take a particle prepared by Alice to be

$$|\Psi\rangle = \chi \int \psi(p)|p\rangle d\mu(p), \qquad \chi = \begin{pmatrix} \zeta \\ \eta \end{pmatrix}, \tag{44}$$

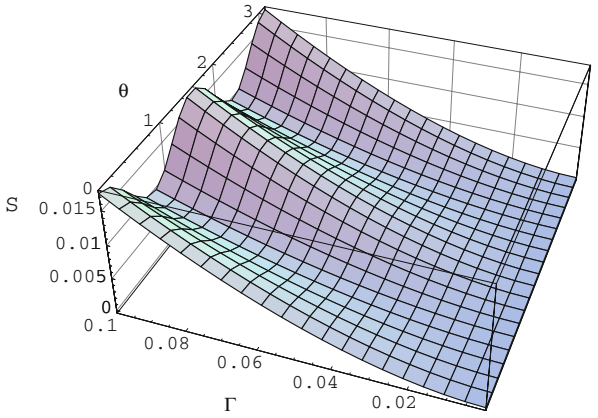


Fig. 2 Dependence of the spin entropy S, in Bob's frame, on the values of the angle θ and a parameter $\Gamma = [1 - (1 - v^2)^{1/2}]\Delta/mv$

where ψ is concentrated near zero momentum and has a characteristic spread Δ. Spin density matrices of all the states that are given by (44) are

$$\rho = \begin{pmatrix} |\zeta|^2 & \zeta\eta^* \\ \zeta^*\eta & |\eta|^2 \end{pmatrix}, \tag{45}$$

and are independent of the specific form of $\psi(p)$. To make calculations explicit (and simpler) I take the wave function to be Gaussian, $\psi(p) = N \exp(-\mathbf{p}^2/2\Delta^2)$, where N is a normalization factor. Spin and momentum are not entangled, and the spin entropy is zero. When that particle is described in Bob's Lorentz frame, moving with velocity v at the angle θ with Alice's z-axis, a detailed calculation shows that the the spin entropy is positive [1]. This phenomenon is illustrated in Fig. 2. A relevant parameter, apart from the angle θ, is in the leading order in momentum spread:

$$\Gamma = \frac{\Delta}{m} \frac{1 - \sqrt{1 - v^2}}{v}, \tag{46}$$

where Δ is the momentum spread in Alice's frame. The entropy has no invariant meaning, because the reduced density matrix τ has no covariant transformation law, except in the limiting case of sharp momenta. Only the complete density matrix transforms covariantly.

I outline some of the steps in this derivation. First, we calculate the rotation parameters $(\omega, \hat{\mathbf{n}})$ of the orthogonal matrix $W(\Lambda, p)$ for a general momentum. The rotation axis and angle are given by

$$\hat{\mathbf{n}} = \hat{\mathbf{v}} \times \hat{\mathbf{p}}, \qquad \cos\theta = \hat{\mathbf{v}} \cdot \hat{\mathbf{p}}, \qquad 0 \le \theta \le \pi\,, \tag{47}$$

where $\hat{\mathbf{v}}$ is boost's direction, while the leading order term for the angle is

$$\omega = \frac{1-\sqrt{1-v^2}}{v}\frac{p}{m}\sin\theta - \mathcal{O}\left(\frac{p^2}{m^2}\right). \tag{48}$$

Without a loss of generality we can make another simplification. We can choose our coordinate frame in such a way that both ζ and η are real. The matrix $D[W(\Lambda, \Lambda^{-1}p)]$ takes the form

$$D[W(\Lambda, p')] = \sigma_0 \cos\frac{\omega}{2} - \mathrm{i}\sin\frac{\omega}{2}(-\sin\phi\,\sigma_x + \cos\phi\,\sigma_y), \tag{49}$$

where (θ, ϕ) are the spherical angle of $\mathbf{p}'$ (to be consistent with Eq. (24) momentum in Alice frame carries a prime, $p' = \Lambda^{-1}p$). The reduced density matrix in Bob's frame is

$$\rho^B_{\sigma\xi} = \int d\mu(p) D_{\sigma\nu} D^*_{\xi\lambda} \psi_\nu(p')\psi^*_\lambda(p'). \tag{50}$$

The symmetry of $\psi(\Lambda^{-1}p)$ is cylindrical. Hence the partial trace is taken by performing a momentum integration in cylindrical coordinates. This simplification is a result of the spherical symmetry of the original ψ. The two remaining integrations are performed by first expanding in powers of p/Δ and taking Gaussian integrals. Finally,

$$\rho' = \begin{pmatrix} \zeta^2(1-\Gamma^2/4) + \eta^2\Gamma^2/4 & \zeta\eta^*(1-\Gamma^2/4) \\ \zeta^*\eta(1-\Gamma^2/4) & \zeta^2\Gamma^2/4 + \eta^2(1-\Gamma^2/4) \end{pmatrix}. \tag{51}$$

Fidelity can be used to estimate the difference between the two density matrices. It is defined as

$$f = \chi^\dagger \rho' \chi, \tag{52}$$

and it is easy to get an analytical result for this quantity. Set $\zeta = \cos\theta$ and $\eta = \sin\theta$. Then

$$f = 1 - \frac{\Gamma^2}{2}\left(3 + \frac{\cos 4\theta}{8}\right). \tag{53}$$

Consider now a pair of orthogonal states that were prepared by Alice, e.g., the above state with $\chi_1 = (1, 0)$ and $\chi_2 = (0, 1)$. How well can moving Bob distinguish them? I use the simplest criterion, namely the probability of error P_E, defined as follows: an observer receives a single copy of one of the two known states and performs any operation permitted by quantum theory in order to decide which state was supplied. The probability of a wrong answer for an optimal measurement is

$$P_E(\rho_1, \rho_2) = \tfrac{1}{2} - \frac{1}{4}\,\mathrm{tr}\sqrt{(\rho_1 - \rho_2)^2}. \tag{54}$$

In Alice's frame $P_E = 0$. In Bob's frame the reduced density matrices are

$$\rho_1^B = \begin{pmatrix} 1 - \Gamma^2/4 & 0 \\ 0 & \Gamma^2/4 \end{pmatrix}, \qquad \rho_2^B = \begin{pmatrix} \Gamma^2/4 & 0 \\ 0 & 1 - \Gamma^2/4 \end{pmatrix}, \tag{55}$$

respectively. Hence the probability of error is $P_E(\rho_1, \rho_2) = \Gamma^2/4$.

3.3 Photons

The relativistic effects in photons are essentially different from those for massive particles that were discussed above. This is because photons have only two linearly independent polarization states. As we know, polarization is a *secondary variable*: states that correspond to different momenta belong to distinct Hilbert spaces and cannot be superposed (an expression such as $|\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle + |\boldsymbol{\epsilon}_{\mathbf{q}}^{\pm}\rangle$ is meaningless if $\mathbf{k} \neq \mathbf{q}$). The complete basis $|p, \boldsymbol{\epsilon}_{\mathbf{p}}^{\pm}\rangle$ does not violate this superselection rule, owing to the orthogonality of the momentum basis. The reduced density matrix, according to the usual rules, should be

$$\rho = \int d\mu(p)|\psi(p)|^2 |p, \boldsymbol{\alpha}(\mathbf{p})\rangle\langle p, \boldsymbol{\alpha}(\mathbf{p})|. \tag{56}$$

However, since ξ in Eq. (29) depends on the photon's momentum even for ordinary rotations, this object will have no transformation law at all. It is still possible to define an "effective" density matrix adapted to a specific method of measuring polarization [30, 31]. I describe one such scheme.

The labeling of polarization states by Euclidean vectors $\boldsymbol{\epsilon}_{\mathbf{p}}^{\pm}$ suggests the use of a 3×3 matrix with entries labeled x, y and z. Classically, they correspond to different directions of the electric field. For example, a component ρ_{xx} would give the expectation values of operators representing the polarization in the x direction, seemingly irrespective of the particle's momentum.

To have a momentum-independent polarization is to admit longitudinal photons. Momentum-independent polarization states thus consist of physical (transverse) and unphysical (longitudinal) parts, the latter corresponding to a polarization vector $\boldsymbol{\epsilon}^{\ell} = \hat{\mathbf{p}}$. For example, a generalized polarization state along the x-axis is

$$|\hat{\mathbf{x}}\rangle = x_+(\mathbf{p})|\boldsymbol{\epsilon}_{\mathbf{p}}^{+}\rangle + x_-(\mathbf{p})|\boldsymbol{\epsilon}_{\mathbf{p}}^{-}\rangle + x_\ell(\mathbf{p})|\boldsymbol{\epsilon}_{\mathbf{p}}^{\ell}\rangle, \tag{57}$$

where $x_\pm(\mathbf{p}) = \hat{\mathbf{x}} \cdot \boldsymbol{\epsilon}_{\mathbf{p}}^{\pm}$ and $x_\ell(\mathbf{p}) = \hat{\mathbf{x}} \cdot \hat{\mathbf{p}} = \sin\theta \cos\phi$. It follows that $|x_+|^2 + |x_-|^2 + |x_\ell|^2 = 1$, and we thus define

$$\mathbf{e}_x(\mathbf{p}) = \frac{x_+(\mathbf{p})\boldsymbol{\epsilon}^+_{\mathbf{p}} + x_-(\mathbf{p})\boldsymbol{\epsilon}^-_{\mathbf{p}}}{\sqrt{x_+^2 + x_-^2}}, \tag{58}$$

as the polarization vector associated with the x direction. It follows from (57) that $\langle\hat{\mathbf{x}}|\hat{\mathbf{x}}\rangle = 1$ and $\langle\hat{\mathbf{x}}|\hat{\mathbf{y}}\rangle = \hat{\mathbf{x}}\cdot\hat{\mathbf{y}} = 0$, and likewise for the other directions, so that

$$|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| + |\hat{\mathbf{y}}\rangle\langle\hat{\mathbf{y}}| + |\hat{\mathbf{z}}\rangle\langle\hat{\mathbf{z}}| = \mathbb{1}_p, \tag{59}$$

where $\mathbb{1}_p$ is the unit operator in momentum space.

To the direction $\hat{\mathbf{x}}$ there corresponds a projection operator

$$P_{xx} = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \mathbb{1}_p = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \int d\mu(k)|\mathbf{p}\rangle\langle\mathbf{p}|\,. \tag{60}$$

The action of P_{xx} on $|\Psi\rangle$ follows from (57) and $\langle\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}|\boldsymbol{\epsilon}^{\ell}_{\mathbf{p}}\rangle = 0$. Only the transverse part of $|\hat{\mathbf{x}}\rangle$ appears in the expectation value:

$$\langle\Psi|P_{xx}|\Psi\rangle = \int d\mu(p)|\psi(p)|^2|x_+(\mathbf{p})\alpha^*_+(\mathbf{p}) + x_-(\mathbf{p})\alpha^*_-(\mathbf{p})|^2. \tag{61}$$

It is convenient to write the transverse part of $|\hat{\mathbf{x}}\rangle$ as

$$|\mathbf{b}_x(\mathbf{p})\rangle \equiv (|\boldsymbol{\epsilon}^+_{\mathbf{p}}\rangle\langle\boldsymbol{\epsilon}^+_{\mathbf{p}}| + |\boldsymbol{\epsilon}^-_{\mathbf{p}}\rangle\langle\boldsymbol{\epsilon}^-_{\mathbf{p}}|)|\hat{\mathbf{x}}\rangle = x_+(\mathbf{p})|\boldsymbol{\epsilon}^+_{\mathbf{p}}\rangle + x_-(\mathbf{p})|\boldsymbol{\epsilon}^-_{\mathbf{p}}\rangle. \tag{62}$$

Likewise define $|\mathbf{b}_y(\mathbf{p})\rangle$ and $|\mathbf{b}_z(\mathbf{p})\rangle$. These three state vectors are neither of unit length nor mutually orthogonal.

Finally, a POVM element E_{xx} which is the physical part of P_{xx}, namely is equivalent to P_{xx} for physical states (without longitudinal photons) is

$$E_{xx} = \int d\mu(k)|p, \mathbf{b}_x(\mathbf{p})\rangle\langle p, \mathbf{b}_x(\mathbf{p})|, \tag{63}$$

and likewise for the other directions. The operators E_{xx}, E_{yy} and E_{zz} indeed form a POVM in the space of physical states, owing to (59).

To complete the construction of the density matrix, we introduce additional directions. Following the standard practice of state reconstruction, we consider $P_{x+z,x+z}$, $P_{x+\mathrm{i}z,x+\mathrm{i}z}$ and similar combinations. For example,

$$P_{x+z,x+z} = \tfrac{1}{2}(|\hat{\mathbf{x}}\rangle + |\hat{\mathbf{z}}\rangle)(\langle\hat{\mathbf{x}}| + \langle\hat{\mathbf{z}}|) \otimes \mathbb{1}_p. \tag{64}$$

The diagonal elements of the new polarization density matrix are defined as

$$\rho_{mm} = \langle\Psi|E_{mm}|\Psi\rangle, \qquad m = x, y, z, \tag{65}$$

and the off-diagonal elements are recovered by combinations such as

$$\rho_{xz} = \langle\Psi(|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{z}}| \otimes \mathbb{1}_p)|\Psi\rangle = \langle\Psi|E_{x+z,x+z} - \mathrm{i}E_{x-\mathrm{i}z,x-\mathrm{i}z} + (1-\mathrm{i})(E_{xx} - E_{zz})/2|\Psi\rangle. \tag{66}$$

Denote $|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{z}}| \otimes \mathbb{1}_p$ as P_{xz}, and its "physical" part by E_{xz}. Then the effective polarization density matrix is $(m, n, = x, y, z)$:

$$\rho_{mn} = \langle\Psi|E_{mn}|\Psi\rangle = \int d\mu(k)|f(\mathbf{p})|^2\langle\boldsymbol{\alpha}(\mathbf{p})|\mathbf{b}_m(\mathbf{p})\rangle\langle\mathbf{b}_n(\mathbf{p})|\boldsymbol{\alpha}(\mathbf{p})\rangle. \tag{67}$$

It is interesting to note that this derivation gives a direct physical meaning to the naive definition of a reduced density matrix,

$$\rho_{mn}^{\text{naive}} = \int d\mu(p)|\phi(p)|^2\boldsymbol{\alpha}_m(\mathbf{p})\boldsymbol{\alpha}_n^*(\mathbf{p}) = \rho_{mn}\,. \tag{68}$$

It is possible to show that this POVM actually corresponds to a simple photodetection model [32].

The basis states $|p, \boldsymbol{\epsilon}_\mathbf{p}\rangle$ are direct products of momentum and polarization. Owing to the transversality requirement $\boldsymbol{\epsilon}_\mathbf{p}\cdot\mathbf{p} = 0$, they remain direct products under Lorentz transformations. All the other states have their polarization and momentum degrees of freedom entangled. As a result, if one is restricted to polarization measurements as described by the above POVM, *there do not exist two orthogonal polarization states*. In general, any measurement procedure with finite momentum sensitivity will lead to the errors in identification, as demonstrated as follows.

Let two states $|\Phi\rangle$ and $|\Psi\rangle$ be two orthogonal single-photon states. Their reduced polarization density matrices, ρ_Φ and ρ_Ψ, respectively, are calculated using (67). Since the states are entangled, the von Neumann entropies of the reduced density matrices, $S = -\mathrm{tr}(\rho\ln\rho)$, are positive. Therefore, both matrices are at least of rank two. Since the overall dimension is 3, it follows that $\mathrm{tr}(\rho_\Phi\rho_\Psi) > 0$ and these states are not perfectly distinguishable. An immediate corollary is that photon polarization states cannot be cloned perfectly, because the no-cloning theorem forbids exact copying of unknown non-orthogonal states.

In general, any measurement procedure with finite-momentum sensitivity will lead to the errors in identification. First I present some general considerations and then illustrate them with a simple example. Let us take the z-axis to coincide with the average direction of propagation so that the mean photon momentum is $k_A\hat{\mathbf{z}}$. Typically, the spread in momentum is small, but not necessarily equal in all directions. Usually the intensity profile of laser beams has cylindrical symmetry, and we may assume that $\Delta_x \sim \Delta_y \sim \Delta_r$ where the index r means radial. We may also assume that $\Delta_r \gg \Delta_z$. We then have

$$f(p) \propto f_1[(p_z - k_A)/\Delta_z]\, f_2(p_r/\Delta_r). \tag{69}$$

We approximate

$$\theta \approx \tan\theta \equiv p_r/p_z \approx p_r/k_A. \tag{70}$$

In pictorial language, polarization planes for different momenta are tilted by angles up to $\sim \Delta_r/k_A$, so that we expect an error probability of the order Δ_r^2/k_A^2. In the density matrix ρ_{mn} all the elements of the form ρ_{mz} should vanish when $\Delta_r \to 0$. Moreover, if $\Delta_z \to 0$, the non-vanishing xy block goes to the usual (monochromatic) polarization density matrix,

$$\rho_{\text{pure}} = \begin{pmatrix} |\alpha|^2 & \beta & 0 \\ \beta^* & 1-|\alpha|^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{71}$$

As an example, consider two states which, if the momentum spread could be ignored, would be $|k_A\hat{\mathbf{z}}, \boldsymbol{\epsilon}^{\pm}_{k_A\hat{\mathbf{z}}}\rangle$. To simplify the calculations we assume a Gaussian distribution:

$$f(p) = Ne^{-(p_z-k_A)^2/2\Delta_z^2} e^{-p_r^2/2\Delta_r^2}, \tag{72}$$

where N is a normalization factor and $\Delta_z, \Delta_r \ll k_A$. In general the spread in p_z may introduce an additional incoherence into density matrices, in addition to the effect caused by the transversal spread. However, when all momentum components carry the same helicity, this spread results in corrections of the higher order. In the example below we take the polarization components to be $\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}} \equiv R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^{\pm}_{k_S}$. That means we have to analyze the states

$$|\Psi_{\pm}\rangle = \int d\mu(p) f(p) |p, \boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}\rangle, \tag{73}$$

where $f(p)$ is given above.

It is enough to expand $R(\hat{\mathbf{p}})$ up to second order in θ. The reduced density matrices are calculated by techniques similar to those for massive particles, using rotational symmetry around the z-axis and normalization requirements. At the leading order in $\Omega \equiv \Delta_r/k_A$

$$\rho_+ = \tfrac{1}{2}(1-\tfrac{1}{2}\Omega^2) \begin{pmatrix} 1 & -\mathrm{i} & 0 \\ \mathrm{i} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \tfrac{1}{2}\Omega^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{74}$$

and $\rho_- = \rho_+^*$. At the same level of precision,

$$P_E(\rho_+, \rho_-) = \Delta_r^2/4k_A^2. \tag{75}$$

It is interesting to note that the optimal strategy for distinguishing between these two states is a polarization measurement in the xy-plane. Then the effective 2×2 density matrices are perfectly distinguishable, but there is a probability $\Omega^2/2$ that no photon will be detected at all. The above result is valid due to the special form of the states that we had chosen. Potential errors in the upper 2×2 blocks are

averaged out in the integration over ϕ. The effect becomes important when, e.g., a plane monochromatic wave undergoes a strong focusing. Then $\Omega \approx l/f$, where l is the aperture radius and f is the focal length [32].

Now let us turn to the distinguishability problem from the point of view of a moving observer, Bob. The probability of an error is still given by (54), but the parameters depend on the observer's motion. Assume again that Bob moves along the z-axis with a velocity v. To calculate Bob's reduced density matrix, we must transform the complete state, and then take a partial trace.

Reduced density matrices of $|\Psi_\pm\rangle$ in both frames are given by the expression

$$(\rho_\pm)^{A,B}_{mn} = \int d\mu(p) |f(p)^{A,B}|^2 \langle R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^\pm_{k_S} | \mathbf{b}_m(\mathbf{p})\rangle \langle \mathbf{b}_n(\mathbf{p}) | R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^\pm_{k_S}\rangle. \tag{76}$$

This is due to the following two reasons. First, $|\mathbf{b}_m(\mathbf{p})\rangle$ are defined by (62) in any frame, while pure boosts preserve the orientation of the coordinate axes in 3-space, and therefore do not affect the indices of ρ_{mn}. Second, phases acquired by polarization states cancel out, since we choose the states $|\mathbf{p}, \boldsymbol{\alpha}(\mathbf{p})\rangle$ to be the helicity eigenstates.

Calculation of Bob's density matrix is similar to the previous cases. The only frame dependent expression in (76) is $f^B(p) = f^A(\Lambda^{-1}p)$. A boost along the z-axis preserves k_r and ϕ. On the other hand,

$$k^B_z \approx k_A \sqrt{\frac{1-v}{1+v}}. \tag{77}$$

Since everything else in the integral remains the same, the effect of relative motion is given by a substitution

$$\Omega^B = \sqrt{\frac{1+v}{1-v}}\,\Omega^A, \qquad P^B_E = \frac{1+v}{1-v}\,P^A_E, \tag{78}$$

so Bob can distinguish the signals either better or worse than Alice [30].

Remarks

1. A modification of the spin operator [33] will allow for a momentum-independent transformation of the spin density matrix between two frames that are related by a fixed Lorentz transformation Λ_{12}. Its relation to our scheme is discussed in [34].
2. An additional motivation for introduction of effective polarization density matrices comes from the analysis of one-photon scattering [31].
3. While the effects of the relative motion of Alice and Bob clearly belong to the realm of a science fiction, the order-of-magnitude estimates for the effect of a strong focusing [32] should be experimentally observable.

4. I have discussed only discrete variables. To explore the relativistic effects with continuous variables [35] it is convenient to express the quantum Lorentz transformations in terms of mode creation and annihilation operators [36].

4 Communication Channels

What happens when Alice and Bob, who are in relative motion, try to communicate? Assume that they use qubits that were described above. Under a general Lorentz transformation Λ that relates Alice's and Bob's frames, the state of this qubit will be transformed due to three distinct effects, which are:

(i) A Wigner rotation due to the Lorentz boost Λ, which occurs even for momentum eigenstates. If Λ is known, then to the extent that the wave-packet spread can be ignored, this is inconsequential.

(ii) A decoherence due to the entangling of spin and momentum under the Lorentz transformation Λ because the particle is not in a momentum eigenstate. Although reduced or effective density matrices have no general transformation rule, such rules can be established for particular classes of experimental procedures. We can then ask how these effective transformation rules, $\rho' = T(\rho)$, fit into the framework of general state transformations. For example, for the massive qubit of Sect. 4.2 the effective transformation is given by

$$\rho' = \rho(1 - \frac{\Gamma^2}{4}) + (\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y)\frac{\Gamma^2}{8}. \tag{79}$$

If Λ is known and it is possible to implement the operators that were mentioned in the Remark 1 above, then this effect is absent. Otherwise, this noise is unavoidable. Still, it is worth to keep in mind that the motion can improve the message fidelity, as in (78).

(iii) Another kind of decoherence arises due to Bob's lack of knowledge about the transformation relating his reference frame to Alice's frame. Using the techniques of the decoherence-free subspaces, it is possible to eliminate this effect completely. For example, for massive particles four physical qubits may be used to encode a logical qubit, while for photons $2 \to 1$ encoding is sufficient. In both cases using the block encoding it is possible to reach an asymptotically unit efficiency [37].

Entanglement between the "qubit" and spatial degrees of freedom leads to an interesting complication of the analysis. It is known [1] that the dynamics of a subsystem may be not completely positive; there is a prior entanglement with another system and the dynamics is not factorizable. Since in (78) and in the discussion following (54) we have seen that distinguishability *can* be improved, we conclude that these transformations are *not completely positive*. The reason is that the Lorentz transformation acts not only on the "interesting" discrete variables, but also on the

primary momentum variables that we elected to ignore and to trace out, and its action on the interesting degrees of freedom depends on the "hidden" primary ones. Of course, the complete state, with all the variables, transforms unitarily and distinguishability is preserved.

This technicality has one important consequence. In quantum information theory quantum channels are described by completely positive maps that act on qubit states. Qubits themselves are realized as discrete degrees of freedom of various particles. If relativistic motion is important, then not only does the vacuum behave as a noisy quantum channel, but the very representation of a channel by a CP map fails.

5 Entanglement and Different Lorentz Observers

In this section I consider only two-particle states. Even in this simple setting there are several possible answers to the question what happens to the entanglement, depending on the details of the question. Since the quantum Lorentz transformation is given by a tensor product $U_1(\Lambda) \otimes U_2(\Lambda)$, the overall entanglement between the states is Lorentz-invariant.

Let us assume that the states can be approximated by momentum eigenstates. Then, the same conclusion applies to the spin–spin (or polarization–polarization) entanglement between the particles, and it is possible to write an appropriate entanglement measures that capture the effects of particle statistics and Lorentz-invariance of the entanglement [24]. However, it does not mean that this invariance will be observed in an experiment or that the violation of Bell-type inequalities that is observed in an experiment performed in Alice's frame will be observed if the same equipment is placed in Bob's frame.

While a field-theoretical analysis shows that violations of Bell-type inequalities are generic, there are conditions that are imposed on the experimental procedures that are used to detect them. Consider the CHSH inequality. For any two spacelike separated regions and any pairs of of operators, a, b, there is a state ρ such that the CHSH inequality is violated, i.e., $\zeta(a, b, \rho) > 1$. With additional technical assumptions the existence of a maximally violating state ρ_m can be proved:

$$\zeta(a, b, \rho_m) = \sqrt{2}, \tag{80}$$

for any spacelike separated regions $\mathcal{O}_L$ and $\mathcal{O}_R$. It follows from convexity arguments that states that maximally violate Bell inequalities are pure. What are then the operators that lead to the maximal violation? It was shown [38] that the operators A_j and B_k that give $\zeta = \sqrt{2}$ satisfy $A_j^2 = \mathbb{1}$ and $A_1 A_2 + A_2 A_1 = 0$, and likewise for B_k. If we define $A_3 := -\mathrm{i}[A_1, A_2]/2$, then these three operators have the same algebra as Pauli spin matrices.

In principle the vacuum state may lead to the maximal violation of Bell-type inequalities. Their observability was discussed in [40].

The operators [39]

$$A_i = 2\left[\frac{m}{p^0}\mathbf{a}_i + \left(1 - \frac{m}{p^0}\right)(\mathbf{a}\cdot\mathbf{n})\mathbf{n}\right]\cdot\mathbf{S} \equiv 2\boldsymbol{\alpha}(\mathbf{a},\mathbf{p})\cdot\mathbf{S}, \tag{81}$$

where $\mathbf{S}$ is the Wigner spin operator and $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$ appear quite naturally as the candidates for the measurement description. The length of the auxiliary vector $\boldsymbol{\alpha}$ is

$$|\boldsymbol{\alpha}| = \frac{\sqrt{(\mathbf{p}\cdot\mathbf{a})^2 + m^2}}{p^0}, \tag{82}$$

so generically $A_i^2 = \boldsymbol{\alpha}^2\mathbb{1} < \mathbb{1}$, and indeed, the degree of violation decreases with the velocity of the observer. Nevertheless, it is always possible to compensate for a Wigner rotation by an appropriate choice of the operators [1].

Realistic situations involve wave packets. For example, a general spin-$\frac{1}{2}$ two-particle state may be written as

$$|\Upsilon_{12}\rangle = \sum_{\sigma_1,\sigma_2}\int d\mu(p_1)d\mu(p_2)g(\sigma_1\sigma_2,\mathbf{p}_1,\mathbf{p}_2)|\mathbf{p}_1,\sigma_1;\mathbf{p}_2,\sigma_2\rangle. \tag{83}$$

For particles with well-defined momenta, g sharply peaks at some values $\mathbf{p}_{10}$, $\mathbf{p}_{20}$. Again, a boost to any Lorentz frame S' will result in a unitary $U(\Lambda)\otimes U(\Lambda)$, acting on each particle separately, thus preserving the entanglement. Nevertheless, since they can change entanglement between different degrees of freedom of a single particle, the spin-spin entanglement is frame-dependent as well. Having investigated the reduced density matrix for $|\Upsilon_{12}\rangle$ and made explicit calculations for the case where g is a Gaussian, as in the Sect. 4.2 above, it is possible to show that if two particles are maximally entangled in a common (approximate) rest frame (Alice's frame), then the concurrence, as seen by a Lorentz-boosted Bob, decreases when $v \to 1$. Of course, the inverse transformation from Bob to Alice will increase the concurrence [41]. Thus, we see that the spin–spin entanglement is not a Lorentz invariant quantity, exactly as spin entropy is not a Lorentz scalar. Relativistic properties of the polarization entanglement are even more interesting [29], since there is no frame where polarization and momentum are unentangled.

Remarks

The question if there is entanglement and what happens to it crucially depends on how we define our subsystems. Consider, e.g., the inequivalent ways to identify two qubits in a four-state system. One choice may lead to an entangled state, while another may give zero entanglement. In addition, in our relativistic problem the choice of subsystems is influenced by the motion of the observer: definitions of spin involve momenta, and so the state of spin depends on the motion.

6 Beyond Special Relativity

6.1 Information in Relativistic Physics

Because of the lack of space I am only going to mention the various fascinating areas of the interplay between quantum information theory and relativistic physics. Quantum field theory provides us with new situations that should be investigated. For example, it is possible to ask all the usual questions about entanglement, distillability, etc., and their invariance[1, 42–44]. So far we discussed only observers that move with constant velocity. An accelerated observer sees Unruh radiation. It leads to a host of interesting effects if we consider a teleportation between a stationary and accelerated observer [45–47]. Dynamical entanglement—the one appearing in the scattering processes or between the decay products—also has been investigated [25, 48, 49].

Going to more exotic settings, I just mention that black hole physics [1, 50–53] and cosmology [54–56], either from the point of view of loop quantum gravity or string theory, provide extremely interesting scenarios where the questions of information can and should be asked.

As a teaser, I would like to show how very simple considerations (only slightly more sophisticated than a single CNOT gate) help to select possible candidates for quantum causal histories [57].

6.2 CNOT and Quantum Causal Histories

The quest for a quantum theory of gravity produced a variety of approaches that include string theory, loop quantum gravity, spin foams, causal sets, and causal dynamical triangulations. The successful theory has to provide a coherent structure that accommodates both classical relativity and quantum mechanics. In some appropriate limit it should produce the familiar physical phenomena on the flat spacetime background and make predictions on the kind and magnitude of the departures from this picture.

Those goals have not yet been achieved by any of the approaches, but these attempts have brought many insights and led to a better understanding of the problem's complexity. Quantum causal histories (QCHs) approach [58] to the quantization of gravity is a background-independent formalism that satisfies many of the conditions that are argued for by the above models. The idea is to use a causal set to describe the casual structure while a quantum theory being introduced through the assignment of finite-dimensional Hilbert spaces to the elementary events. Originally motivated by quantum cosmology, and providing a description of the causal spin foam models, QCHs make a direct contact with quantum information theory.

To apply the quantum-informational considerations I have to begin with a brief outline of QCHs in the Hilbert space language, roughly following [58]. If a spacetime is time-orientable and has no closed timelike curves, then its causal structure

can be completely described as a partial order relation on its points. The relation $x \preceq y$ is defined if there exists a future-directed non-spacelike curve from x to y. It is transitive, and the absence of CTCs means that $x \preceq y$ and $y \preceq x$ are simultaneously true if and only if $x = y$. Those two conditions make the relation "$\preceq$" into a partial order.

A discrete analogue of a smooth chronology-respecting spacetime is a causal set $\mathcal{C}$, which is a locally finite and partially ordered set. That is, for any two events $x, y \in \mathcal{C}$, there exist (at most) finitely many events $z \in \mathcal{C}$ such that $x \preceq z \preceq y$. If the events x and y are not related, i.e., neither $x \preceq y$ nor $y \preceq x$ holds, then they are spacelike separated, this fact being denoted as $x \sim y$. An acausal set is a subset $\xi \in \mathcal{C}$ such that all events in it are spacelike separated from one another. Then the maximal acausal sets are the discrete analogues of spacelike hypersurfaces.

A causal set can be represented by a directed graph of elementary relations, as in Fig. 3. Its vertices are the points of $\mathcal{C}$, while the edges $x \to y$ represent the elementary causal relations, namely $x \preceq y$ without any intermediate z such that $x \preceq z \preceq y$.

A future-directed path is a sequence of events such that there exists an edge from each event to the next. It is an analogue of a future-directed non-spacelike curve. A future-directed path is future (past) inextendible if there exists no event in $\mathcal{C}$ which is in the future (past) of the entire path. Then one can define complete future and complete past of an event. An acausal set ξ is a complete future of an event x if ξ intersects any future-inextendible future-directed path that starts at x, and a complete past is defined similarly. If an acausal set ζ is a complete future of an acausal set ξ and at the same time the set ξ is a complete past of ζ, then the sets form a complete pair, $\xi \preceq \zeta$. The resulting structure is a discrete, locally finite, causal pre-spacetime.

A local quantum structure on it is introduced by attaching a finite-dimensional Hilbert space $\mathcal{H}(x)$ to every event $x \in \mathcal{C}$. For only two spacelike separated events x and y the composite state space is $\mathcal{H}(x, y) = \mathcal{H}(x) \otimes \mathcal{H}(y)$, with an obvious generalization to larger sets. In ordinary quantum mechanics (of closed systems) time evolution is a unitary map of Hilbert spaces. In a QCH approach one introduces a unitary evolution between complete pairs of acausal sets ξ and ζ, where, e.g., ζ is the complete future of ξ, $\xi \preceq \zeta$. One can think of a complete pair as successive Cauchy surfaces of an isolated component of spacetime, or of all spacetime, with a unitary map U relating $\mathcal{H}(\xi)$ and $\mathcal{H}(\zeta)$.

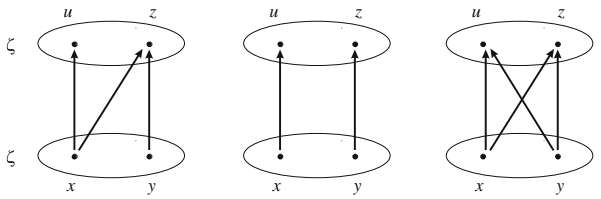


Fig. 3 Three possible causal histories

Hence a QCH consists of a causal set $\mathcal{C}$, a finite-dimensional Hilbert space $\mathcal{H}(x)$ at every $x \in \mathcal{C}$ and a unitary map

$$U(\xi, \zeta) : \mathcal{H}(\xi) \to \mathcal{H}(\zeta), \tag{84}$$

for any complete pair $\xi \preceq \zeta$. The maps have a natural composition property

$$U(\varsigma, \zeta)U(\xi, \varsigma) = U(\xi, \zeta), \qquad \text{for } \xi \preceq \varsigma \preceq \zeta. \tag{85}$$

Different possible causal relations between complete sets are shown in Fig. 3.

Note that the precise meaning of arrows in the relation $x \preceq y$ is as follows. For $x \in \xi$ and $y \in \zeta$ the existence of $U(\xi, \zeta)$ implies that if there is no future-directed path between x and y, i.e., $x \sim y$, then the reduced final state on $\mathcal{H}(y)$, $\rho_y^\zeta = \mathrm{tr}_{\zeta \setminus y}\rho^\zeta = \mathrm{tr}_{\zeta \setminus y} U\rho^\xi U^\dagger$ is independent of the initial reduced state ρ_x^ξ.

In the example I deal with here it is possible to identify the initial and final Hilbert spaces pointwise: in Fig. 3 we assume that $d_x \equiv \dim \mathcal{H}(x) = d_u \equiv \dim \mathcal{H}(u)$, etc. A general case is considered in [57].

Consider three possible causal relations that are shown in Fig. 3. Despite its intuitive appeal the causal history (a) is incompatible with the definition (84). It can be observed on a simple example of two qubits. Label the basis of each of the spaces by $|0\rangle, |1\rangle$. The CNOT gate [2] apparently fits the described scheme: the value of the qubit x remains the same, while the qubit y may be flipped. However, a textbook exercise shows that in the basis $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, the roles of the source and the target bits are reversed.

Table 1 CNOT gate

$\|\psi\rangle$	$U\|\psi\rangle$	$\|\psi\rangle$	$U\|\psi\rangle$
00	00	++	++
01	01	−+	−+
10	11	+−	−−
11	10	−−	+−

In general, if an operation on $\mathcal{H}(y)$ is controlled by the state on $\mathcal{H}(x)$ which remains unchanged,

$$U(|\psi\rangle \otimes |\phi\rangle) = |\psi\rangle \otimes V_\psi|\phi\rangle, \tag{86}$$

for all states $|\phi\rangle$, then it is actually independent of $|\psi\rangle$, $V_\psi \equiv V$. Indeed, consider two possible initial states $|\psi\rangle|\phi\rangle$ and $|\psi'\rangle|\phi\rangle$. Their overlap is preserved under U, hence

$$\langle\phi|V_\psi^\dagger V_{\psi'}|\phi\rangle = 1, \tag{87}$$

for an arbitrary state $|\phi\rangle$. Hence $V_\psi = V_{\psi'}$.

As a result, only the second and the third histories of Fig. 3 are consistent with the existence of an arbitrary unitary evolution on causal sets. The alternative is to introduce an external (classical) observer who is restricted to do measurements in a (given) particular basis, say (0,1) as in the above example. The CNOT example shows that the observer will prescribe different causal relations depending on its choice of measurement basis. Then, by restricting the allowed unitary evolutions between the complete sets, the asymmetric causal structure is made compatible with quantum mechanics.

Adhering to the latter option is not only too restrictive, but not always possible. Consider the projectors P_0 and P_1, $P_0 + P_1 = \mathbb{1}$, on the singlet (spin-0) and triplet (spin-1) states. Then for generic values of the parameters α, β there is no basis in which the unitary

$$U = e^{\mathrm{i}\alpha} P_0 + e^{\mathrm{i}\beta} P_1, \tag{88}$$

can be represented as in (86). This particular set of unitary operators is relevant to computational universe models and to loop quantum gravity.

More complicated causal structures can be analyzed in the same spirit. Moreover, it is possible to prove the following general theorem.

Theorem 1 *A discrete, locally finite, causal pre-spacetime structure admits a unitary evolution between its acausal surfaces if and only if it can be represented as a quantum computational network.*

References

1. A. Peres, D.R. Terno: Rev. Mod. Phys. **76**, 93 (2004) 221, 223, 225, 233, 240, 242, 243
2. M.A. Nielsen, I.L. Chuang: *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000) 221, 245
3. M. Keyl: Phys. Rep. **369**, 431 (2002) 221
4. P. Busch, M. Grabowski, P.J. Lahti: *Operational Quantum Physics* (Springer, Berlin, 1995) 221
5. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, e-print quant-ph/0702225, to be bublished in Rev. Mod. Phys. (2007) 221
6. H.-P. Breuer, F. Petruccione (eds.): *Relativistic Quantum Measurement and Decoherence* (Springer, Berlin, 2000) 221, 225
7. C.H. Bennett, D.P. DiVincenzo, C.A. Fuchs, T. Mor, E. Rains, P.W. Shor, J.A. Smolin, W.K. Wootters: Phys. Rev. A **59**, 1070 (1999) 223
8. D. Beckman, D. Gottesman, M.A. Nielsen, J. Preskill: Phys. Rev. A **64** 052309 (2001) 223, 224
9. T. Eggeling, Schlingemann, R.F. Werner: Europhys. Lett. **57**, 782 (2002) 223
10. B. Groisman, B. Reznik: Phys. Rev. A **66**, 022110 (2002) 224
11. L. Vaidman: Phys. Rev. Lett. **90**, 010402 (2003) 224, 225
12. M. Florig, S.J. Summers: J. Math. Phys. **38**, 1318 (1997) 225
13. R. Haag: *Local Quantum Physics: Fields, Particles, Algebras* (Springer, Berlin, 1996) 225
14. A. Kent: Phys. Rev. Lett. **83**, 1447 (1999) 225
15. A. Kent: Phys. Rev. Lett. **90**, 237901 (2003) 225
16. O. Keller: Phys. Rep. **411**, 1 (2005) 225
17. M. Toller: Phys. Rev. A **59**, 960 (1999) 225
18. S. Mazzucchi: J. Math. Phys. **42**, 2477 (2001) 225

19. W.-K. Tung: *Group Theory in Physics* (World Scientific, Singapore, 1985) 226
20. S. Weinberg: *The Quantum Theory of Fields* vol. 1 (Cambridge University Cambridge, 1995) 226
21. A.O. Barut, R. Raczka: *Theory of Group Representations and Applications* (World Scientific, Singapore, 1987) 230
22. N.N. Bogolubov, A.A. Logunov, A.I. Oksak, I.T. Todorov: *General Principles of Quantum Field Theory* (Kluwer, Dordrecht, 1990) 230, 231
23. F.R. Halpern: *Special Relativity and Quantum Mechanics* (Prentice-Hall, Englewood Cliffs, 1968) 230
24. C. Soo, C.C.Y. Lin: Int. J. Quant. Info. **2**, 183 (2004) 230, 241
25. N.L. Harshman: Phys. Rev. A **71**, 022312 (2005) 230, 243
26. D.R. Terno: Phys. Rev. A **67**, 014102 (2003) 231
27. P.M. Alsing, G.J. Milburn: Quant. Info. Comp. **2**, 487 (2002) 231
28. N.H. Lindner, A. Peres, D.R. Terno: J. Phys. A **36**, L449 (2003) 231
29. A.J. Bergou, R.M. Gingrich, C. Adami: Phys. Rev. A **68**, 042102 (2003) 231, 242
30. A. Peres, D.R. Terno: J. Mod. Opt. **50**, 1165 (2003) 235, 239
31. A. Aiello, J.P. Woerdman: Phys. Rev. A **70**, 023808 (2004) 235, 239
32. N.H. Lindner, D.R. Terno: J. Mod. Opt. **52**, 1177 (2005) 237, 239
33. M. Czachor, M. Wilczewski: Phys. Rev. A **68**, 010302(R) (2003) 239
34. M. Czachor: Phys. Rev. Lett. **94**, 078901 (2005) 239
35. P. Kok, T.C. Ralph, G.J. Milburn: Quant. Info. Comp. **5**, 239 (2005) 240
36. P. Kok, S.L. Braunstein: Int. J. Quant. Info. **4**, 119 (2006) 240
37. S.D. Bartlett, D.R. Terno: Phys. Rev. A **71**, 012302 (2005) 240
38. S.J. Summers, R. Werner: J. Math. Phys. **28**, 2440 (1987) 241
39. M. Czachor: Phys. Rev. A **55**, 72 (1997) 241
40. B. Reznik, A. Retzker, J. Silman: Phys. Rev. A **71**, 042104 (2005) 241
41. R.M. Gingrich, C. Adami: Phys. Rev. Lett. **89**, 270402 (2002) 242
42. R. Verch, R.F. Werner: Rev. Math. Phys. **17**, 545 (2005) 243
43. D.R. Terno: Phys. Rev. Lett. **93**, 051303 (2004) 243
44. M.B. Plenio, J. Eisert, J. Dreissig, M. Cramer: Phys. Rev. Lett. **94**, 060503 (2005) 243
45. P.M. Alsing, G.J. Milburn: Phys. Rev. Lett. **91**, 180404 (2003) 243
46. I. Fuentes-Schuller, R.B. Mann: Phys. Rev. Lett. **95**, 120404 (2005) 243
47. R. Schützhold, W.G. Unruh: e-print quant-ph/0506028 243
48. E.B. Manoukian, N. Yongram: Eur. J. Phys. D **31**, 137 (2004) 243
49. L. Lamata, J. Leon, E. Solano: Phys. Rev. A **73**, 012335 (2006) 243
50. O. Dreyer, F. Markopoulou, L. Smolin: Nucl.Phys. B **744**, 1 (2006) 243
51. E.R. Livine, D.R. Terno: Phys. Rev. A **72**, 022307 (2005) 243
52. E.R. Livine, D.R. Terno: Nucl. Phys. B **741**, 131 (2006) 243
53. D.R. Terno: Int. J. Mod. Phys. D **14**, 2307 (2005) 243
54. D. Campo, R. Parentani: Phys. Rev. D **74**, 025001 (2006) 243
55. R. Kallosh, A. Linde: Phys. Rev. D**73**, 104033 (2006) 243
56. R. Dijkgraaf, R. Gopakumar, H. Ooguri, C. Vafa: Phys. Rev. D **73**, 066002 (2006) 243
57. E.R. Livine and D.R. Terno: Phys. Rev. D **75**, 084001 (2007) 243, 245
58. E. Hawkins, F. Markopoulou, H. Sahlmann: Class. Quant. Grav. **20**, 3839 (2003) 243

Index

A
adiabatic
  passage, 115
  window, 88
algebra, 8
approximation
  Born, 82, 126
  Born-Markov , 130
  Markov, 83, 131
  Wigner-Weisskopf, 145

B
barrier spins, 120
beam splitter, 189
Bell
  state, 24
Bloch
  ball, 22
  or Poincaré sphere, 188
  sphere, 101
Bohr frequencies, 29, 85
boundary
  convex, 16
  of state space, 21
broker-client model, 209

C
C*-algebra, 8
C*-norm, 8
causal, 223
causality constraints, 221
Choi matrix, 34
CHSH–Bell type inequality, 158
classical probability model, 41
Clifford group, 193
communication channels, 240
commutation relations
  canonical, 188
complete positivity, 32, 58, 179
completely positive, 33
  evolution, 222
composite systems, 22
concurrence, 100, 103, 112, 168
  loss of, 182
condition
  Markov, 60, 131
conservation laws, 135
controlled
  closed systems, 29
  NOT (or CNOT), 101
convex, 16
correlation
  functions, 131
  time, 131
covariance, 59
CP
  eigenstates, 155
  violating parameter, 145, 155
CPT invariance, 145
cross
  Kerr nonlinearity, 211
  phase modulation, 211
cumulant
  expansion, 82
  ordered, 129

D
decoherence, 62, 66, 68, 72, 74, 125
  models, 179
density matrix, 13, 126
  reduced, 127
dissipation, 62, 66, 68, 72, 73
distribution
  B-, 5
  function, 6
double-heralding
  entangling procedure, 206
  protocol, 205

Brüning, E., Petruccione, F.: *Index*. Lect. Notes Phys. **787**, 249–253 (2010)
DOI 10.1007/978-3-642-02871-7
© Springer-Verlag Berlin Heidelberg 2010

dressed Hamiltonian, 93
dynamics
  embedding into a Lindblad -, 137
  reduced, 32
Dyson expansion, 61

E
effect, 45
encoding, 113
  dual-rail, 115
  Jamiołkowski-Choi, 34
engineering of couplings, 113
entanglement, 99, 141
  distillation, 112
  localizable , 106
  long distance, 105
  magnetic, 104
  measures of, 181
  of formation, 167
    loss of, 182
  quantifiers, 25
  thermal, 104
  transmission of, 111
  witnesses, 25
entropy
  Shannon, 18
  von Neumann, 100, 167
environment, 126
EPR
  correlations, 142
  reality criterion, 156
equation
  Lindblad, 130, 163
    non-Markovian generalization, 136, 138
  master, 61
  Nakajima-Zwanzig, 128
eraser
  active, 174
  partially passive quantum, 175
  passive, 176
erasing the past and impacting the future, 173
extreme points, 16

F
ferromagnet, 107
fidelity, 100, 109, 234
Fourier picture, 90
fringe visibility, 152

G
G-space, 47
gate, 187
  controlled phase, 101
  entangling, 121
  fusion, 196
    type-I, 196
  logic, 122
  quantum, 101
generator
  Lindblad, 163
  LKGS, 83
  of group of automorphisms, 27
  TCL, 129
group of automorphisms, 27

H
harmonic oscillator
  classical, 2
  damped, 62
Heisenberg
  chain, 119
  equation of motion, 28
  picture, 32
heralded perfect transfer, 115
Hilbert-Schmidt scalar product, 135
HOM effect, 194

I
instrument, 30, 54
interaction
  exchange, 101
  Heisenberg, 101
  Ising, 102
  XY, 102
interferometer, 191
Ising chain, 114

J
Jordan-Wigner fermions, 122

K
K-qubits, 143
kaonic–photonic analogy, 152
Kerr nonlinearities, 210
KLM approach, 192
KMS condition, 84
Kraus
  operators/effects, 216
  representation, 35, 54

L
Lévy processes, 74

limit
van Hove, 85
weak coupling, 85
LOCC, 223

M
map
adjoint, 135
completely positive, 131, 133, 134, 137
CPT -, 133
error -, 90
positive, 32
positive and trace-preserving, 132
mapping
covariant, 47
marginals, 23
Markovian, 32
mass
eigenstate basis, 150
eigenstates, 155
measurement
active, 149, 160, 174
generalized quantum , 133
passive, 149, 177
single-qubit, 198
verification, 224
memory kernel, 128
minimax principle, 18

O
observable, 7
as registration procedures, 44
automorphisms of, 26
position, 48
one-way model of quantum computing, 200
operation, 54
adjoint, 8
arbitrary local unitary, 101
entangling unitary, 101
local, 22
localizable, 223
semilocalizable, 223
operational partition of unity, 30
operator
creation and annihilation, 188
fusion
type-II, 197
Kraus, 59
Lindblad, 61
statistical, 43
Weyl, 93

P
path predictability, 152
PBS – polarizing beam splitter, 190
phase
shift, 189
space, 22
photon, 188
polarization, 235
rotation, 190
positive, 32
and state preserving, PTP, 32
positive operator valued measure, POVM, 30, 44, 216, 222
positivity, 8
primary variables, 231
projection
operator technique, 125, 127
Nakajima-Zwanzig, 128
valued measure, PVM, 45
projection operator method
time-convolutionless, 129
propagator, 80
PUP, positive unity preserving, 32

Q
quantum
Brownian motion, 68
causal histories, 243
channel, 133
communication, 107
computation, 118, 119
computer, 100
dynamical map, 131, 137
dynamical semigroup, 131
dynamics
non-Markovian, 131
ensemble, 14
entanglement, 99
eraser, 173
error correction, 193
fluctuation-dissipation theorem, 91
linear Boltzmann equation, 73
Lorentz transformation, 225
memory, 204
operation, 26, 30
oscillator, 3
phase transitions, 104
probability space, 13
response, 118
stimulus, 118
system
open, 126, 178
quantum computing
cluster state, 198

measurement-based , 198
one-way model of, 198
quasi–spin, 147
qubit, 187
memories, 204
states, 21

R
relaxation, 125
time, 131
representation
single/dual rail, 190

S
Schmidt
decomposition, 25
number, 26
Schrödinger
picture, 32
time dependent equation, 3
time independent equation, 4
Schwarz inequality for positive maps, 36
secondary variables, 232
semigroup evolution, 60
sequence
more mixed, 18
of probabilities, 14
stochastically dominating or majorizing, 18
shift-covariance, 64
simplices, 17
spectral density, 82
spooky action at distant, 162
standard matrix units, 11
state, 7, 12
a posteriori, 55
a priori, 55
as preparation procedures, 43
cluster, 198
computational basis, 188
conditional, 24
correlated initial, 130, 134
entangled, 134
entangled , 133
Jamiołkowski, 34
mixed, 20
product, 133
pure, 20
separable, 133, 134
separable or non-entangled, 25
single particle, 226
single photon, 228
space, 13
space, extended, 136
universal cluster, 200
vacuum, 189
strangeness, 143
basis, 150
eigenstates, 155
oscillation, 146
quantum numbers, 144
subsystems
independent, 25
super-operator, 28, 34
correlated projection, 132
Liouville, 128
super-propagator, 80
system of imprimitivity, 49

T
teleportation, 112
identity, 224
protocol, 225
trick, 192
tensor product, 126
theorem
Bell's, 157
Birkhoff's, 17
Carathéodory's, 17
Choi, 34
Gelfand, 10
Gelfand-Naimark-Segal representation, 14
Lidskii's, 18
Lindblad, 37
of Krein and Milman, 16
Rado's, 18
representation, 133
Riesz's representation, 13
thermal memory time, 84
trace
partial, 23
transposition, 33, 134
two-positivity inequality, 36

U
unitary dilation, 83
unity preserving, 32

V
von Neumann
equation, 127
measurement, 29
projection postulate, 56

W
Weyl
form, 72
operators, 51
white noise, 84
Wigner
rotation matrix, 228
spin, 227

X
XY spin chains, 122

Z
Zeno
effect, 82, 214, 215
gate, 214