Lecture Notes in Physics

Editorial Board

R. Beig, Wien, Austria
B.-G. Englert, College Station, TX, USA
U. Frisch, Nice, France
P. Hänggi, Augsburg, Germany
K. Hepp, Zürich, Switzerland
W. Hillebrandt, Garching, Germany
D. Imboden, Zürich, Switzerland
R. L. Jaffe, Cambridge, MA, USA
R. Kippenhahn, Göttingen, Germany
R. Lipowsky, Golm, Germany
H. v. Löhneysen, Karlsruhe, Germany
I. Ojima, Kyoto, Japan
D. Sornette, Nice, France, and
Los Angeles, CA, USA
H. A. Weidenmüller, Heidelberg, Germany
W. Weise, Villazzano/Trento, Italy, and
Garching, Germany
J. Wess, München, Germany
J. Zittartz, Köln, Germany

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Physics and Astronomy ONLINE LIBRARY

http://www.springer.de/phys/

Editorial Policy

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 Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to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

Acceptance

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

Contractual Aspects

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

Manuscript Submission

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

Online Version/ LNP Homepage

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

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

Dieter Heiss (Ed.)

Fundamentals of Quantum Information

Quantum Computation, Communication, Decoherence and All That

Springer

Editor

Dieter Heiss
Department of Physics
University of Stellenbosch
7602 Matieland, South Africa

Cover picture: with courtesy of Mrs. J.A. de Waal, University of Stellenbosch, South Africa

Library of Congress Cataloging-in-Publication Data.

Fundamentals of quantum information : Quantum computation, communication, decoherence, and all that / Dieter Heiss (ed.).
p. cm. -- (Lecture notes in physics, ISSN 0075-8450 ; 587)
Includes bibliographical references.
ISBN 3540433678 (alk. paper)
1. Quantum computers. 2. Quantum theory. I. Heiss, Dieter, 1939- II. Series.

QA76.889 .F86 2002
004.1--dc21

2002021847

ISSN 0075-8450
ISBN 3-540-43367-8 Springer-Verlag Berlin Heidelberg New York

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-Verlag. Violations are liable for prosecution under the German Copyright Law.

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

http://www.springer.de

© Springer-Verlag Berlin Heidelberg 2002
Printed in Germany

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.

Typesetting: Print-ready by the authors/editor
Camera-data conversion by Steingraeber Satztechnik GmbH Heidelberg
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
SPIN: 10869082 54/3141/du - 5 4 3 2 1 0

Preface

Quantum computers do not exist as of yet, it may take more than a decade until usable devices come to realisation; some people are even doubtful as to whether a large-scale quantum computer is a possibility in principle. However, the basic ideas formulated just a bit more than 10 years ago had a tremendous impact upon research in the special field in particular, and they have initiated substantial progress towards deeper understanding of quantum mechanics in general. Modern experimental techniques have provided convincing evidence about various aspects of "quantum weirdness" in that, for instance, entanglement or teleportation are established as physical realities. Also, the mysterious collapse of the wave function is now replaced by a thorough understanding of the dynamical process which is decoherence.

It is in the spirit and tradition of the South African Chris Engelbrecht Summer School in Theoretical Physics that the topical subject implied in the title of the present issue was chosen as the theme of the School in 2001. This volume presents pertinent contributions of some of the leading researchers at the front of this highly topical field. It comprises material that is well balanced between theoretical aspects and experimental realisation. The broad spreading of contributions in this fast-growing field is expected to ensure, for this volume, a special place in an exploding proliferation of monographs, books and other forms of literature. This volume also reflects nicely the speakers' endeavour to guide the audience from the beginners level to the present state of the art.

The participants came from a larger variety of backgrounds than ever before at this School owing to the interdisciplinary character of the topic. We are also delighted that a substantial number came from adjacent countries and from further afield. The togetherness and the friendly working atmosphere of speakers and listeners were noted by all participants.

The Organising Committee is indebted to the National Research Foundation for its financial support, without which such high-level courses would be impossible. We also wish to express our thanks to the Editors of Lecture Notes in Physics and Springer-Verlag for their agreement to publish and for assistance in the preparation of these proceedings.

Johannesburg, South Africa *W.D. Heiss*
October 2001

List of Contributors

Guido Burkard
Department of Physics
and Astronomy,
University of Basel,
Klingelbergstrasse 82,
CH-4056 Basel, Switzerland

Dik Bouwmeester
Centre for Quantum Computation,
Clarendon Laboratory,
University of Oxford,
Parks Road, OX1 3PU Oxford, UK

Juan Ignacio Cirac
Max-Planck-Institut
für Quantenoptik,
Hans-Kopfermann Str. 1,
D-85748 Garching, Germany

Artur Ekert
Centre for Quantum Computation,
Clarendon Laboratory,
University of Oxford,
Parks Road, OX1 3PU Oxford, UK

Hans-Andreas Engel
Department of Physics
and Astronomy,
University of Basel,
Klingelbergstrasse 82,
CH-4056 Basel, Switzerland

John C. Howell
Centre for Quantum Computation,
Clarendon Laboratory,
University of Oxford,
Parks Road, OX1 3PU Oxford, UK

Antia Lamas-Linares
Centre for Quantum Computation,
Clarendon Laboratory,
University of Oxford,
Parks Road, OX1 3PU Oxford, UK

A.J. Leggett
Department of Physics,
University of Illinois,
1110 West Green Street,
Urbana, IL 61801-3080, USA

Daniel Loss
Department of Physics
and Astronomy,
University of Basel,
Klingelbergstrasse 82,
CH-4056 Basel, Switzerland

Juan Pablo Paz
Departamento de Física
J.J. Giambiagi,
FCEN, UBA,
Pabellón 1, Ciudad Universitaria,
1428 Buenos Aires, Argentina

Wojciech Hubert Zurek
Theoretical Astrophysics, MS B288,
Los Alamos National Laboratory,
Los Alamos, NM 87545, USA

Table of Contents

A Guide for the Reader 1

Qubits, Cbits, Decoherence, Quantum Measurement and Environment
A.J. Leggett 3
1 Lecture 1 – A Single Qubit: Basic Notions 3
2 Lecture 2 – Distinguishing Qubits from Cbits: Multi-qubit Systems and Quantum Computation 10
3 Lecture 3 – The Fundamental Theorem of Decoherence: “Isolated” Systems 16
4 Lecture 4 – Candidate Systems for Qubits: Preliminary Estimate of Decoherence Times 22
5 Lecture 5 – The Environment as a Bath of Oscillators: More Systematic Treatment of Decoherence 30
6 Lecture 6 – The Quantum Measurement Problem 37

Introduction to Quantum Computation
Artur Ekert 47
1 Physical Representation of Information 47
2 Qubits, Gates and Networks 48
3 Quantum Arithmetic and Function Evaluations 53
4 Algorithms and Their Complexity 56
5 From Interferometers to Computers 59
6 The First Quantum Algorithms 63
7 Quantum Search 65
8 Optimal Phase Estimation 67
9 Periodicity and Quantum Factoring 69
10 Cryptography 71

Environment-Induced Decoherence and the Transition from Quantum to Classical
Juan Pablo Paz, Wojciech Hubert Zurek 77
1 Introduction and Overview 77
2 Quantum Measurements 80
2.1 Bit-by-Bit Measurement and Quantum Entanglement 82

2.2 Interactions and the Information Transfer in Quantum Measurements . . . 85
2.3 Monitoring by the Environment and Decoherence . . . 87
2.4 One-Bit Environment for a Bit-by-Bit Measurement . . . 88
2.5 Decoherence of a Single (Qu)Bit . . . 90
2.6 Decoherence, Einselection, and Controlled Shifts . . . 94
3 Dynamics of Quantum Open Systems: Master Equations . . . 96
3.1 Master Equation: Perturbative Evaluation . . . 97
3.2 Example 1: Perturbative Master Equation in Quantum Brownian Motion . . . 100
3.3 Example 2: Perturbative Master Equation for a Two-Level System Coupled to a Bosonic Heat Bath . . . 102
3.4 Example 3: Perturbative Master Equation for a Particle Interacting with a Quantum Field . . . 105
3.5 Exact Master Equation for Quantum Brownian Motion . . . 106
4 Einselection in Quantum Brownian Motion . . . 112
4.1 Decoherence of a Superposition of Two Coherent States . . . 112
4.2 Predictability Sieve and Preferred States for QBM . . . 115
4.3 Energy Eigenstates Can Also Be Selected by the Environment! 117
5 Deconstructing Decoherence: Landscape Beyond the Standard Models . . . 118
5.1 Saturation of the Decoherence Rate at Large Distances . . . 119
5.2 Decoherence at Zero Temperature . . . 120
5.3 Preexisting Correlations Between the System and the Environment . . . 122
6 Decoherence and Chaos . . . 125
6.1 Quantum Predictability Horizon: How the Correspondence Is Lost . . . 125
6.2 Exponential Instability vs. Decoherence . . . 127
6.3 The Arrow of Time: A Price of Classicality? . . . 129
6.4 Decoherence, Einselection, and the Entropy Production . . . 132
7 How to Fight Against Decoherence: Quantum Error Correcting Codes . . . 133
7.1 How to Protect a Classical Bit . . . 134
7.2 How to Protect a Quantum Bit . . . 134
7.3 Stabilizer Quantum Error-Correcting Codes . . . 140
8 Discussion . . . 143

Quantum Information Science Using Photons
Dik Bouwmeester, John C. Howell, Antia Lamas- Linares . . . 149
1 Introduction . . . 149
1.1 A Humble Point of View . . . 149
1.2 Quantum Mystery . . . 150
1.3 Maxwell's Demon . . . 151
1.4 Shannon Entropy . . . 154

1.5 Von Neumann Entropy ... 156
2 Einstein–Podolsky–Rosen Paradox and Bell's Inequalities ... 157
3 Producing Entangled Particles ... 159
3.1 Introduction ... 159
3.2 Parametric Down-Conversion ... 159
3.3 Franson's Proposal ... 160
3.4 Polarization Entanglement ... 161
4 The Beam Splitter Action on a Two-Photon State ... 163
4.1 Beamsplitter Transformation ... 163
4.2 Bell-State Analyzer ... 164
5 No-Cloning Theorem ... 165
6 Quantum Cryptography ... 167
7 Quantum Dense Coding ... 169
7.1 Theoretical Scheme ... 169
7.2 Experimental Dense Coding with Qubits ... 170
8 Quantum Teleportation ... 171
8.1 Theoretical Scheme ... 171
8.2 Experimental Quantum Teleportation of Qubits ... 173
8.3 Teleportation of Entanglement ... 174
8.4 A Two-Particle Scheme for Quantum Teleportation ... 174
9 Teleportation of Continuous Quantum Variables ... 177
9.1 Theoretical Scheme ... 177
9.2 Quantum Optical Implementation ... 178
10 Quantum Error Detection and Correction ... 181
10.1 Introduction ... 181
10.2 Quantum Error Detection ... 181
10.3 Avoiding Controlled-NOT Operations ... 183
10.4 Post-selection ... 184
11 Stimulated Entanglement ... 185
11.1 Theory ... 185
12 Bohm-Type Spin-s Entanglements ... 193

Quantum Information: Entanglement, Purification, Error Correction, and Quantum Optical Implementations
Juan Ignacio Cirac ... 199
1 Introduction ... 199
2 Quantum Computers: Quantum Optical Implementations ... 200
2.1 Introduction ... 200
2.2 How to Construct a Quantum Computer ... 201
2.3 Physical Implementations for Quantum Computation ... 202
2.4 Quantum Optical Systems for Quantum Computation ... 202
2.5 Quantum Computation with Trapped Ions ... 203
3 Decoherence and Error Correction ... 205
3.1 Introduction ... 205

3.2 Decoherence 205
3.3 Error Correction 206
3.4 Error Correction and Decoherence 209
4 Entanglement of Pure States 210
4.1 Introduction 210
4.2 2-Partite Entanglement: Definition 211
4.3 2-Partite Entanglement: Qualifying and Quantifying 212
4.4 Multipartite Entanglement 214
5 Entanglement of Mixed States 215
5.1 Introduction 215
5.2 2-Partite Entanglement: Definition 217
5.3 2-Partite Entanglement: Entanglement Witnesses 219
5.4 2-Partite Entanglement: Positive Maps 220
5.5 2-Partite Entanglement: Entangled States with Positive Partial Transposition 224
5.6 Multipartite Entanglement 224
6 Entanglement Distillation 226
6.1 Introduction 226
6.2 Generalized Measurements 227
6.3 Distillation of Pure States 229
6.4 Distillation of Mixed States 231
6.5 Multipartite Distillation 236

Spintronics, Quantum Computing, and Quantum Communication in Quantum Dots
Guido Burkard, Hans-Andreas Engel, Daniel Loss 241
1 Introduction 241
1.1 Quantum Computing 242
1.2 Quantum Communication 242
1.3 General Remarks on Quantum Dots 243
2 Phase Coherence 243
3 Pulsed Switching and Adiabaticity 245
4 Initialization of a Quantum Register 247
5 Two-Qubit Operations in Coupled Quantum Dots 247
5.1 Laterally Coupled Dots 248
5.2 Vertically Coupled Dots 250
5.3 Singlet–Triplet Entangling Gate 253
6 Single-Spin Operations 253
7 Single-Spin Measurement 254
7.1 Spin Measurements Through Spontaneous Magnetization 254
7.2 Spin Measurements via the Charge 254
7.3 Quantum Dot as Spin Filter and Read-Out/Memory Device 255
7.4 Optical Measurements 257
8 Quantum Communication with Entangled Electrons 257
8.1 Adding Entangled Electrons to the Fermi Sea 257
8.2 Shot Noise of Entangled Electrons 258

8.3 Spin-Dependent Current Through a Double Dot–Probing Entanglement 260
8.4 Double Dot with Superconducting Leads 262

A Guide for the Reader

When the term 'quantum information' is used nowadays in general discussions, most people associate the expression with quantum computers, some may have heard of the mysterious-sounding 'teleportation', but only a few would usually have a thorough understanding of these subtle concepts and be aware of the conceptional difficulties, let alone their possible physical implementation in the laboratory. The selection of contributions in the present volume addresses all aspects that are essential for a basic theoretical understanding of the terms mentioned, as well as the most promising experimental implementations known at present. While each of the six lectures appears to have a different emphasis or even topic, they are closely intertwined thematically.

The first contribution provides the basic and elementary quantum mechanical prerequisites essential for the whole theme. The concept of the qubit is presented and the ensuing discussion of multi-qubit systems clarifies the crucial difference between a classical and a prospective quantum computer. Decoherence, the major fundamental obstacle that has so far prevented the existence of a large-scale quantum computer, is dealt with next. Its mechanism and its ubiquitous occurrence, owing to an ever present environment, is addressed. Simple models provide quantitative estimates of the extremely fast decoherence times.

The next contribution presents the basics of quantum computation. The definition of a qubit is revisited in a more abstract way, followed by an introduction to gates, networks and quantum arithmetic function evaluation. The necessary elementary background of number theory is provided, enabling the reader to follow Shor's famous procedure of finding large prime factors of large digit numbers. A basic introduction into quantum cryptography concludes these lectures. A considerable expansion of this last topic is given in some of the following lectures.

The all important topic of decoherence is dealt with in great depth in the third lecture. In a variety of settings the dynamical process of decoherence and its significance for the transition from the quantum mechanical to the classical world is discussed in great detail. Using the appropriate technical language – the density matrix formulation – we demonstrate how the superselection rules (einselection) are induced by the environment to transform a pure state into a mixture. Master equations – describing the time evolution of the density matrix – are formulated for a few significant examples to demonstrate this dynamical process. Considerations of chaotic systems suggest a similar physical origin for quantum decoherence and the second law of thermodynamics. Needless to say that a thorough understanding of decoherence is crucial in the attempt to remove this main obstacle to computation.

While the first two lectures provide the reader with the very basics of quantum information, this third lecture is already more specialised in that one very central aspect of quantum computation, decoherence and its possible remedies, is presented up to the present frontiers of our understanding. The following three contributions are devoted to the actual physical realisations of prototype quantum computers, or, in technical terms, of realisations of entangled states, teleportation, error detection and correction. Where necessary, and for the sake of self-consistency, the important theoretical background is of course given. Much of the results presented are new and thus constitute the up-to-date state of the art.

After a pertinent introduction into basic principles of information transfer, including various definitions of entropy, the actual implementation of entangled states using polarised photons is presented in detail in the fourth lecture. Communication and teleportation between the notorious players Alice and Bob are discussed for a variety of physical settings. In this way, quantum dense coding, cryptography, error detection and correction are dealt with thoroughly. The same theme is addressed in the fifth lecture, albeit from different viewing angles, and in many cases with a different emphasis on details. Needless to say that the problem of the eavesdropper features in the discussion of secret communication channels.

While photons, and perhaps trapped ions, bear some promise for an actual implementation of a quantum computer there is of course the spin degree of freedom that can be used as a further tool. Historically it was, after all, the singlet state of two electrons that served as *the* paradigm for entanglement and nonlocality. The pertinent subject *spintronics* is treated in the sixth lecture of this volume. Quantum dots are the obvious candidates for manipulating electrons, as great progress has been made in handling them experimentally. A thorough treatise of theoretical aspects as well as of the experimental implementation is presented here. The working of the relevant gates, switching times, and memory devices are some of the important key issues featured here.

The series of lectures presented here give a comprehensive overview of the state of the art of quantum information. The issues at the forefront of research in the field are discussed in great detail. While each of the six contributions can be read on its own, the reader is encouraged to read first the very basics contained in the first two lectures. Each contribution carries the personal stamp of the author, and this is the additional and special charm of this volume.

Qubits, Cbits, Decoherence, Quantum Measurement and Environment

A.J. Leggett

Abstract. I present a tutorial review of the behavior of a simple 2-state quantum system ("qubit") in interaction with a dissipative environment, with particular attention to the questions of when we can regard the system as effectively isolated and of "true" versus "false" decoherence. The last lecture discusses the quantum measurement paradox in the light of recent experiments on Josephson systems.

1 Lecture 1 – A Single Qubit: Basic Notions

Let's start by considering the simplest quantum-mechanical system known to man, namely a free particle of spin 1/2. This is often said to constitute a single "qubit". In the following, it should always be remembered that the formalism of quantum mechanics is usually taken to apply not to individual systems but to ensembles of systems, that is the collection of systems prepared according to some specified recipe. For brevity in the following, I shall often speak of "the system", but this should always be read, unless explicitly otherwise specified, as "the relevant ensemble of similarly prepared systems". We recall some basic features of this system:

(1) The system can be described in terms of the three operators s_i $(i = x, y, z)$ corresponding to the Cartesian components of spin, or more conveniently by the Pauli operators $\sigma_i \equiv 2s_i/\hbar$; the latter satisfy the commutation relations (a special case of the general relation for angular momentum)

$$[\hat{\sigma}_{\mathrm{i}}, \hat{\sigma}_j] = 2\mathrm{i}\varepsilon_{ijk}\hat{\sigma}_k \qquad (\text{i.e.}[\sigma_x, \sigma_y] = 2i\sigma_z\text{etc.}) \tag{1.1}$$

and also the anticommutation relations

$$\{\hat{\sigma}_i, \hat{\sigma}_j\} \equiv \hat{\sigma}_i\hat{\sigma}_j + \hat{\sigma}_j\hat{\sigma}_i = 2\delta_{ij} \tag{1.2}$$

(2) The Hilbert space is two-dimensional; the basis vectors are conventionally chosen as the eigenstates of $\hat{\sigma}_z$ corresponding to eigenvalue ±1, and labelled $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively; thus,

$$\hat{\sigma}_z \, |\uparrow\rangle \; = |\uparrow\rangle \qquad \hat{\sigma}_z \, |\downarrow\rangle = - \, |\downarrow\rangle$$

It follows that the operator $\hat{\sigma}_z$ has in this basis the matrix representation

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.3}$$

W. Dieter Heiss (Ed.): LNP 587, pp. 3–45, 2002.
© Springer-Verlag Berlin Heidelberg 2002

Note that this choice does not define the relative phase of the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, i.e. we can make the transformation

$$|\uparrow\rangle \to \exp\ \mathrm{i}\varphi_\uparrow\ |\uparrow\rangle,\ |\downarrow\rangle \to \exp\ \mathrm{i}\varphi_\downarrow\ |\downarrow\rangle \tag{1.4}$$

($\varphi_\uparrow \neq \varphi_\downarrow$ in general) without affecting the form of the matrix representing $\hat{\sigma}_z$.

The conventional choice of relative phase is such that the matrix representing $\hat{\sigma}_x$ has the form

$$\hat{\sigma}_x = \begin{pmatrix} 0\ 1 \\ 1\ 0 \end{pmatrix} \tag{1.5}$$

Then it follows from the commutation relation (1.1) that the form of the matrix $\hat{\sigma}_y$ is uniquely determined:

$$\hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \tag{1.6}$$

(3) Consider a general normalized pure state of the system, represented by a state vector which is a coherent superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$,

$$|\ \psi\rangle = \alpha\ |\uparrow\rangle + \beta\ |\downarrow\rangle \qquad \left(\equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right) \tag{1.7}$$

where α and β are complex coefficients which to ensure normalization of $|\ \psi\rangle$ must satisfy the condition

$$|\ \alpha\ |^2 + |\ \beta\ |^2 = 1. \tag{1.8}$$

Then we have the theorem:

There exists a (real) unit vector $\mathbf{n}$ in ordinary (Cartesian) space such that $|\ \psi\rangle$ is an eigenstate of $\mathbf{n}\cdot\hat{\boldsymbol{\sigma}}$ with eigenvalue $+1$, i.e.

$$\mathbf{n}\cdot\hat{\boldsymbol{\sigma}}\ |\ \psi\rangle = +\ |\ \psi\rangle \tag{1.9}$$

i.e., intuitively, any pure state $|\ \psi\rangle$ corresponds to the spin "pointing" along direction $\mathbf{n}$ (though see below).

To prove the theorem, we solve for $\mathbf{n}$ in terms of the complex numbers α and β. Using the explicit matrix forms of the operators $\hat{\sigma}_i$, we have

$$\mathbf{n}\cdot\hat{\boldsymbol{\sigma}}\ |\ \psi\rangle \equiv (n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z)\ |\ \psi\rangle = \begin{pmatrix} n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & -n_z \end{pmatrix} |\ \psi\rangle \tag{1.10}$$

$$\equiv \begin{pmatrix} n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & -n_z \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{1.11}$$

We therefore require

$$\begin{pmatrix} n_z & n_x - \mathrm{i}n_y \\ n_x + \mathrm{i}n_y & -n_z \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{1.12}$$

or explicitly

$$n_z\alpha + (n_x - \mathrm{i}n_y)\beta = \alpha \tag{1.13a}$$
$$(n_x + \mathrm{i}n_y)\alpha - n_z\beta = \beta \tag{1.13b}$$

The condition that $\mathbf{n}$ is a unit vector guarantees the vanishing of the determinant, so that this pair of equations always has a solution for any α, β, namely

$$(n_x + \mathrm{i}n_y)/(1 + n_z) = \beta/\alpha \tag{1.14b}$$

or since all components of $\mathbf{n}$ are real

$$n_x/(1 + n_z) = \ \mathrm{Re}\,(\beta/\alpha) \tag{1.15a}$$
$$n_y/(1 + n_z) = \ \mathrm{Im}\,(\beta/\alpha) \tag{1.15b}$$

It is convenient to introduce two "angles" θ, φ by the definitions

$$\theta \equiv 2\tan^{-1} \mid \beta/\alpha \mid, \qquad \varphi \equiv 2\arg\,(\beta/\alpha) \tag{1.16}$$

so that in view of the normalization condition (1.8) we have

$$\alpha = \cos\ \theta/2 \cdot \exp\,(-\mathrm{i}\varphi/2), \qquad \beta = \sin\ \theta/2 \cdot \exp\,(+\mathrm{i}\varphi/2) \tag{1.17}$$

Then it follows from eqns. (1.15) and the normalization condition $n_x^2 + n_y^2 + n_z^2 = 1$ that (recall that $\tan\ \theta/2 \equiv \sin\ \theta/(1 + \cos\ \theta)$!)

$$n_x = \sin\ \theta \cdot \cos\ \varphi \tag{1.18a}$$
$$n_y = \sin\ \theta \cdot \sin\ \varphi \tag{1.18b}$$
$$n_z = \cos\ \theta \tag{1.18c}$$

i.e., if α and β are parametrized by θ, φ as in (1.17), then θ and φ are just respectively the azimuthal and polar angles of the direction along which the spin is "pointing."

Actually, the above phrase is a bit ambiguous. The value of the total spin squared in units of $\hbar/2$ is $\hat{\sigma}^2 \equiv \hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2$, and since $\hat{\sigma}_i^2 \equiv 1$ (eqn. (1.2)) this gives $\hat{\sigma}^2 = 3$. Since $(\mathbf{n}{\cdot}\boldsymbol{\sigma})^2 = 1$, this means that the component of σ transverse to $\mathbf{n}$ must have a value of $\sqrt{2}$, however we do not know in which direction it lies. This is, of course, a general feature of any problem involving a nonzero angular momentum, and is a reflection of the fact that the different components of the angular momentum operator fail to commute with one another. The pure state gives maximum possible information about the state of the system; how much this is, we investigate below.

It is straightforward to verify that in a state described as in (1.17) by the wave function

$$\mid \psi\rangle = \begin{pmatrix} \cos\ \frac{\theta}{2} \cdot e^{-\mathrm{i}\varphi/2} \\ \sin\ \frac{\theta}{2} \cdot e^{\mathrm{i}\varphi/2} \end{pmatrix} \tag{1.19}$$

so that it is an eigenfunction of $\mathbf{n}\cdot\hat{\boldsymbol{\sigma}}$ with eigenvalue +1, where θ and φ are the angles of $\mathbf{n}$ as above, the expectation values of the various spin components are given, as we should intuitively expect, by

$$\langle\sigma_x\rangle = \sin\ \theta\cos\ \varphi \equiv n_x \tag{1.20a}$$
$$\langle\sigma_y\rangle = \sin\ \theta\sin\ \varphi \equiv n_y \tag{1.20b}$$
$$\langle\sigma_z\rangle = \cos\ \theta \equiv n_z \tag{1.20c}$$

(4) Any Hermitian operator $\hat{\Omega}$ on the 2D Hilbert space on the system can be written in the form

$$\hat{\Omega} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}^* & \Omega_{22} \end{pmatrix} \tag{1.21}$$

It can therefore be represented in the form

$$\hat{\Omega} \equiv \Omega_0\hat{1} + \Omega_z\hat{\sigma}_z + \Omega_x\hat{\sigma}_x + \Omega_y\hat{\sigma}_y \equiv \Omega_0\hat{1} + \mathbf{\Omega}\cdot\boldsymbol{\sigma} \tag{1.22}$$

where Ω_0 and all components of the vector $\boldsymbol{\Omega}$ are real. In particular, since the Hamiltonian $\hat{H}$ must be a Hermitian operator, and any term proportional to the unit matrix is irrelevant as regards the dynamics of the system, we can write generally (introducing a - sign and a factor of 1/2 for convenience)

$$\hat{H}(t) \equiv -\frac{1}{2}\boldsymbol{\mathcal{H}}(t)\cdot\hat{\boldsymbol{\sigma}} \tag{1.23}$$

where $\boldsymbol{\mathcal{H}}$(t) can be regarded formally as an effective "magnetic field" acting on the system.

(5) The most general state of our spin-1/2 system is described not by a wave function of the form (1.7) but by a 2 × 2 density matrix $\hat{\rho}$, which is by definition Hermitian and of unit trace (see e.g. ref. [1], §12). For the special case of a pure state of the form (1.7) $\hat{\rho}$ takes the form

$$\hat{\rho}_{\text{pure}} = \begin{pmatrix} |\,\alpha\,|^2 & \alpha^*\beta \\ \alpha\beta^* & |\,\beta\,|^2 \end{pmatrix} \qquad (=|\,\psi\rangle\langle\psi\,|) \tag{1.24}$$

By an appropriate choice of basis (such that the "direction of spin" $\mathbf{n}$ is taken as the z-axis) it becomes

$$\hat{\rho}_{\text{pure}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{1.25}$$

In the general case, the Hermitian property guarantees that the density matrix can always be diagonalized, and it is convenient for the present discussion to choose the basis so that it is diagonal in the "standard" (σ_z-) representation. With this convention the most general form of $\hat{\rho}$ is then, using the property of unit trace,

$$\hat{\rho} = \begin{pmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{pmatrix} \quad , \quad \rho_{11} + \rho_{22} = 1. \tag{1.26}$$

Note that the special case of a pure state with $\sigma_z = +1(-1)$ simply corresponds to the choice $\rho_{22} = 0$ ($\rho_{11} = 0$).

It is possible to interpret the density matrix (1.26) intuitively in a number of different ways:

(a) Suppose we know for sure that the ensemble we are trying to describe is actually a mixture of two different subensembles which are known to correspond to $\sigma_z = +1$ and $\sigma_z = -1$ respectively. (Such a situation might arise, for example, if we are talking about a beam of particles obtained by combining the output from two different "filters" having the property of allowing through only particles with $\sigma_z = +1$ and $\sigma_z = -1$ respectively). Then the density matrix of the complete ensemble is of the form (1.26), with ρ_{11} and ρ_{22} respectively the "weights" of the $+1$ and -1 ensembles (i.e. the probability that a particle chosen at random from the complete ensemble came from the $+1(-1)$ subensemble).

(b) As we saw in (3) above, if the ensemble were in a pure state there would be a unit vector $\mathbf{n}$ such that $\mathbf{n} \cdot\hat{\boldsymbol{\sigma}} \mid \psi\rangle = +\mid \psi\rangle$, i.e. the spin "lies along" $\mathbf{n}$ and the expectation value of the spin component σ_i (i $= x, y, z$) is just n_i. Now, it may be that we have reason to believe that for the ensemble in question $\langle\sigma_z\rangle$ has a value intermediate between $+1$ and -1, but $\langle\sigma_x\rangle = \langle\sigma_y\rangle = 0$; such a situation would naturally arise, for example, for a gas of particles in thermal equilibrium at temperature T in a magnetic field directed along the z-axis. One way of interpreting this result is to regard the complete ensemble as a mixture of different subensembles each corresponding to a definite value of $\mathbf{n}$, with a fixed azimuthal angle θ but random polar angle φ, equally weighted over φ. Since the value of (e.g.) $\langle\sigma_x\rangle$ for the whole ensemble is the appropriately weighted average over that for the subensembles, we have

$$\langle\sigma_x\rangle = \frac{1}{2\pi}\int_0^{2\pi} \sin\ \theta \cos\ \varphi \,\mathrm{d}\varphi \equiv 0 \tag{1.27}$$

as required, (and similarly for $\langle\sigma_y\rangle$), while

$$\langle\sigma_z\rangle = \frac{1}{2\pi}\int_0^{2\pi} \cos\ \theta \,\mathrm{d}\varphi = \cos\ \theta \neq 0. \tag{1.28}$$

(in general).

This is a special case of the general result that a quantum-mechanical mixture of two states is equivalent to a mixture of coherent superpositions of the two states with random relative phase (in this case the relative phase is just the polar angle φ). Incidentally this example shows that the "basis" with respect to which a mixture-type density matrix is defined is in some intuitive sense not unique, (i.e. it can be represented either as a "mixture" of the two states $|\uparrow\rangle$ and $|\downarrow\rangle$, or of the continuum of (non-mutually orthogonal) coherent superpositions of these two). Note that θ is in some sense still well-defined for the mixture even though φ is not: we can write $\rho_{11} \equiv cos^2\theta/2$, $\rho_{22} \equiv sin^2\theta/2$.

(c) Finally, a density matrix of the form (1.21) may arise if the system in question is "entangled" with something else and we wish to average over the behavior of the “something else”. I postpone discussion of this interpretation for a little while.

(6) Let’s now investigate the form of $\hat{\rho}$ in a general representation. Its Hermiticity guarantees that it can be put in the form (1.22), and since Tr $\hat{\sigma}_i = 0$ the property of unit trace implies that Ω_0 is 1/2. Thus, relabelling the real vector $\boldsymbol{\Omega}$ for this case as $\frac{1}{2}\ \boldsymbol{\rho}$, we have quite generally

$$\hat{\rho} = \frac{1}{2}(\hat{1} + \boldsymbol{\rho} \cdot \hat{\boldsymbol{\sigma}}) \tag{1.29}$$

so that the density matrix is completely specified by the real vector $\boldsymbol{\rho}$. From the condition Tr $\hat{\rho}^2 \leq 1$ (most easily proved by using the representation (1.26) it follows that $|\boldsymbol{\rho}| \leq 1$, the equality holding only for a pure state. It is also easy to see that the expectation value $\langle \boldsymbol{\sigma} \rangle$ of the spin in units of $\hbar/2$ is given by

$$\langle \boldsymbol{\sigma} \rangle \equiv \mathrm{tr}\ \hat{\rho}\hat{\boldsymbol{\sigma}} = \boldsymbol{\rho} \qquad (|\langle \boldsymbol{\sigma} \rangle| \leq 1) \tag{1.30}$$

so that we can equally well characterize the state completely by the value of $\langle \boldsymbol{\sigma} \rangle$.

(7) Dynamics: the dynamics of the two-state system is extremely simple. As we have just seen, the state can be completely characterized by the value of $\langle \boldsymbol{\sigma} \rangle$, so it is sufficient to calculate the rate of change of this quantity in time: with units chosen so that $\hbar = 1$,

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{\sigma} \rangle = \langle [\boldsymbol{\sigma}, \hat{H}(t)] \rangle . \tag{1.31}$$

Since, however, the most general form of $\hat{H}(\mathrm{t})$ with which we need be concerned is (1.23), we immediately find, using the commutation relations (1.2), the equation

$$\frac{d\langle \boldsymbol{\sigma} \rangle}{\mathrm{d}t} = \langle \boldsymbol{\sigma} \rangle \times \boldsymbol{\mathcal{H}}(t) \tag{1.32}$$

which is formally nothing but the equation of precession of a classical spin in a magnetic field $\boldsymbol{\mathcal{H}}(\mathrm{t})$. Note that eqn. (2.32) conserves $|\langle \boldsymbol{\sigma} \rangle|$ and hence the “degree of purity” (see below) of the state (as must be the case for any unitary evolution).

(8) Finally, let us consider from a rather intuitive point of view the “amount of information” contained in the description of a given ensemble by a pure state or mixture form of density matrix. Let us start for simplicity with the case $\theta = \pi/2$ ($\mid \alpha \mid = \mid \beta \mid$). Then, both for the pure case and the mixture, we know that the spin lies in the x-plane; however, for the pure state we also know (from the relative phase φ of α and β) the direction in this plane, and can by a suitable transformation of coordinates reduce the wave function to the simple form $|\uparrow\rangle$. For the mixture, by contrast, we have no knowledge of the direction in the xy-plane (cf. above) and no transformation will yield a form of $\hat{\rho}$ corresponding to (1.26) with $\rho_{22} = 0$. So, qualitatively speaking, it is clear that the pure state encodes more information than the mixture.

On the other hand, let us consider a density matrix of the form (1.26) with $\rho_{11} \gg \rho_{22}$, i.e. $\theta \ll 1$. It is clear intuitively that this form of $\hat{\rho}$ carries a great deal more “information” than the one discussed above, in fact almost as much as the pure state which it approaches as $\theta \to 0$. We would like to have some definition of the "amount of information" or "degree of purity" of the state which reflects this fact. Moreover, this measure should be invariant under the choice of representation. Many different measures which play this role qualitatively are available, and which of them one chooses depends to a large degree on matters of technical convenience which arise when one goes on to the case of a $2N$-dimensional Hilbert space (see next lecture). Purely for illustration, I quote here one possible definition of the "degree of purity" η:

$$\eta = 2\,\mathrm{Tr}\,\hat{\rho}^2 - 1 = |\langle \boldsymbol{\sigma} \rangle|^2 \tag{1.33}$$

It is easily checked that the value of η ranges between a maximum of 1 (realized for any pure state) and a minimum of 0 (realized for the "boring" case $\hat{\rho} = 1/2 \cdot \hat{1}$). For an arbitrary mixture expressed in the diagonal form (1.26) with $\rho_{11} = cos^2\theta/2$, etc., we have $\eta = cos^2\theta$, thus recovering the above two limits.

The General 2-State System

So far, we have assumed that the system we are trying to describe "really is" a physical particle of spin 1/2. However, it is important to emphasize that any system whose dynamics is effectively restricted to lie within a particular 2D Hilbert space can be put into a formal one-one correspondence with this system. For example, if we deal with a particle moving in a symmetric double-well potential and the temperature is much less than the energy of excited states within each potential separately, then the motion is effectively restricted to lie within the Hilbert space spanned by the two approximate ground states $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ in the left and right wells respectively, and it is conventional to regard these as the eigenstates of a "particle" of spin 1/2 corresponding to $\sigma_z = +1$ ($|\uparrow\rangle$) and $\sigma_z = 1$ ($|\downarrow\rangle$) respectively.[1] Then everything we have said so far goes through for this system, except that the interpretation of the symbols is somewhat different; for example, a "magnetic field" in the x-direction corresponds to an energy proportional to $\hat{\sigma}_x$, i.e. such that it tends to prefer a state which is a linear superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$ ($|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$) with equal weight. The physical realization of such an energy in the double-well problem corresponds to the splitting of the even and odd-parity combinations of $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ by tunneling through the barrier between them.

A second implementation, which I believe will be discussed in much more detail by Dr. Bouwmeester, is the polarization states of a photon. In that case we could associate the horizontal ($|\mathrm{H}\rangle$) and vertical ($|\mathrm{V}\rangle$) polarization states with $|\uparrow\rangle$ and $|\downarrow\rangle$; the eigenstates of $\hat{\sigma}_x$ then correspond to the "45°" linear polarization states, while the eigenstates of $\hat{\sigma}_y$, which in the spin-1/2 language are $2^{-1/2}(|\uparrow\rangle \pm i\,|\downarrow\rangle)$, correspond to the left- and right-circularly polarized states of the photon.

A point which is worth emphasizing is that in some cases when we deal with a general 2-state system the concept of "ensemble" may have to be generalized

[1] Warning: this choice of "axes" is not universal in the literature.

somewhat. When we are dealing with a real spin-1/2 particle, it is natural to think of the relevant "ensemble" as constituted by, say, the collection of atoms in a atomic beam, and this is indeed the sense in which, historically, the concept has been employed to interpret a wide variety of experiments. However, if our individual "2-state system" is actually a macroscopic physical system such as a SQUID ring (see lecture 4) then it is much more likely that we have only one such system and subject it to repeated experiments starting from the same initial conditions: the "ensemble" for which the quantum formalism makes predictions in this case is a "time" ensemble, that is the collection of experimental runs with the same physical object. Such a situation may obtain even when the individual system is microscopic, for example a single ion in a Penning trap.

2 Lecture 2 – Distinguishing Qubits from Cbits: Multi-qubit Systems and Quantum Computation

To motivate the topic of this lecture, let's stay for the moment with the single-qubit (spin-1/2) system discussed in the last lecture, and choose a specific representation (i.e. make a specific choice of the z-axis); note that in many of the realizations of a generalized 2-state system, there is a unique "natural" choice (e.g. in the case of a system moving between two potential wells separated by a barrier, the natural basis identifies the eigenstates of σ_z with the ground states $|\mathrm{L}\rangle$ and $|\mathrm{R}\rangle$ in the separate wells). Now we ask the question: How can we tell that our system really is a genuine qubit, rather than a classical "bit"? That is, how do we know that the possible states of the system include not only the two basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, but arbitrary linear combinations of the form $\alpha\ |\uparrow\rangle + \beta\ |\downarrow\rangle$ with $|\ \alpha\ |^2 + |\ \beta\ |^2 = 1$?

If we ask the question without qualification, the answer is obvious from the considerations of the last lecture: we need simply measure some component of the spin (or Pauli) vector other than the z-component, e.g., σ_x. If the system is indeed always in one of the states $|\uparrow\rangle$ or $|\downarrow\rangle$, or the ensemble is described by a classical probabilistic mixture of these as specified by the density matrix (1.26), then while individual measurements may yield values of σ_x different from zero the expectation value over the ensemble $\langle\sigma_x\rangle$, must always be zero. On the other hand, for a pure state with the parameters α and β specified by eqn. (1.17), the expectation value $\langle\sigma_x\rangle$ is clearly $n_x = \sin\ \theta \cos\ \varphi$.

However, in practice such a method of discrimination may not be realistic, since in many real-life experimental cases it may be difficult or impossible to measure $\langle\sigma_x\rangle$ directly: this is likely to be the case, for example, in the case of a two-well system. So let us impose the condition that the only measurements we are allowed to make are in the "classical" basis, i.e. of σ_z. Can we then distinguish a true "qubit" from a classical bit ("cbit")? The answer is yes, provided that the Hamiltonian contains terms which do not commute with σ_z; in this case the time correlations of the measurable property, σ_z, turn out under certain circumstances to have properties which those of a classical bit can never have.

A simple example is the following. Suppose the Hamiltonian is of the form $-\frac{1}{2}\sigma_x\mathcal{H}$, where $\mathcal{H}$ is a fictitious "magnetic field" in the x-direction which for example in the two-well problem might correspond to the splitting of the ground state due to tunneling. (The factor of 1/2 is inserted because apart from a factor of the magnetic moment, a physical magnetic field couples to the spin rather than the Pauli operator.) Classically, a "spin" described by this Hamiltonian would precess around the x-axis at angular frequency $\mathcal{H}$. What is the prediction of a quantum-mechanical calculation? First, when actually measured, the quantity σ_z will always have one of the values ± 1, never an intermediate value. Secondly, consider the correlation of the values of spin measured at different times, that is the quantity

$$\langle \sigma_z(t_2)\sigma_z(t_1)\rangle \equiv K(t_2,t_1) \qquad (t_2 > t_1 > 0) \tag{2.1}$$

This is measurable in a series of experimental runs, starting from identical initial conditions at $t = 0$, measuring σ_z at time t_1, then leaving the system undisturbed (this is crucial!) until t_2 when σ_z is again measured. On each run i a value $\eta_i = +1$ is assigned if the two measurements yield the same value of σ_z, and a value $\eta_i = -1$ if they are opposite: $K(t_2,t_1)$ is then the average of η over all the runs. The value of $K(t_2,t_1)$ predicted by a quantum-mechanical calculation is fairly obviously independent of the initial conditions at $t = 0$, since it can be expressed in terms of four conditional probabilities ($K(t_2,t_1) \equiv 1/2[P(1 \mid 1) + P(-1 \mid -1) - P(1 \mid -1) - P(-1 \mid 1)]$) each of which can be calculated simply from the consideration that immediately after a measurement at t_1 which yields (say) $+1$ the state of the system is "collapsed" on the state $|\uparrow\rangle$; this statement is completely independent of what the description was before the measurement. By integrating the time-dependent Schrödinger equation with the Hamiltonian $-\frac{1}{2}\sigma_x\mathcal{H}$, we find that at a time Δt later the state of the system is given by a superposition of the form (1.7), where the angles θ and φ parametrizing the coefficients α and β are given by $\varphi = \pi/2, \theta = \mathcal{H}\Delta t$. Consequently, the conditional probability $P(1 \mid 1)$ of finding $\sigma_z = +1$ at a later time $t_2 \equiv t_1 + \Delta t$, given that σ_z at time t_1 was measured to be $+1$, is given by

$$P(1 \mid 1) = cos^2\theta/2 = cos^2(\mathcal{H}\Delta t)/2. \tag{2.2}$$

By symmetry the quantity $P(-1 \mid -1)$ is given by the same expression, while the conditions $P(1 \mid 1) + P(-1 \mid 1) \equiv 1$ (etc.) give $P(-1 \mid 1) = \sin^2(\mathcal{H}\Delta t/2)$. Thus the quantity $K(t_2,t_1)$ is a function only of the difference $t_2 - t_1 (\equiv= \Delta t)$ and independent of the initial conditions on the ensemble: in fact we can write

$$K(t_j,t_i) = \cos\ \mathcal{H}(t_j - t_i) \qquad \text{for all } t_i, t_j. \tag{2.3}$$

At first sight the result (2.3) is in no way remarkable. What is surprising about it is that it cannot be obtained with a classical bit, that is, by the assumption that at any particular time the system is definitely in one or other of the two states $|\uparrow\rangle$ and $|\downarrow\rangle$ (with allowance made for very rapid "transits" between the two). The argument to this conclusion, which is almost a word-for-word transcription of the "CHSH" generalization of Bell's theorem, is extremely straightforward: Consider any four quantities A,B,C,D which each possess a definite value which is either +1 or −1. It is then a matter of simple exhaustion of the 16 possibilities to show that those quantities must satisfy the "CHSH" inequality

$$AB + BC + CD - AD \leq 2. \tag{2.4}$$

Now, if we believe that the "bit" is classical, then, with the above minor reservation which can be allowed for, the value of σ_z is defined at all times t_1 and takes one of the values ±1. Thus, let us choose four different times t_1, t_2, t_3, t_4 and define

$$\sigma_z(t_1) = A, \sigma_z(t_2) = B, \sigma_z(t_3) = C, \sigma_z(t_4) = D. \tag{2.5}$$

Then on any one run we must obviously have

$$\sigma_z(t_1)\sigma_z(t_2) + \sigma_z(t_2)\sigma_z(t_3) + \sigma_z(t_3)\sigma_z(t_4) - \sigma_z(t_1)\sigma_z(t_4) \leq 2 \tag{2.6}$$

Now imagine an ensemble of runs, and assume (this is a crucial step) that the measurement or not of σ_z at (e.g.) time t_1 does not "disturb" the ensemble, i.e. change the statistical predictions for the results of subsequent measurements. It should be emphasized that this assumption is quite false in a quantum-mechanical picture; however, I believe in a classical picture it is not only very natural but almost inescapable, provided we specify that the measurement is done in a suitably "noninvasive" way. Given this assumption (which I will call "noninvasive measurability"), the runs on which measurements are made at different times can be regarded as belonging to the same ensemble and the result (2.6) can be averaged over the ensemble, giving the result

$$K(t_2, t_1) + K(t_3, t_2) + K(t_4, t_3) - K(t_4, t_1) \leq 2 \tag{2.7}$$

But it is clear that the form of $K(t_j, t_i)$ given by (2.3) does not satisfy the inequality (2.7); for example, if we take $t_2 - t_1 = t_2 - t_2 = t_4 - t_3 = \pi/(4\mathcal{H})$, then the expression on the left-hand side of (2.7) is actually equal to $2\sqrt{2}$. Thus, even if we are allowed measurements only in the "classical" (σ_z-) basis, quantum spin-1/2 ensembles have at least one qualitative property which their classical counterparts do not.

This difference may be regarded as arising from the fundamental cbit/qubit difference: a classical system with two states available to it must at all times definitely be in one of these two states, and thus in information-theoretic terms can indeed only encode one "bit" of information. A quantum 2-state system, by contrast, is characterized by a complex number (α/β) which is the ratio of the amplitudes for the two states to be occupied.

Let us now move on to the case which is interesting in the context of quantum computing (and more generally "quantum information", a term which covers also e.g. quantum communication and cryptography), namely a set (or, more precisely, as usual, an ensemble of sets) of N 2-state systems which in general are mutually interacting. The difference expressed in the last paragraph now becomes much more spectacular: In the classical case, each 2-state system individually encodes one bit of information, so, trivially, the set of N systems encodes N bits. (The whole is exactly the sum of its parts.) In the quantum case, consider first the special case of a pure state which is a simple product of the wave functions of individual systems, i.e. such that

$$\psi(1, 2 \dots N) = \chi_1(1)\chi_2(2) \dots \chi_N(N) \tag{2.8}$$

where the χ_i's need not in general be identical. It is clear that such a wave function is parameterized by the N complex amplitudes α_i/β_i, so that the ratio of "information" in the quantum and classical cases, however we choose to define it, is no greater (and no less) than for a single system.

However, now comes the crunch: If we regard the collection of N quantum 2-state systems ("N qubits") as a single system, then the dimensionality of the relevant Hilbert space is 2^N. There is no general reason why we should not consider an arbitrary state vector in this space; thus, if we denote the states $|\uparrow\rangle_i$ and $|\downarrow\rangle_i$ respectively by values ± 1 of a dichotomic variable η_i, and $\psi\{\eta_i\}$ denotes the product state vector $| \eta_1\rangle \mid \eta_2 \dots \mid \eta_N\rangle$, we can consider the general state vector

$$\psi(1, 2 \dots N) = \sum_{\eta_i} C_{\{\eta_i\}} \psi\{\eta_i\} \tag{2.9}$$

where $C_{\eta_i} = C_{\eta_1 \eta_2 \dots \eta_N}$. Since each η_i has 2 different values, the total number of complex coefficients C_{η_i} is 2^N, a number which grows exponentially with N. It is this exponential growth of the information which can be encoded, by comparison with the linear growth in the classical case, which is what makes a "quantum computer" so enormously more efficient (at least in principle!) then its classical counterpart. In the general quantum case, the whole is indeed in a sense much greater than the sum of its parts!

I will now sketch very briefly the basic idea of a quantum computer, leaving a more sophisticated discussion to my fellow lecturers who are far more expert then I in this area; my main purpose here is to motivate the discussion of decoherence and related topics which will occupy the bulk of the rest of my lectures.

The term "quantum computer" describes a device (at present almost entirely on the drawing board, at least for "interesting" values of N!) which guides the N-qubit system through a series of states of the form (2.9), while maintaining the phase coherence expressed by the complex coefficients C_{η_i}; the motivation for this guiding process is of course to perform a computation, but there is no need to go into the details here of how this works. The condition that the phase coherence be preserved is crucial: if it is lost, then the system behaves just "as if" it were definitely in configuration $\{\eta_i\}$ with some probability $p\{\eta_i\}$, and even though this probability may not be a product of individual $P_i(\eta_i)$, i.e. there may be strong correlations between the classical states of the different bits, the computational power is now no greater than that of an old-fashioned classical N-bit system.

As we have seen, the number of complex coefficients required to characterize a product state, i.e. one of the form (2.6), is N, while that required for a general state of the form (2.7) is 2^N (more precisely, $2^N - 1$, since the overall normalization and arbitrariness of overall phase removes one complex degree of freedom).

Thus, for N appreciable, the vast majority of states of the N-qubit system are not product states; they are said to exhibit the phenomenon of "entanglement". A very well-known example of an "entangled" state for $N = 2$ is the state of 2 particles of spin corresponding to total spin 0 ("singlet" state):

$$\psi_{\text{sing}}(1,2) = \frac{1}{\sqrt{2}}(|\uparrow\rangle_1 \; |\downarrow\rangle_2 - \; |\downarrow\rangle_1 \; |\uparrow\rangle_2). \tag{2.10}$$

The role played by the state (2.10) and related ones in the discussion of "quantum nonlocality" will be discussed at this School by Dr. Bouwmeester.

A quantum computer may be regarded as involving, conceptually, three elements with different roles: the "system" (S), the "control" (C)[2] and the "environment" (E). The "system" is simply the collection of N qubits which will cycle through states of the form (2.9). The "control" indicates schematically the external device employed by the experimenter to drive the system through this cycle (we include in the "control" any device necessary to make a measurement on the final state, and/or to prepare the initial state). The "environment" denotes all those other components of the Universe with which the system may interact in the course of its evolution. The system and the control are of course essential and hence "wanted" ingredients in the problem; the environment is an unwanted but unavoidable complication. Let us briefly discuss the nature of the control and the environment.

The simplest form of "control" is just a set of time-dependent magnetic fields (real or "pseudo") which act independently on the different qubits; this corresponds to a Hamiltonian of the simple form

$$\hat{H}(t) = -\frac{1}{2}\sum_{i=1}^{N} \boldsymbol{\mathcal{H}}_i(t) \cdot \boldsymbol{\sigma}_i \tag{2.11}$$

where for the moment we assume that the quantities $\boldsymbol{\mathcal{H}}_i(t)$ can be treated as classical and under the control of the experimentalist (but cf. next lecture). Actually, a Hamiltonian of the form (2.11) is not very interesting, since it is easy to convince oneself that if the initial state of the N-qubit system is of the product form (2.11) then evolution under the influence of (2.9) will preserve the product nature of the wave function.

A more interesting form of Hamiltonian is

$$\hat{H}(t) = -\sum_{i,j=1}^{N} J_{ij}^{\alpha\beta}(t)\sigma_{i\alpha}\sigma_{j\beta} \tag{2.12}$$

(where i, j label the different qubits and α, β Cartesian components). Again, for the moment we assume that the quantities $J_{ij}^{\alpha\beta}(t)$ may be treated as c-numbers

[2] As far as I know there is no standard terminology in the literature for what I call the "control"; it should be carefully distinguished from the "control bit" which is a qubit which plays a particular role in computation (see e.g. the lectures of Dr. Ekert).

which, at least to an extent, are under the control of the experimenter. (In practice, J is likely to have some contribution which is intrinsic to the system and fixed in time, plus another component which may be controlled by turning various "knobs" – see for example Daniel Loss' lectures in this School.) Obviously, terms which couple three or more qubits are also possible, but in practice they tend to be harder to generate, and a Hamiltonian of the form (2.12) (possibly plus (2.11)) is usually enough to generate a degree of entanglement sufficient to be interesting in the context of quantum computing. A familiar special case of (2.12) is the Heisenberg Hamiltonian, defined by the choice

$$J_{ij}^{\alpha\beta} = J_{ij}\delta_{\alpha\beta} \tag{2.13}$$

It is straightforward to verify that such a Hamiltonian will evolve (e.g.) the state $|\uparrow\rangle_1 \, |\downarrow\rangle_2$ into an entangled state of the form $\alpha \, |\uparrow\rangle_1 \, |\downarrow\rangle_2 + \beta \, |\uparrow\rangle_1 \, |\downarrow\rangle_2$.

Now I turn to the "environment". As implied above, this term is effectively a shorthand for "all unwanted interactions with the system", that is, interactions with just about anything else in the Universe except for the control. Typical examples of an "environment" would be, in the case of nuclear spins, the blackbody radiation field or the phonons of the solid in question; for an atom in a cavity it might be interaction with the conduction electrons in the cavity walls, for a SQUID ring the normal electrons (if any) in the ring, and so on: cf. lecture 4. In general, the "environment" must be described quantum-mechanically, although in certain, special cases (e.g. the 50 Hz rf background) a classical description may be a good approximation. Note that the "environment" need not be something which is physically external to the system (cf. the SQUID example above); all the term really implies is one or more degrees of freedom other than the one(s) we are primarily interested in.

In order to perform a quantum computation, it is necessary to guide the N-qubit system very accurately through the 2^N-dimensional Hilbert space. Thus, any interactions which change the coefficients $c_{\{\eta_i\}}(t)$ in a way which the experimenter cannot know or cannot control are liable to spoil the operation. From this point of view it is sometimes natural to think of the environment as liable to contribute unknown extra terms to the quantities $\mathcal{H}_i(t)$ in eqn. (2.11), or to the quantities $J_{\alpha\beta}^{ij}(t)$ in (2.12), and so on. (I say "unknown" rather than "random"; even if one is confident that the extra terms are characteristic of the particular environment and thus in principle reproducible from run to run, this is unlikely to help very much in practice so long as one does not know their actual values[3].) In the case where the environment can be modelled by a (large-amplitude) classical field (as e.g. in the case of the 50 Hz background) the effect of such terms is easy to visualize intuitively; they have the same effect as an uncontrolled error in the "pulse length" of the fields $\mathcal{H}_i(t)$ (or coupling constants $J_{ij}^{\alpha\beta}(t)$, etc.) applied by the control. Although in a real-life experiment such effects may well be the dominant ones (cf. Lecture 4), they have not been very much discussed in the theoretical literature, probably because it is realized that the business

[3] However, it should be stressed that a failure to calculate these a priori does not hurt if we can instead obtain them from experiment: cf. Lecture 5.

of minimizing them is very much a practical problem to be solved by trial and error (e.g. in the case of the 50 Hz background, by appropriate rf shielding of the apparatus), rather than something to which theoreticians can contribute much deep insight.

What has been discussed in considerable detail in the theoretical literature, and is perceived as a major problem for the practical implementation of quantum computing, is the phenomenon known as "decoherence"; while this term is somewhat ambiguous (and not used uniformly in the literature), it is crudely speaking a shorthand for the unwanted effects of interaction with an environment which must be described quantum-mechanically. I will introduce and discuss this idea in the first part of the next lecture.

3 Lecture 3 – The Fundamental Theorem of Decoherence: "Isolated" Systems

Consider two quantum systems in interaction; for simplicity let one of them (1) be a 2-state system (qubit) while the nature of the other (2) is arbitrary, and denote the two chosen basis states of the qubit by $|\uparrow\rangle_1$ and $|\downarrow\rangle_1$ as usual. (This most common application is when system 2 is the "environment", but for generality I will not assume this at this stage.) Let us do a thought experiment in which the interaction is initially switched off; moreover, assume (as is natural under this condition) that the initial description of the two systems is factorizable, e.g.

$$\Psi(1,2) =|\uparrow\rangle_1 \chi(2) \tag{3.1}$$

where $\chi(2)$ is some normalized state of system 2 which we need not specify in detail. Now imagine that by a sequence of appropriate control pulses we prepare system 1 in a linear superposition of the states $|\uparrow\rangle$ and $|\downarrow\rangle$ with complex amplitudes α and β, without affecting the state of system 2:

$$\Psi(1,2) = (\alpha\,|\uparrow\rangle_1 + \beta\,|\downarrow\rangle_1)\chi(2) \tag{3.2}$$

We can verify that the state of 1 is indeed a linear superposition and not a mixture by measuring the expectation value of (say) σ_x; from the results of Lecture 1 we have[4]

$$\langle\sigma_x\rangle_{\text{sup}} = 2\,\mathrm{Re}\,\alpha^*\beta = \sin\ \theta\cos\ \varphi \tag{3.3}$$

whereas for a classical mixture of $|\uparrow\rangle$ and $|\downarrow\rangle$ we would have obtained simply

$$\langle\sigma_x\rangle_{\text{mixt}} = 0 \tag{3.4}$$

Now consider the effects of switching on the interaction between 1 and 2. In the general case the state (3.2) is not an eigenstate of the interaction Hamiltonian, so the resulting state vector will be time-dependent and its detailed form

[4] In the special case cos $\varphi = 0$ it is convenient instead to measure $\langle\sigma_y\rangle(=\sin\ \theta\sin\ \varphi)$.

will of course depend on the precise form of $\hat{H}_{\text{int}}$. However, the crucial point is that except under very special circumstances it will not remain a product, i.e. the interaction will generate entanglement between systems 1 and 2. Generically, the resulting state can be written in the form

$$\Psi(1,2:t) = \alpha(t)\chi_{\uparrow}(2:t)\;|\uparrow\rangle + \beta(t)\chi_{\downarrow}(2:t)\;|\downarrow\rangle \tag{3.5}$$

where the states $\chi_{\uparrow,\downarrow}$ of system 2 are normalized but have no special relation to one another (in particular, in the generic case they are neither identical nor mutually orthogonal). In the following I shall omit the parametric dependence of the various quantities on time.

The "fundamental theorem of decoherence" is now the following: Suppose that we measure, on the coupled system described by the state (3.5), the expectation value of (say) σ_x, the x-component of spin of system 1. What does this mean, formally, in terms of the coupled state vector (3.5)? We have to evaluate the expectation value of the appropriate operator on the combined system, and that operator is $\hat{\sigma}_x$ as regards system 1, multiplied by the unit operator as regards system 2. Thus,

$$\begin{aligned}\langle\sigma_x\rangle &\equiv \langle(\hat{\sigma}_x)_1 \otimes (\hat{1})_2\rangle \qquad (3.6)\\ &\equiv \int \mathrm{d}2\{\alpha^*\chi^*_{\uparrow}(2)\langle\uparrow| \;+\beta^*\chi_{\downarrow}(2)\langle\downarrow||\;\hat{\sigma}_x\;|\;\alpha\chi_{\uparrow}(2)\;|\uparrow\rangle + \beta\chi_{\downarrow}(2)\;|\downarrow\rangle\}\end{aligned}$$

and since $\langle\uparrow|\;\hat{\sigma}_x\;|\uparrow\rangle = \langle\downarrow|\;\hat{\sigma}_x\;|\downarrow\rangle = 0$, $\langle\uparrow|\;\sigma_x\;|\downarrow\rangle = \langle\downarrow|\;\sigma_x\;|\uparrow\rangle = 1$, this becomes

$$\begin{aligned}\langle\sigma_x\rangle &= \int \mathrm{d}2\{\alpha^*\beta\chi^*_{\uparrow}(2)\chi_{\downarrow}(2) + \alpha\beta^*\chi^*_{\downarrow}(2)\chi_{\uparrow}(2)\} \qquad (3.7)\\ &\equiv 2\,\mathrm{Re}\,\{\alpha^*\beta(\chi_{\uparrow}(2), \chi_{\downarrow}(2)\}\end{aligned}$$

(where (ψ, φ) denotes the usual scalar product in Hilbert space). In particular, if the states $\chi_{\uparrow}(2)$,$\chi_{\downarrow}(2)$ happen to be mutually orthogonal then we have simply

$$\langle\sigma_x\rangle = 0 \tag{3.8}$$

Similarly under these conditions we have

$$\langle\sigma_y\rangle = 2\,\mathrm{Im}\,\{\alpha^*\beta(\chi_{\uparrow}(2), \chi_{\downarrow}(2)\} = 0 \tag{3.9}$$

and since any operator on system 1 can be expressed in the form $\Omega_0\hat{1} + \boldsymbol{\Omega}\cdot\hat{\boldsymbol{\sigma}}$, and the expectation values of $\hat{1}$ and $\hat{\sigma}_z$ do not distinguish the pure state from a mixture, it follows that the state of system 1 is under these conditions completely equivalent to a classical mixture of $|\uparrow\rangle_1$ and $|\downarrow\rangle_1$ with probability $|\;\alpha\;|^2$ and $|\;\beta\;|^2$ respectively. Indeed, this result can be seen from an explicit calculation of the reduced density matrix of system 1: since in general we have

$$\rho_1(1,1') = \int \mathrm{d}2\;\Psi^*(1,2)\Psi(1',2) \tag{3.10}$$

we find that for a coupled wave function $\Psi(1,2)$ of the form (3.5) with $(\chi_\uparrow, \chi_\downarrow) = 0$ we have

$$\rho_1 = \begin{pmatrix} \mid \alpha \mid^2 & 0 \\ 0 & \mid \beta \mid^2 \end{pmatrix} \tag{3.11}$$

which is indeed identical to the density matrix of a system which is in one or other of the states $|\uparrow\rangle$, $|\downarrow\rangle$ with probability $\mid \alpha \mid^2$, $\mid \beta \mid^2$ respectively. (This corresponds to interpretation (3) of the density matrix in point (5) of Lecture 1.)

This result for the 2-state system is of course a special case of a much more general theorem, which may be stated in words as follows: As soon as any pair of (mutually orthogonal) states i and j of the system of interest become correlated to mutually orthogonal states of another system, all effects of phase coherence between them become lost, so that as regards any effects which depend only on these two states the situation is equivalent to a classical mixture (technically, the relevant element ρ_{ij} of the reduced density matrix of the system of interest is zero). This is what I call the "fundamental theorem of decoherence"; it has a very long history, and has played an essential role not only in discussions of quantum computing but also of the quantum measurement problem (see Lecture 6). Several comments and cautions are in order.

First, while coherence is transitive, decoherence is not. That is, if states i and j are mutually coherent (have a definite phase relation) and similarly j and k, then it follows that i and k are mutually coherent; on the other hand, it is perfectly possible for i to be decoherent with j and j with k while i and k remain mutually coherent.

Secondly, it is clear that the effects of "decoherence" due to interaction with a quantum environment are qualitatively very similar to the "dephasing" effects of a suitable classical random field. Suppose for example that we apply to the 2-state system in question a "random" field $\mathcal{H}(t)$ in the x-direction. The effect is to cause a precession, over a time interval Δt, of the relative phase of the coefficients β and α by an angle $\int_0^{\Delta t} \mathcal{H}(t) \mathrm{d}t$. Since the value of $\mathcal{H}(t)$ is random, in order to establish the statistical behavior of the ensemble we have to average over this angle with the relevant probability distribution for $\mathcal{H}(t)$. That is, the quantity $\alpha^* \beta$ should be replaced by

$$\overline{\alpha^* \beta} \equiv \int \mathcal{D}\mathcal{H}(t) P[\mathcal{H}(t)] \alpha_0^* \beta_0 \exp -\mathrm{i} \int_0^{\Delta t} \mathcal{H}(t) \mathrm{d}t \tag{3.12}$$

where α_0, β_0 denote the values in the absence of the fluctuating field. Evidently, with any "reasonable" (e.g. Gaussian) statistics for $\mathcal{H}(t)$, the quantity $\overline{\alpha^* \beta}$ decreases rapidly as soon as $\langle \mathcal{H}^2 \rangle \Delta t^2 \gtrsim 1$, so that the reduced density matrix of the system becomes diagonal to a very good approximation; thus the effect is indistinguishable from that of correlation to (nearly) orthogonal states of a second system. Indeed, there is no way that by making measurements at the time in question on the system alone we could distinguish (quantum) decoherence from classical dephasing. See however below.

Thirdly, it is absolutely crucial to appreciate that the mere fact that at some time t two states are effectively "decohered", that is that in the relevant representation the off-diagonal elements $\rho_{12} = \rho_{21}$ of the reduced system density matrix are $\ll 1$, by no means implies that no effects of phase coherence between 1 and 2 can ever be seen at any subsequent time. (Indeed, if this were the case then the whole idea of a quantum computer would be a nonsense, since at intermediate times during the computation the relevant density matrices of the individual qubits are typically nearly or totally diagonal!) The reason is quite a subtle one: in many cases of practical interest, even though a high degree of entanglement of our system with a second system occurs, that entanglement is often "adiabatic" in nature and thus reversible. To illustrate this point, consider the example of a 2-state system interacting with a harmonic oscillator of mass m and frequenty ω_0, in such a way that the parts of the Hamiltonian which involve the system are

$$\hat{H}_S = \frac{1}{2}\varDelta\hat{\sigma}_x + g\hat{X}\hat{\sigma}_z \tag{3.13}$$

where X is the coordinate operator of the oscillator; we assume that $\varDelta \ll \hbar\omega_0$, and moreover that the condition $g \gg (m\hbar\omega_0^3)^{1/2}$ is satisfied, so that in the absence of the "tunneling" term $\varDelta$ the two degenerate ground states would correspond to $|\uparrow\rangle\chi_\uparrow(X)$ and $|\downarrow\rangle\chi_\downarrow(X)$ with the centers of the wave packets $\chi_\uparrow(\mathrm{X})$ and $\chi_\downarrow(X)$ separated by a distance large compared to the zero-point width, so that their overlap is small. Suppose we consider the dynamics of this system in the presence of a finite tunneling matrix element $\varDelta$, starting from the configuration $|\uparrow\rangle\chi_\uparrow(X)$. It is intuitively clear (and can of course be justified more quantitatively) that under the stated conditions the coupled system still effectively forms a 2-state system with the (coupled) basis $|\uparrow\rangle\chi_\uparrow(X)$, $|\downarrow\rangle\chi_\downarrow(X)$; however, the effective tunneling matrix element is not $\varDelta$ but rather $\varDelta$ times the overlap (Franck-Condon factor) $\langle\chi_\uparrow \mid \chi_\downarrow\rangle \equiv \epsilon$, which under the stated conditions is $\ll 1$. Let us set $\varDelta \cdot \epsilon \equiv \varDelta_{\text{eff}}$; then apart from an irrelevant term proportional to the unit matrix the Hamiltonian in the coupled basis is simply

$$\hat{H} = \begin{pmatrix} 0 & \varDelta_{\text{eff}} \\ \varDelta_{\text{eff}} & 0 \end{pmatrix} \tag{3.14}$$

Evolving the coupled system from the initial state $|\uparrow\rangle\chi_\uparrow(\mathrm{X})$ according to the Hamiltonian (3.14), we find that at a time $t = \pi/2\varDelta_{\text{eff}}$ the wave function of the combined system will be strongly entangled:

$$\varPsi(1,2) = \frac{1}{\sqrt{2}}|\uparrow\rangle\chi_\uparrow(X) + \mathrm{i}\,|\downarrow\rangle\chi_\downarrow(X) \tag{3.15}$$

(corresponding to a rotation of the "spin" through $\pi/2$ around the x-axis). It is clear that the reduced density matrix of the 2-state system is

$$\hat{\rho}_1 = \frac{1}{2}\begin{pmatrix} 1 & +\mathrm{i}\epsilon \\ -\mathrm{i}\epsilon & 1 \end{pmatrix} \tag{3.16}$$

and since $\epsilon \ll 1$, this means that (at least at first sight!) that $|\uparrow\rangle$ and $|\downarrow\rangle$ states are almost totally decohered. However, it is clear that if we wait for a further quarter-cycle $\pi/(2\varDelta_{\text{eff}}\ )$ the wave function will return to a product form, i $|\downarrow\rangle\chi_\downarrow(X)$. More generally, the correlation $K(t_2, t_1)$ of σ_z predicted by a quantum-mechanical calculation will be exactly of the form (2) except that $\mathcal{H}$ (i.e. Δ) is replaced by Δ_{eff}:

$$K(t_2, t_1) = \cos\ \Delta_{\mathrm{eff}}\ (t_2 - t_1) \tag{3.17}$$

and thus cannot be mimicked by any classical system (see lecture 2). In other words, the effect of interaction with the oscillator has been to lengthen the time scale of the oscillation between the states $|\uparrow\rangle$ and $|\downarrow\rangle$ enormously but to leave its qualitative nature completely unchanged.

More generally, the inference from the (near)-vanishing at time t_0 of the off-diagonal elements of the single-particle density matrix $\hat{\rho}_1$ to the absence of any effects of interference between the two basis states at a later time is in general invalid. It is justified only in two special cases: (1) there is no term in the Hamiltonian for times later than t_0 which couples the two systems (b) there is an interaction term, but the nature of the two states χ which are involved in the entanglement is such that it is ineffective in restoring coherence. An example of case (a) is the operation of a quantum computer when the control couples the various qubits together with a time-dependent interaction which is nonzero before time t_0 (thus inducing the relevant entanglement) but vanishes thereafter. Case (b) is typified by an electron which is directed through a Young's slits device where the screen in the neighborhood of one slit is strongly conducting and in the neighborhood of the other insulating, so that in the first case the electron is almost certain to radiate a plasma excitation (plasmon) into the metal and in the other will do nothing. In this case, the plasmon effectively "goes away to infinity" so that the overlap of the states of the screen corresponding to passage through the L(R) slit is not just small but effectively zero.

The distinction between entanglement which is adiabatic in nature and thus reversible, and that which is as it were "permanent", is sometimes expressed as the difference between "false" and "true" decoherence; note once again that inspection of the form of $\hat{\rho}_1$ at time t does not allow us to distinguish these possibilities, we need at least a minimum of information about the states $\chi_\uparrow, \chi_\downarrow$ and also about the nature of the interaction at times subsequent to t. In the context of the parallel between decoherence and classical dephasing, "false" decoherence may be regarded as the (much more ubiquitous) analog of the classical spin-echo phenomenon, where although the field $\mathcal{H}(t)$ is in some sense random, its values for $t > \ \Delta t$ and $t < \ \Delta t$ are strongly correlated.

I now turn to a fundamental but surprisingly rarely discussed question, namely: Why is the concept of an "isolated system" in quantum mechanics ever valid at all? After all, every physical system in nature, however microscopic, is always interacting with other parts of the physical world, i.e. an "environment", be it only the blackbody radiation field, and moreover in the case of the systems of interest in the present context they must also interact strongly with the "control". So how come that it is ever possible to treat them as described by a wave function in their own right - will they not immediately get so entangled with the environment and/or the control that this approximation becomes quite invalid? From now on, when discussing the case of a quantum computer, I shall treat the whole N-qubit system rather than the individual bits as the "system" and discuss the question of how far that system as a whole does or does not get entangled with anything else in the universe. (As already pointed out, the whole point of a quantum computer is to entangle the individual bits with one another!) Note that once we go beyond a single qubit, not only the pseudo magnetic fields felt by the individual qubits, but also the strength of their interactions, may be operators with respect to the environment variables, and thus lead to decoherence.

As regards the control, the reason why there is no (serious) problem of its entanglement with the system is relatively well known: In a typical real-life application, when the system is say a set of N nuclear spins and thus the "field" $\boldsymbol{\mathcal{H}}(t)$ is literally a magnetic field, this is typically produced by the currents in a macroscopic object such as an rf coil, and thus the initial state of the (relevant degrees of freedom of the) radiation field is to a very high degree of approximation a coherent state, and moreover one of very large amplitude (mean number of photons $\gg 1$.) Now, it is a characteristic of a large amplitude coherent state that it is very nearly an eigenstate of the field operator $\hat{\boldsymbol{\mathcal{H}}}(t)$; thus the action of the relevant term in the Hamiltonian, namely - $\frac{1}{2}\sum_i \hat{\boldsymbol{\sigma}}_i \cdot \hat{\boldsymbol{\mathcal{H}}}_i(t)$, certainly in general results in a change of state of the system but in negligible change of the state of the radiation field, even though the latter may have exchanged quantities such as angular momentum and energy with the system. (A spectacular illustration of this state of affairs is the neutron interferometry experiment of ref. [2], in which the rf field flips the neutron spin in one beam; were this flip to induce any substantial degree of system-control entanglement, the observed interference of the two beams would not occur.) Needless to say, it is only this circumstance which justifies us in treating the electromagnetic field put out by (say) a radio antenna as a classical object.

A more delicate question arises concerning the interaction of our system with the ever-present quantum-mechanical "environment". Consider first a case which we are tempted to think of as the paradigm of a "free" particle, namely a neutron passing through an interferometer. In face, the neutron is interacting, because of its finite magnetic moment, with the radiation field, and as a result gets entangled with it: thus, at the intermediate stage the correct description is strictly speaking not the pure state of the neutron usually assumed, namely $2^{-1/2}(\psi_L(x)+\psi_R(x))$ where ψ_L and ψ_R are localized near the $L(R)$ "slit" respectively, but rather (schematically: I neglect complications connected with the previous history of the radiation field)

$$\Psi(x : E) = 2^{-1/2}(\psi_L(x)\chi_L(E) + \psi_R(x)\chi_R(E)) \qquad (3.18)$$

where χ_L and χ_R are the relevant states of the radiation field ("environment"). While the actual value of the overlap $(\chi_L \mid \chi_R)$ depends on the way the high-energy cutoff is treated, and is in general nonzero, it is certainly not unity, and thus a naive argument based on the form of $\hat{\rho}_1$ at this stage would lead one to expect a fringe visibility substantially reduced from 100%, in disagreement with experiment.

In the light of the above discussion, the resolution in this case is rather obvious; this is a typical example of "false" decoherence, and as the two neutron beams reconverge to the final screen where they interfere, the associated states of the radiation field as it were reconverge with them. As to the effect of the interaction in changing the parameters of the neutron dynamics, e.g. the "effective" neutron mass, this effect is present everywhere and is precisely what is calculated in the classic work of the 40's and 50's on renormalization of quantum electrodynamics; so it is already buried in the "experimental" mass of the neutron, and no further correction is necessary, at least at the kind of level in which we are interested in the present context.

Now let's consider some of the systems more directly relevant to quantum computing, such as the photons in an optical cavity or the Cooper pairs in a "Coulomb box" device (these will be described in more detail in the next lecture). Are these systems really isolated? A surprisingly widespread misconception is that they are not. In the case of the photons, the argument runs as follows: In order to be confined in the cavity at all, a photon must interact strongly with the atoms or conduction electrons in the cavity walls, and as a result its state will become strongly entangled with theirs. In the case of the Cooper pairs, it is tempting to argue that as a single pair tunnels, the rest of the electrons within the "receiving" box recoil, and as a result the states of these other electrons corresponding to the Cooper pair being in the $L(R)$ box are nearly orthogonal, i.e. we have a strong degree of entanglement.

These arguments are quite wrong: a quantitative calculation (which I hope to present elsewhere) shows that e.g. in the case of the photon, even though the interaction with the conduction electrons in the walls is from the point of view of the photon very "strong" (in that it confines it entirely within the cavity) from the point of view of the electrons it is so weak that the "environmental" (wall) states corresponding to the presence/absence of the photon have an overlap of very close to 100%. The same is true of the states of the "environment" (other electrons) in the Cooper pair case. Indeed, were this not so it would be hard to understand why the "naive" application of quantum mechanics to these systems seems to work so well. To paraphrase the proverb about walking like a duck, if it behaves like an isolated system, it probably <u>is</u> an isolated system, at least to a very high degree of approximation!

4 Lecture 4 – Candidate Systems for Qubits: Preliminary Estimate of Decoherence Times

I will start by reviewing briefly a few of the systems which are currently seriously considered as candidates for the ingredients in a quantum computer (qubits). In some cases, rudimentary quantum computation has already been performed using these systems, while in others even this first step remains in the future. Some of these systems will be discussed in much more depth by other lecturers in this School; my primary purpose here is to motivate the ensuing discussion of decoherence, so I shall pay particular attention to the possible mechanisms for this.

(1) **Quantum-optical systems**. This includes the photons trapped in an optical cavity (note that anything which is going to play the role of a qubit in a practical computer must presumably be confined to a relatively small region of real space!) and also any atoms which may be subjected to a controlled interaction with them. If we consider first the photons themselves, then the two states involved can be taken as corresponding to orthogonal polarization directions; the polarization can be manipulated, in principle, by "control" devices such as Kerr cells which are switchable and can rotate the polarization of a photon incident on them. The principal mechanism of decoherence for the photons is likely to be dissipation in the cavity walls, which have a large but still finite conductivity at optical frequencies.

A very elegant series of experiments [3] by the Haroche group at the ENS in Paris has shown that it is possible to produce an interesting degree of entanglement between photons and atoms which are passed through the relevant cavity in a controlled way; furthermore, this process can be used also to entangle photons or atoms with one another. In the case of the atoms the two basic states of the qubit are the ground state and a particular excited state to which it is coupled by the relevant photon; that these two states have very different energies is of minor importance, since one can work in the so-called rotating frame in which this is taken into account. For the atoms, the principal mechanism of decoherence is likely to be the quantized radiation field, which causes spontaneous transitions from the excited state back to the ground state and thus destroys any superposition.[5] At higher densities atom-atom collisions may play a role, but this is unlikely to be significant under the conditions required for a practical quantum computer.

(2) **Ions in Penning traps.** It is possible to keep a single ion in a trap formed by electric and magnetic fields for long periods. One can then produce superpositions of ground and excited states as in (1), and furthermore entangle them with the spatial motion of the ion in the trap. If several such ions are stored in neighboring traps, it is possible to use their strong Coulomb interaction to entangle their spatial degrees of freedom and thus indirectly their internal ones. Obviously, a major source of decoherence for an ion is likely to be random fluctuating electric fields, produced e.g. by the walls of the experimental cell; unwanted vibrational degrees of freedom may also be a problem.

(3) **Nuclear spins in solids or liquids.** This is a rather different kind of case from the above ones, in the sense that the "ensemble" is of the traditional kind rather than a "time" ensemble, i.e. one measures simultaneously the properties of a collection of independent "quantum computers", each of which itself consists of a number N of qubits (nuclear spins). (In the experiments actually conducted to date, N has always been a single-digit number.) The principal

[5] It is not immediately obvious that this process fits into the scheme outlined in Lecture 3. To discuss this question would require more technical detail then I want to go into here.

mechanisms of decoherence in this case are probably random fluctuating magnetic fields produced by the physical environment, and possibly also phonons (which modulate the spin-spin interaction in an unwanted and uncontrollable way).

(4) **Electron spins in solids.** The general features of this system (which is discussed in detail by Daniel Loss at this School) are similar to those of the nuclear spin system, the principal difference being that the magnetic moment, i.e. the coupling to external magnetic fields, is a factor 2000 larger (and the coupling of the qubits to one another, ceteris paribus, is about a factor of 4×10^6 larger). Again, the principal decoherence mechanisms are likely to be random magnetic fields and/or phonons, with strong electric fields also possibly playing a role (e.g. by modulating the shape of the electron wave packet and hence the strength of the electron-electron coupling).

(5) **Single electrons, or Cooper pairs, tunneling** between two (or more) "quantum dots" (that is, boxes holding a mesoscopic ($\sim 10^9$) number of conduction electrons). In this case the "2-state system" is formed by the positional variable of the electron (or pair), which schematically takes two values corresponding to being in the $L(R)$ box respectively. In practice, one usually tunes the potential difference between the two boxes so that there are only two relevant many-body states, namely those corresponding (1) to $N+1$ electrons on the left and $N_{\text{tot}} - (N+1)$ on the right ("$\mid L\rangle$") and (2) to N on the left and $N_{\text{tot}} - N$ on the right ("$\mid R\rangle$"); other states, corresponding to the transfer of more than one electron (pair) are energetically disfavored by the capacitance energy, which is of order e^2/C and for reasonably attainable values of the capacitance C can be made much larger than kT under accessible cryogenic conditions.

One obvious possible mechanism for decoherence in this case is the interaction with the $N-1(\sim 10^9)$ electrons which are not involved in the hopping from L to R, but readjust within the individual boxes; however, as we have seen at the end of lecture 3, this interaction is surprisingly weak, and in existing experiments it appears that the primary mechanism is dissipation in the external leads (which in these experiments are inevitable if one wants to actually measure anything!)

(6) **SQUID rings.** The simplest system of this type is a bulk superconducting ring interrupted by a single Josephson junction ("rf SQUID"). This system has been of great interest not only in the context of quantum computing (where its usefulness is still somewhat speculative) but also, long before this became a topic of urgent interest, in the context of tests of quantum mechanics at the macroscopic level, a topic which I shall address in lecture 6. For present purposes, it is sufficient to note that the two states which form the basis of the qubit correspond to currents circulating of the order of 1 μA respectively clockwise and anticlockwise; thus, in strong contrast to the other candidates considered above, these two states are by a reasonable criterion macroscopically distinct (it is of course just this feature which endows the SQUID system with such interest in the context of the quantum measurement problem). In this case the possible mechanisms of decoherence are many (which may be why the whole idea of looking for superpositions of different current (flux) states of a SQUID, when originally

proposed 20 years ago, initially evoked such a sceptical reaction from much of the "quantum measurement" community): normal electrons, phonons, nuclear spins and these of electronic impurities, the blackbody radiation field, even the "passing truck" (a shorthand for various more or less classical noise sources, such as the 50 $\boldsymbol{\mathcal{H}}z$ rf background, which while perhaps not strictly classifiable as giving rise to "decoherence" in the technical sense, have as we have seen in Lecture 2 much the same effect in practice). Actually, as we shall see below, it turns out that the decoherent effect of many of these mechanisms is very much less than one might have mainly guessed, and very recent experimental advances (see Lecture 6) have made it at least plausible that a SQUID ring may indeed be a viable candidate for a qubit.

Finally, it is worth pointing out that there are many other examples of 2-state systems interacting with complex environments, and thus liable to experience decoherence, which for one technical reason or another may not be good candidates for ingredients in a quantum computer but have been well studied in their own right; these range from hydrogen atoms tunneling between equivalent impurity sites in a solid (or electrons tunneling between equivalent sites in a large molecule while interacting with the vibronic degrees of freedom) to solar neutrinos oscillating between "electron" and "muon" states while interacting with the matter in the Sun. Thus, the problem of a two-state system interacting with a complex environment is one which occurs in virtually every sub-field of physics, and its formulation and solution is of very wide interest. I now turn to this question.

In the following we shall be concerned until further notice with the question of the mutual decoherence of two specific states of our "system", even though that system may itself be composed of $N(\gg 1)$ qubits and thus be described by a 2^N-dimensional Hilbert space. For example, we might be concerned with the question whether, in a given $N = 3$ (mini-) quantum computer, the two states $|\uparrow_1, \downarrow_2, \uparrow_3\rangle$ and $|\downarrow_1, \uparrow_2, \uparrow_3\rangle$ remain mutually coherent over a given time. It is therefore tempting to treat this special pair of states as itself effectively constituting a "qubit", and to truncate the Hamiltonian to the terms found in the corresponding 2×2 subspace of the Hilbert space. It is not clear that this is always an adequate approximation, but it is what is usually done in the discussion of decoherence in the literature, so we should use it with appropriate reservations. It is, of course, not necessarily implied that the c-number part of the "field" acting on the system is along the (notional) z-axis; in fact, in the case of a symmetric two-well potential the conventional (though not universal) choice is to take it along the x-axis, cf. lecture 1 and eqn. (4.4) below).

The next question concerns the form of the interaction between the "system" (qubit) and the environment. It is clear that an arbitrary interaction can be written in the generic form

$$\hat{H}_{\text{int}} = \hat{\Omega}_0 \hat{1} + \hat{\boldsymbol{\Omega}} \cdot \hat{\boldsymbol{\sigma}} \tag{4.1}$$

where the quantities $\Omega_i (i = 0, 1, 2, 3)$ are arbitrary Hermitian operators (not necessarily mutually commuting) on the Hilbert space of the environment. The term proportional to the unit matrix clearly has no effect on the dynamics of the system and can be safely neglected in what follows.

If the components of the vector $\hat{\boldsymbol{\Omega}}$ were just numbers, we could of course always choose the basis states of the system so that the interaction term becomes simply $\mid \boldsymbol{\Omega} \mid \hat{\sigma}_z$. Since the components are actually operators (matrices), there is no guarantee that we can do this, and indeed in (for example) the case of nuclear spin it is typically impossible to avoid having environmental "fields" which act on more than one component of $\boldsymbol{\sigma}$. However, in many cases of practical interest in the context of quantum computing it turns out that it is indeed possible to choose the basis so that the environment couples only to $\hat{\sigma}_z$, and moreover the basis so defined is in some sense the "natural" one. This is not entirely an accident: it is connected with the fact that one usually wants to choose the "natural" basis so that the states in question are fairly stable, meaning inter alia that the "intrinsic" transitions between them should be weak. One obvious way of guaranteeing that it is indeed weak is to arrange for the states in question to be connected only by tunneling through a potential barrier, so that they correspond to configurations which are substantially displaced relative to one another either in real space (as in the "quantum box" example) or in some abstract space (as in the example of the rf SQUID). Under these circumstances, any interaction with the environment which is to "flip" the system from one of the two basis states to the other has to involve many elementary "steps", i.e. many powers of the basic interaction, and is thus likely to be negligible.

This rather hand-waving argument can be made a bit more quantitative in the case when the qubit is obtained by truncation of a problem involving a continuous coordinate q with which is associated a simple canonical momentum p; this is the case not only where q represents the real Cartesian coordinate (as in the quantum-box case), but also in some cases (e.g. the rf SQUID) where q represents an abstract coordinate (in that case, the total magnetic flux through the ring). The argument is rather technical and I refer to refs. (6,8) for the details[6], but the crucial point is that in such a case, subject to a couple of very generic and plausible assumptions, the coupling to the environment can always be chosen to be a function only of the coordinate q (i.e. independent of the conjugate momentum p). It is then immediately obvious that when the extended problem (which typically would involve two degenerate or nearly degenerate potential wells separated by a potential barrier (cf. above), is truncated to a 2-state one, where the two states correspond to widely different values of q, the truncated coupling must be diagonal in that representation and thus must (apart from a term proportional to the unit matrix, which as we have seen is irrelevant to the dynamics) be proportional to $\hat{\sigma}_z$.

Essentially the same conclusion can be obtained in an explicit quantum computing context if we consider the two states in question (the eigenstates of σ_z) to be e.g. $|\uparrow\uparrow \ldots \uparrow\rangle$ and $|\downarrow\downarrow \ldots \downarrow\rangle$, with a total of N qubits. Suppose that each of the qubits is acted on by some random (classical or quantum) "noise" corresponding to a "magnetic field" in a random direction. Then the noise which modulates

[6] I anticipate at this point some results which will be discussed further in Lecture 5.

the energy of each of the states relative to the other (the x-component of the random "magnetic field") will be of order $N^{1/2}\epsilon$, while the amplitude for flipping all the qubits so as to induce a transition between the σ_z-eigenstates (the x-component of the "magnetic field") is of order ϵ^N and hence, for small ϵ and large N, completely negligible with respect to the x-component. This consideration is somewhat related to the question of quantum error correction, see the lectures by Cirac.

The upshot of the above rather turgid argument is that in many (though not all) cases of practical interest the interaction of the system with the environment can be expressed, in the "natural" basis, in the simple-form $\frac{1}{2}\hat{\Omega}\hat{\sigma}_z$ where $\hat{\Omega}$ is some Hermitian operator on the Hilbert space of the environment. I will assume until further notice that we deal with this case. Since an arbitrary Hamiltonian of the system by itself can be cast (up to an irrelevant constant) in the form $-\frac{1}{2}\boldsymbol{\mathcal{H}}(t)\cdot\hat{\boldsymbol{\sigma}}$ where the components of the vector $\boldsymbol{\mathcal{H}}(t)$ are real c-numbers, it follows that the most general case we need discuss is described by the Hamiltonian

$$\hat{H} = -\frac{1}{2}\boldsymbol{\mathcal{H}}(t)\cdot\hat{\boldsymbol{\sigma}} + \hat{\Omega}\hat{\sigma}_z + \hat{H}_{\mathrm{env}} \tag{4.2}$$

where $\hat{H}_{\mathrm{env}}$ is that part of the Hamiltonian of the environment which does not involve the system variables.

In the case where $\boldsymbol{\mathcal{H}}$ is time-independent (e.g. if we wish to study the decoherence of our "effective qubit" between the control pulses) we can actually simplify the Hamiltonian a little further, since by an appropriate choice of the relative phases of the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$ we can eliminate the term in $\mathcal{H}_y$. Then denoting $\mathcal{H}_x$ as Δ and $\mathcal{H}_z$ as ϵ, we have

$$\hat{H} = -\frac{1}{2}\Delta\hat{\sigma}_x - \frac{1}{2}\epsilon\hat{\sigma}_z + \hat{\Omega}\hat{\sigma}_z + \hat{H}_{\mathrm{env}} \tag{4.3}$$

Of course, it is always possible to perform a further rotation of axes so as to get rid of the term $-\epsilon\hat{\sigma}_z$, but this comes at the expense of a more complicated form for the interaction term and is usually not worthwhile (though cf. note (2) at end of this section.) In many cases of practical interest ϵ is zero (or at any rate $\ll \Delta$), and it is instructive to study this case as a guide to the qualitative features of decoherence in the more general problem. Thus in the rest of this lecture (and much of the next) we study the "canonical" problem defined by the $\epsilon = 0$, version of (4.3), namely

$$\hat{H} = -\frac{1}{2}\hat{\Delta}\hat{\sigma}_x + \Omega\hat{\sigma}_z + \hat{H}_{\mathrm{env}} \tag{4.4}$$

The question we wish to answer is: Given that we start off with the system in an arbitrary coherent superposition of the states $|\uparrow\rangle$ and $|\downarrow\rangle$ (for example, the state $2^{-1/2}(|\uparrow\rangle + \mathrm{i}|\downarrow\rangle))$ with the environment in some specified initial state (e.g. in thermodynamic equilibrium at temperature T), how long is it before this phase coherence is destroyed, i.e. the off diagonal elements of the reduced system density matrix relax to their equilibrium value (which, it should be noted, is not zero in the representation corresponding to the σ_z-eigenstates! (see below))? In the next lecture I will try to treat this question with some rigor; in the rest of this one I give a simple argument (to be found in many places in the literature) which I believe captures most (though not all) of the essential physics.

In the following, a special role will be played by the spectral density of the operator $\hat{\Omega}$: we define

$$J'(\omega : \beta) \equiv \mathrm{Z}^{-1} \sum_m e^{-\beta E_m} \sum_n \mid \Omega_{mn} \mid^2 \delta(\omega - (E_n - E_m)) - \delta(\omega - (E_m - E_n)) \tag{4.5}$$

where as usual $\beta \equiv 1/k_{\mathrm{B}}T$, Z is the environment partition function and we have set $\hbar = 1$. We note that J' is, up to a numerical factor, the imaginary part of the susceptibility (response function) of the operator $\hat{\Omega}$. Purely for pedagogical simplicity, I will assume in the following that J' is not a function of temperature, so that it can be written in the simpler form

$$J'(\omega) = \sum_n \mid \Omega_{0n} \mid^2 \delta(\omega - (E_n - E_0)) \tag{4.6}$$

This is in fact the case for many models of interest (cf. Lecture 5). If we define the quantities

$$P_{\pm}(\omega) \equiv 2\pi Z^{-1} \sum_m e^{-\beta E_m} \sum_n \mid \Omega_{mn} \mid^2 \{\delta(\omega \pm (E_n - E_m)\} \tag{4.7}$$

which for positive ω represent the probability of absorbing (+) energy $\hbar\omega$ from the environment or losing (−) energy $\hbar\omega$ to it (see below), then we have by the fluctuation-dissipation theorem

$$P_+(\omega) = 2\pi n(\omega) J'(\omega) \qquad P_-(\omega) = 2\pi(n(\omega) + 1) J'(\omega) \tag{4.8}$$

where $n(\omega) \equiv (\exp \beta\omega - 1)^{-1}$.

Let us now suppose that the system-environment interaction is very weak, and treat it by standard second-order perturbation theory. Consider first the equilibrium state of the system neglecting this interaction: then the energy eigenstates are the eigenstates of $\hat{\sigma}_x$ (not $\hat{\sigma}_z$!), namely $2^{-1/2}(|\uparrow\rangle + |\downarrow\rangle)$ (energy $-\Delta/2$) and $2^{-1/2}(|\uparrow\rangle - |\downarrow\rangle)$ (energy $+\Delta/2$) and have relative population $\exp \beta\Delta$. Thus, the equilibrium value of σ_x is

$$\langle \sigma_x \rangle_{\mathrm{eq}} = \tanh \ \beta\Delta/2 \tag{4.9}$$

while the equilibrium values of σ_y and σ_z are zero (since they are zero in each eigenstate of $\hat{\sigma}_x$). Now suppose, first, that while leaving the density matrix diagonal in the σ_x-representation we create a nonequilibrium population, and ask how it relaxes under the influence of the interaction with the environment. Since in second-order perturbation theory the rate of "down" transitions ($|\leftarrow\rangle \to |\rightarrow\rangle$) is $P_-(\Delta)$ and that of "up" transitions $P_+(\Delta)$, we easily find

$$\mathrm{d}\langle \sigma_x \rangle / \mathrm{d}t = -\frac{(\langle \sigma_x \rangle - \langle \sigma_x \rangle_{\mathrm{eq}})}{T_1} \tag{4.10}$$

where the relaxation rate T_1^{-1} is given by

$$T_1^{-1} = 2\pi \coth\,(\beta\Delta/2) J'(\Delta) \tag{4.11}$$

What of the relaxation of $\langle\sigma_y\rangle$ and $\langle\sigma_z\rangle$, which both have equilibrium value zero? The simplest argument is from symmetry: since the fluctuating "field" $\hat{\Omega}$ is always along the z-axis, it cannot relax $\langle\sigma_z\rangle$, and by symmetry its effect on $\langle\sigma_y\rangle$ should be the same as on $\langle\sigma_x\rangle$. However, even for the free system the spin precesses around the x-axis with frequency Δ. Thus we obtain

$$\frac{\mathrm{d}\langle\sigma_y\rangle}{\mathrm{d}t} = \Delta\langle\sigma_z\rangle - \frac{\langle\sigma_y\rangle}{T_2} \tag{4.12}$$

$$\frac{\mathrm{d}\langle\sigma_z\rangle}{\mathrm{d}t} = -\Delta\langle\sigma_y\rangle \tag{4.13}$$

with T_2 identical to T_1, i.e. T_2^{-1} is given by (4.11).

A second, somewhat intuitive way of obtaining the relaxation term for $\langle\sigma_y\rangle$, which makes contact with the general analysis of lecture 3, is to argue that the phase coherence between the σ_z-eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ is destroyed by any process in which one state as it were changes the environment while the other does not. (It is probably easier to follow the argument if one adds to the Hamiltonian a term of the form $\hat{\Omega}\hat{1}$, which as we have seen has no effect on the system dynamics but means that $|\uparrow\rangle$ tends to induce changes in the environment while $|\downarrow\rangle$ does not. Keeping track of the factors of 2, etc., is however quite delicate.) Since it is irrelevant whether energy is gained or lost in the process, the relevant probability is $P_+(\Delta) + P_-(\Delta) = 2\pi \coth \beta\Delta/2\ J'(\Delta)$, in agreement with eqn. (4.11).

In view of eqns. (4.10) and (4.12) (which are very reminiscent of the Bloch equations for nuclear magnetic resonance)[7], we may regard the time T_2 defined by (4.10) as the time for the states $|\uparrow\rangle$ and $|\downarrow\rangle$ to undergo mutual decoherence. It is now crucial to observe that the behavior of this quantity in the limit $\Delta \to 0$ is sensitive to the behavior of the environmental spectral density $J'(\omega)$: if this quantity behaves as ω^s where $s > 1$ ("superohmic" behavior) then the decoherence time tends to infinity as $\Delta \to 0$, whereas if $J'(\omega) \to \tilde{\alpha}\omega$ as $\omega \to 0$ ("ohmic" case) T_2 tends to the finite limit $(\tilde{\alpha}kT)^{-1}$. As one might expect, this last result can be obtained by regarding the environment as providing a classical random field where Gaussian fluctuations are given by the standard fluctuation-dissipation theorem; cf. eqn. (3.12).

The above analysis has shown that for the two-state system described by the Hamiltonian (4.4) the "decoherence time" T_2 which enters in eqn. (4.12) is identical to the "energy relaxation time" T_1 which occurs in eqn. (4.10) (since for the isolated system the mean energy is proportional to $\langle\sigma_x\rangle$). However, it should be emphasized that this simple result is not generic. For example, if we go through the argument of the last paragraph, but for a simple harmonic

[7] However, there are significant differences, associated inter alia with the fact that in the NMR problem the environment generally couples to all three components of $\boldsymbol{\sigma}$.

oscillator of frequency ω whose mean excitation energy is $\bar{n}\hbar\omega$, we find that the energy relaxation rate is proportional to the difference of the emission probability $(\bar{n}+1)n(\omega)$ and the absorption probability $\bar{n}(n(\omega)+1$ (where $n(\omega)$ is the thermal Bose occupation factor as above) while the decoherence rate is proportional to their sum. Since the mean energy is itself proportional to $\bar{n}$, the upshot is that the decoherence time is a factor of order $(\bar{n}n(\omega))^{-1}$ shorter than the energy relaxation time. (For details, see e.g. ref. [4]). Although quantitative calculations for systems other than the two-state system and the harmonic oscillator are hard to find in the literature, the qualitative result, that the decoherence time is shorter then the energy relaxation time, often by several orders of magnitude, is almost certainly more generally true for any system whose Hilbert space is of dimension $\gg 2$.

Note added in proof: It may be easier to make contact with the recent literature in this area if we note the following points:

(1) In the context of attempts to implement quantum computing, we are usually interested in the limit $\Delta T_2 \gg 1$. Explicit solution of eqns. (4.12-13) in this limit shows that the "average" rate of relaxation of the transverse (yz-plane) component of $\boldsymbol{\sigma}$ is equal to $1/(2T_2)$ rather than $1/T_2$, so one sometimes defines a "dephasing time" $T_\varphi \equiv 2T_2$, and it is this latter time which is most easily measured experimentally.

(2) If in addition to the coupling to the environment represented in eqn. (4.4), which is diagonal in $\hat{\sigma}_z$, there exists also a coupling diagonal in $\hat{\sigma}_x$ then it is clear that this coupling cannot affect T_1 but will in general give an additional contribution at T_2^{-1} which turns out, in the simplest case, to be proportional to $\lim_{\omega\to 0}\coth(\beta\hbar\omega/2)J''(\omega)$, where $J''(\omega)$ is an appropriately defined spectral density analogous to (4.6). In the literature this effect is sometimes called "pure dephasing" or "pure phase noise": it arises automatically when a system originally described by the symmetric Hamiltonian (4.4) is subjected to a bias ϵ as described by (4.3), since we can then if we wish redefine the axes so that the direction of the c-number "field" is still $\hat{x}$. Cf. e.g. ref. [5], section 4.

5 Lecture 5 – The Environment as a Bath of Oscillators: More Systematic Treatment of Decoherence

Although the simple calculation done at the end of the last lecture gives some feeling for the main physical features of decoherence, it is clearly inadequate as it stands, since, for example, it does not allow for the possibility of the adiabatic "following" of the system by the environment and the resulting radical renormalization of the system frequency ($\Delta \to \Delta_{\text{eff}}$) which was discussed in Lecture 3. It is desirable, therefore, to look for a more complete approach, and at least in many cases of physical interest it turns out that such is possible. I will not spend time here on the technical details of the derivation, which can be found in refs. [6-9], but will just summarize the main results. To state the central conclusion right away: In many cases of physical interest, not only can the problem of the interacting system and environment be cast into the "canonical" form of eqn. (4.4), but for the purposes of considering the system dynamics (only!) the environment can be represented as a set of simple harmonic oscillators, with the operator $\hat{\Omega}$ proportional to a linear combination of the oscillator coordinates. That is, the total Hamiltonian can be represented for our purposes in the form

$$\hat{H} = -\frac{1}{2}\Delta\hat{\sigma}_x - \frac{1}{2}\epsilon\hat{\sigma}_z + \frac{1}{2}q_0\hat{\sigma}_z \sum_\alpha C_\alpha \hat{x}_\alpha + \hat{H}_{\text{osc}}, \tag{5.1}$$

$$\hat{H}_{\text{osc}} = \sum_\alpha (\frac{1}{2}m_\alpha \omega_\alpha^2 \hat{x}_\alpha^2 + \frac{1}{2}\hat{p}_\alpha^2/m_\alpha) \tag{5.2}$$

where $\hat{x}_\alpha, \hat{p}_\alpha, m_\alpha$ and ω_α are respectively the coordinate and momentum operators, the mass and the frequency of the α-th oscillator of the "bath" (environment). The problem defined by the Hamiltonian (5.1) is known as the "spin-boson" problem and has been very extensively studied in the literature, both as a generic problem and in particular contexts: see, for example, ref. [8].

In some cases (e.g. an electron tunneling between two discrete sites in a large molecule, while interacting with the vibronic modes of the latter) eqn. (5.1) can be justified directly from a knowledge of the microscopic ingredients. However, it is important to realize that it can also be justified, under rather wide conditions, for a much more general class of problem, namely one in which the two fundamental states are attained by truncation of an extended variable of a certain common type. Examples of such a case which may be interesting in the context of quantum computing include the quantum box and the rf SQUID; the latter is also of great interest in the context of tests of the extrapolation of quantum mechanics to the macroscopic level, see Lecture 6.

In fact, consider a system which in isolation would be described by the Hamiltonian

$$\hat{H}_S = \hat{p}^2/2M + V(\hat{q}) \tag{5.3}$$

$$[\hat{p}, \hat{q}] = -i\hbar \tag{5.4}$$

and $V(q)$ is a potential which in the case of practical interest has two nearly degenerate local minima separated by a potential barrier. The variable q may (as in the quantum box) represent a real Cartesian coordinate, but it need not (for example, in the case of the rf SQUID it represents the total magnetic flux trapped through the ring, see ref. [6]). Note that systems where the principal variable of interest is of the nature of a spin do not, at least prima facie, conform to eqns. (5.2-3), so that the ensuing discussion would have to be generalized somewhat in order to apply to them.

Suppose, now, that the system interacts with an arbitrary environment through an interaction which is linear in p and q:

$$\hat{H}_{\text{int}} = \hat{q}\hat{\Omega}_1 + \hat{p}\hat{\Omega}_2 \tag{5.5}$$

where $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are arbitrary Hermitian operators in the Hilbert space of the environment, and suppose furthermore that the total Hamiltonian is invariant under time reversal T, so that $\hat{\Omega}_1$ is even under T and $\hat{\Omega}_2$ odd. These conditions are actually not at all stringent, and we can be fairly confident that they are met in most of the cases of practical interest[8]

Let us now assume that the environment is, in some sense, "macroscopic", or more precisely that it is described by a large number N of degrees of freedom. Then at first sight, at least, the coupling to the system of each degree of freedom will be of order $N^{-1/2}$, i.e. extremely weak in the limit of large N; more formally, if m, n characterizes two particular energy eigenstates of the (uncoupled) environment, the matrix elements $\Omega^{(1)}_{mn}$, $\Omega^{(2)}_{mn}$ should be much less then the energy difference $E_m - E_n$. Under this condition it may be shown (see refs. (6) and (7) for details) that it is possible, for the purposes of studying the classical or quantum dynamics of the system (only !) to represent the environment as a bath of single harmonic oscillators, and moreover by an appropriate choice of "gauge" (i.e. of the definition of the "position" and "momentum" of the α-th oscillator) to represent the coupling as bilinear in the system and oscillator coordinates. Thus the Hamiltonian can be written in the generic form

$$\hat{H} = p^2/2M^* + V(q) + q\sum_\alpha \tilde{C}_\alpha x_\alpha + \hat{H}_{\text{osc}} + \Delta V(q) \tag{5.6}$$

$$\hat{H}_{\text{osc}} \equiv \sum_\alpha (p_\alpha^2/2m_\alpha + \frac{1}{2}m_\alpha \omega_\alpha^2 x_\alpha^2) \tag{5.7}$$

where the "counterterm" $\Delta V(q)$ may or may not be present depending on the physical nature of the system and environmental considered; in these cases where it does arise, it is generally of the form

$$\Delta V(q) = \frac{1}{2}\sum_\alpha \tilde{C}_\alpha^2/m_\alpha\omega_\alpha^2 \tag{5.8}$$

The effective mass M^* may or may not be the same as the original mass of the uncoupled system (in most cases of practical interest it is). The "bath" parameters[9] m_α, ω_α and C_α are in the most general case functions of temperature.

It should be strongly emphasized that the representation (5.6-7) is valid under the stated conditions, even when the environment looks prima facie nothing like a bath of simple harmonic oscillators; further, even when it does have this nature, the notional "oscillators" of eqn. (5.7) may have little or nothing to do with the physical oscillators. This claim may at first sight look surprising, but it is actually nothing but a generalization of a simple trick which has been employed since the mid-nineteenth century in atomic physics (without, of course, its inventors being aware at the time of its justification!): The physical picture within which the spectroscopic data were originally interpreted regarded the atom ("environment" in our language) as composed of simple harmonic oscillators interacting linearly with the electromagnetic field ("system"). Although we now of course know that the true atomic states are not at all like those of a set of simple harmonic

[8] The final result (5.1) can also be justified for the case $\hat{H}_{\text{int}} = \hat{\Omega}_1 f(\hat{q})$, see ref. [8].

[9] The "masses" m_α can actually be set equal to 1 without loss of generality.

oscillators, this picture was so spectacularly successful (in pre-laser days!) that the language associated with it ("oscillator strengths", etc.) is standard in atomic physics to this day. The reason for its success is essentially that, at least under pre-laser conditions, the electromagnetic field interacts very weakly with any one atomic degree of freedom (the probability of any one atom being excited, even when the field is right on resonance, is very small); this is completely compatible with the fact that the interaction with the whole collection of atoms in, say, a gas sample may be sufficiently strong that the field is completely absorbed in the gas. In exactly the same way, in the generic case we are considering the interaction of our system with any one degree of freedom of the environment is very weak, but this is entirely compatible with the possibility that the total interaction with the environment may be strong enough to change the system dynamics qualitatively.

An objection may be raised to the above argument, to the effect that even though the environment is "macroscopic" it is not immediately obvious that the perturbation of any one degree of freedom need always be small; consider for example a spoon stirring a liquid in a container which is not very much larger then the spoon. In this case we can save the argument by reformulating the problem in the adiabatic (Born-Oppenheimer) basis (i.e. the basis in which the wave functions of the environment respond adiabatically to the motion of the system) and defining our oscillator coordinates with respect to this basis: see for details ref. [6], Appendix C. The upshot of this operation is again to produce an effective Hamiltonian of the form (5.6), with no renormalization of the system "mass" (i.e. $M^* \equiv M$) and a counterterm given by the standard form (5.8).

The Hamiltonian (5.6) can be used to discuss various problems which are of interest for the extended system, such as the classical dynamics (i.e. the dynamics in that parameter regime where quantum mechanics gives predictions indistinguishable from those of classical mechanics) and the rate of tunneling through a potential barrier in $V(q)$. A very important point is that while in general we do not know the individual oscillator parameters $m_\alpha, \omega_\alpha, C_\alpha$ a priori, the only combination of these quantities which enters the dynamics of the system is the spectral density $\tilde{J}(\omega)$ defined by

$$\tilde{J}(\omega) \equiv \pi/2 \sum_\alpha (\mid \tilde{C}_\alpha \mid^2 / m_\alpha \omega_\alpha) \delta(\omega - \omega_\alpha) \tag{5.9}$$

and this quantity can be read off from the experimentally observed classical dynamics. In particular, for the common case where the mass is not renormalized and the counterterm $\varDelta V(q)$ is given by the "standard" expression (5.8), $\tilde{J}(\omega)$ is simply related to the phenomenological complex admittance $Y(\omega)$ which appears in the classical equation of motion:

$$\tilde{J}(\omega) = \operatorname{Im}(\mathrm{i}\omega Y(\omega)). \tag{5.10}$$

For example, in the common case where the classical equation of motion is of the form

$$M\ddot{q} + \eta\dot{q} + \partial V/\partial q = 0 \tag{5.11}$$

we have $Y(\omega) \equiv \eta$, and thus the quantity $\tilde{J}(\omega)$ is simply $\eta\omega$. This observation enables us to make predictions about various aspects of the essentially quantum-mechanical behavior (e.g. tunneling out of a metastable state) in terms of parameters such as the effective friction coefficient which can be determined from experiments conducted entirely in the classical regime, i.e. to predict the quantum behavior in the presence of the environment with no adjustable parameters. For an example of the comparison of such predictions with the measured behavior (in this case the tunneling of a Josephson junction out of the zero-voltage state) see ref. [10].

Eqns. (5.5-6) may be thought of as defining a problem of dynamics in a higher dimensional space whose axes are q and the collection $\{x_\alpha\}$; however, we are interested in the environment coordinates x_α only to the extent that they affect the dynamics of the system variable q. An important point is that if there is no mass renormalization and the potential counterterm $\Delta V(q)$ has the "standard" form (5.5), then by comparison with the uncoupled problem ($C_\alpha \equiv 0$) the minimum potential attainable for given q is unchanged, but occurs at a point in the many-dimensional space which is shifted away from the q-axis. Suppose now that the original potential has two nearly degenerate minima; then in the presence of the coupling these minima remain, but are shifted away from the q-axis in opposite directions. If the original barrier between the two minima is high, so that tunneling between the two minima is a rare event, then this is even more true in the presence of the environmental coupling; it is easy to see that under the stated conditions ($M^* = M, \Delta V(q)$ given by (5.5)) the coupling tends to suppress the tunneling[10]. It then makes eminent sense to do what one would do in the uncoupled case, namely to truncate the variable q to the two discrete values, say $\pm q_0/2$, corresponding to the two minima and to label these in the 2-state notation by the eigenvalues ± 1 of $\hat{\sigma}_z$. There is then a term in the Hamiltonian of the form $-\frac{1}{2}\epsilon\sigma_z$, where ϵ is the difference between the zero-point energies in the two wells in the uncoupled problem; for an originally symmetric well ϵ is trivially zero.

We should also expect a term proportional to σ_x and representing the tunneling between the two wells, but the consideration of this term raises a rather delicate point. If in the uncoupled problem the tunneling amplitude between the two wells is Δ_0, then we should expect that the "bath" oscillators with frequency much larger than Δ_0 (but possibly comparable to the small-oscillation frequency ω_0 in the individual wells) could have two effects, which are actually somewhat related: (1) they can affect the actual process of transmission through the barrier (2) they could "follow" the system adiabatically and thus renormalize Δ downwards by a Franck-Condon factor which may be very small compared to unity. It turns out that effect (1) is associated with oscillations of frequency $\gtrsim \omega_0$, while effect (2) can in general have contributions from oscillations in the whole range $\Delta_0 \ll \omega \ll \omega_0$, (which typically may span several orders of magnitude). Finally, oscillators of frequency $\lesssim \Delta_0$ can have a more complicated effect, since

[10] For a detailed discussion of when tunneling is suppressed and when it is enhanced, see ref. [7].

in general they do not follow the system adiabatically and may contribute to real dissipation.

As a consequence of the above considerations, it is convenient to introduce a cutoff ω_c such that $\Delta_0 \ll \omega_c \ll \omega_0$, and to take into account the effects (1) and (2) produced by oscillations with frequency $< \omega_c$ by renormalization of the effective tunneling matrix element:

$$\Delta_0 \to \Delta(\omega_c) \tag{5.12}$$

where typically $\Delta \ll \Delta_0$. The remaining oscillators are taken into account explicitly in the new effective Hamiltonian by truncating the interaction term in (5.6) to its value for $q = \pm\, q_0$. Noting that the term in $\Delta V(q)$ contributes only an irrelevant constant, we then see that we may indeed write the truncated Hamiltonian in the form (5.1), with C_α given by $q_0 \tilde{C}_\alpha$, the sum over α taken only over the low-energy ($\omega_\alpha < \omega_c$) oscillators and the effective tunneling matrix element Δ a function of ω_c. (As we should expect, the dependence on the relatively arbitrary value of ω_c cancels out in the final results.) For a (relatively) rigorous formulation of the above argument, and a precise prescription for computing the quantity $\Delta(\omega_c)$, see Section 2 and Appendix A of ref. [8]. The upshot is that for a large class of qubits which are obtained by truncation of an originally continuous variable, the spin-boson Hamiltonian (5.1-2) is a valid approximation, with the quantity $J(\omega)$ given by

$$J(\omega) = \tilde{J}(\omega)\theta(\omega_c - \omega) \tag{5.13}$$

where $\tilde{J}(\omega)$ may be obtained, as described above, from experiments conducted in the classical regime. Note that $J(\omega)$ is equal to the $J'(\omega)$ of Lecture 4 up to a constant factor: in fact, comparing (4.6) with (5.9) we find

$$J'(\omega) \equiv (q_0^2 \hbar / 4\pi) J(\omega) \tag{5.14}$$

Once we are assured of the validity of (5.1), the dynamics of the system (i.e. of the "spin" components σ_i) can be obtained by a variety of techniques, since the oscillators can be formally integrated out of the problem; a particularly convenient technique is the two-state functional integral method used in ref. (8). The resulting expressions involve complicated multiple integrals, but these can be evaluated numerically or, in a few limiting cases, analytically. Although the direct output of this method is the dynamics only of $\langle \sigma_z \rangle(t)$ rather than of all three spin components, we can use the results as a check on other more phenomenological methods such as that described at the end of Lecture 4. There is little point in going into the technical details here, so I will just concentrate on a couple of crucial qualitative features.

The dynamics of the spin-boson system turns out to be dominated by two physical effects which are in a sense mutually complementary: (1) As a result of the interaction with oscillators of frequency $\gtrsim \Delta$, the system undergoes "false" decoherence and its effective tunneling frequency is pushed downwards: $\Delta \to \Delta_{\text{eff}}$ (2) As a result of interactions with oscillators of frequency $\sim \Delta_{\text{eff}}$, the system undergoes "true" decoherence involving irreversible loss of both energy and information to the environment. It turns out that what is crucial in determining the final rate of decoherence (and in some cases even the qualitative behavior) is not so much the magnitude of $J(\omega)$ but rather its dependence on frequency for $\omega \lesssim \Delta$. The reason is the following: If we start by trying to make an adiabatic elimination of the oscillators with frequency say $> 10\Delta$, then the result is to depress Δ to a smaller value Δ'. We then iterate the process by eliminating oscillators with frequencies in the range $10\Delta' < \omega < 10\Delta$, and so on. At each stage the renormalization is given by a Franck-Condon type formula of the general form

$$\Delta_{\text{new}} = \Delta_{\text{old}} \cdot \exp - \int_{10\Delta_{\text{old}}} \frac{q_0^2}{2\pi\hbar} \cdot \frac{J(\omega)}{\omega^2} \coth\,(\beta\omega/2)\mathrm{d}\omega \qquad (5.15)$$

where the upper limit on the integral is the lower cutoff at the previous step and need not be specified explicitly for our purposes. Suppose for definiteness that $J(\omega)$ has a simple power-law form for $\omega \lesssim \Delta : J(\omega) = \text{const.}\ \omega^s$. It is clear that if s < 1 at $T = 0$, or $s < 2$ at finite T (with $kT \gtrsim \Delta$), then the Franck-Condon factor exp - ... "avalanches" and the iteration eventually forces Δ_{eff} to zero. Under these circumstances there are no coherent transitions between the eigenstates of $\hat{\sigma}_z$, only thermally activated incoherent hopping (and at zero temperature, for $s < 1$, the system is completely "localized" in one eigenstate or the other). On the other hand, if $s > 2$, then irrespective of the value of T the exponent of the Franck-Condon factor becomes progressively closer and closer to zero as the iteration proceeds, so that Δ_{eff} eventually converges to a finite value. Once this has happened, we can effectively apply the perturbation-theoretic argument of Lecture 4 with Δ replaced by Δ_{eff} , so that we obtain the results (4.10) and (4.12-13) with the inverse decoherence time given by

$$T_1^{-1} = T_2^{-1} = 2\pi \coth\,(\beta\Delta_{\text{eff}}\ /2)J'(\Delta_{\text{eff}}\ ) \equiv \frac{q_0^2}{2\hbar} \coth\,(\beta/\Delta_{\text{eff}}\ )J(\Delta_{\text{eff}}\ ) \quad (5.16)$$

The case $s = 1$ is particularly interesting, since at zero temperature it is clear from the above that 1 is the "critical" value of the exponent which separates two qualitatively different types of behavior. It is also arguably the generic case, since the quantity $J(\omega)$ ($\sim$ Im $\chi_{\Omega}(\omega)$) is formally odd in ω, and thus if it is analytic at $\omega = 0$ the leading term in the limit $\omega \to 0$ must be linear in ω; in the absence of pathology this term should not vanish. In fact, by (5.10) and (5.13) it will be characteristic of any system which in the classical regime is described by a linear dissipation (viscosity) coefficient η: we then have

$$\frac{q_0^2}{2\pi\hbar} J(\omega) = (\eta q_0^2/2\pi\hbar)\omega \equiv \alpha\omega \qquad (5.17)$$

where we note that the quantity α is dimensionless. A system with the spectral density (5.17) is often referred to in the literature as "ohmic": such systems have been intensively studied, in particular in ref. [8]. In that reference it is shown, inter alia, that the value of $\alpha = 1$ is the discriminant (at zero temperature and in the limit $\Delta \ll \omega_c$) between two qualitatively different types of behavior of $\langle\sigma_z\rangle(t)$, corresponding respectively to localization (as in the case s < 1) and to damped (in fact possibly overdamped) oscillation. For $\alpha kT \gg \Delta$ (and actually in a broader region of the phase diagram) the behavior is always overdamped, i.e. $\langle\sigma_z\rangle(t) \sim \exp -\Gamma t$, with a rate Γ which is proportional to $T^{2\alpha-1}$ and thus is an increasing function of temperature for $\alpha > \frac{1}{2}$ but a decreasing one for $\alpha < \frac{1}{2}$.

In the context of quantum computing the most interesting regime is the small-α limit; it seems that in practice it should not be particularly difficult to obtain, in e.g. an rf SQUID device, values of α of the order $10^{-3} - 10^{-4}$. For such small values of α (and reasonable values of the ratio Δ/ω_c) it turns out that the "false" decoherence effects (renormalization $\Delta \to \Delta_{\text{eff}}$) are negligible, and the output of a functional-integral calculation is completely compatible, as we should expect, with that of the simple perturbation scheme outlined at the end of Lecture 4. Thus, reproducing these results for convenience, we have

$$\frac{\mathrm{d}\langle\sigma_x\rangle}{\mathrm{d}t} = -\frac{(\langle\sigma_x\rangle - \langle\sigma_x\rangle_{\text{eq}})}{T_1} \tag{5.18}$$

$$\mathrm{d}\langle\sigma_y\rangle/\mathrm{d}t = \Delta\langle\sigma_x\rangle - \frac{\langle\sigma_y\rangle}{T_2} \tag{5.19}$$

$$\frac{\mathrm{d}\langle\sigma_z\rangle}{\mathrm{d}t} = -\Delta\langle\sigma_y\rangle \tag{5.20}$$

where the decoherence rate T_2^{-1} ($= T_1^{-1}$ in this case) is given by

$$T_2^{-1} = \coth\,(\beta\Delta/2)J(\Delta) \equiv \pi\alpha\Delta \coth\,(\beta\Delta/2) \tag{5.21}$$

which reduces to $\pi\alpha\Delta$ for $kT \ll \Delta$ and to $2\pi\alpha kT$ for $kT \gg \Delta$. Note that the decoherence time is much longer than the period Δ^{-1}, so that the oscillation of $\langle\sigma_z\rangle$ is underdamped, so long as $\alpha kT \ll \Delta$, even if kT is $\gg \Delta$. Thus, if the (many) other practical difficulties can be overcome, it seems that the use of ohmically dissipative systems such as rf SQUIDs as elements in a quantum computer need not automatically be vitiated by decoherence.

6 Lecture 6 – The Quantum Measurement Problem

The topic of this lecture is one on which I have written at length in many places. Rather than repeat all the relevant considerations here, I will try to concentrate on two aspects in particular, namely the extent to which the recent interest in quantum computation puts the issue of quantum measurement in a new light, and the significance of recent experimental advances in the area of Josephson technology.

To recap briefly an argument which has by now become fairly standard: If we go back to the simple 2-state system analyzed in Lecture 1, the difference between a cbit and a qubit is that while the former, by definition, must at all times "really be" in one or other of the two basis states $|\uparrow\rangle$ or $|\downarrow\rangle$, the latter can in addition be in a superposition of the two. As we saw in Lecture 2, this fundamental difference is reflected in the fact that a qubit can show types of behavior (e.g. as regards two-time correlations) which are totally impossible for any cbit. In other words, it is impossible to explain this behavior on the assumption that at all times the qubit "really is" in one or other of the two available states. Despite this, any "measurement" of the qubit in the canonical basis will always reveal it to be in one of the two basis states, i.e. either $|\uparrow\rangle$ or $|\downarrow\rangle$.

But what does this "measurement" entail? We certainly cannot inspect the state of a (microscopic) qubit such as a nuclear spin directly. What we do in practice is to provide a mechanism by which each of the microscopic states $|\uparrow\rangle$, $|\downarrow\rangle$ will be correlated with a particular macroscopic state of a (macroscopic) measuring device, i.e. we formally provide a Hamiltonian which will give rise to a unitary evolution operator $\hat{U}(t)$ having the property[11]

$$\hat{U}(\infty)\,|\uparrow\rangle\mid X_0\rangle =|\uparrow\rangle\mid X_\uparrow\rangle \tag{6.1}$$

$$\hat{U}(\infty)\,|\downarrow\rangle\mid X_0\rangle =|\downarrow\rangle\mid X_\downarrow\rangle \tag{6.2}$$

Here $\mid X_0\rangle$ is the initial state of the measuring device, and $\mid X_\uparrow\rangle$ and $\mid X_\downarrow\rangle$ are states of the device (one of these may be identical to $\mid X_0\rangle$) such that they are not only mutually orthogonal but also macroscopically distinct, so that we can read off with the naked eye which is the final state of the device on a given run (and thus assign (by definition!) a "measured" value of σ_z on this run to the qubit). An often quoted example of this process (though one which is likely to be applicable to the quantum computing case only in a minority of implementations) is a Stern-Gerlach device: here a particle in state $|\uparrow\rangle$ (i.e., literally, with spin "up") follows one trajectory in the magnetic field gradient and triggers, say, counter 1, while a particle in state $|\downarrow\rangle$ follows a different trajectory and triggers counter 2. We can then inspect with the naked eye which of the two counters was triggered on a given run, and thereby assign a value of σ_z to the particle "measured" on that run.

Continuing the standard argument, let us now consider what happens if the qubit (or more accurately the ensemble of qubits) is initially described as being not in state $|\uparrow\rangle$ or state $|\downarrow\rangle$ but rather in the linear superposition state $\alpha\ |\uparrow\rangle + \beta\ |\downarrow\rangle$, where α, β are complex coefficients such that $|\alpha|^2 + |\beta|^2 = 1$. The crucial point, now, is that so long as we stick within the framework of standard quantum mechanics, the evolution operator $\hat{U}(t)$ is strictly linear, i.e. if state 1 evolves into $1'$ and 2 into $2'$, then a linear superposition of 1 and 2 evolves into the corresponding linear superposition of $1'$ and $2'$. When applied to eqns. (6.1), this principle immediately implies that

$$\hat{U}(\infty)\{\alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle\} = \alpha\,|\uparrow\rangle\mid X_\uparrow\rangle + \beta\,|\downarrow\rangle\mid X_\downarrow\rangle \tag{6.3}$$

in other words, the effect of the measurement process is to entangle the (microscopically distinct) states of the qubit with the (macroscopically distinct)

[11] Strictly speaking, (6.1) and (6.2) describe what is technically called an "ideal" measurement, i.e. one in which the state of the microsystem does not change.

states of the measurement apparatus. Provided that we believe that the standard quantum formalism can indeed be extrapolated in the way we have implicitly assumed, then there is absolutely no doubt that the technically correct description of the ensemble of "universes" (I shall simply use the term "universe" as a shorthand for "microsystem plus measuring device plus (possibly) environment") at the end of the measurement process is, under the idealized conditions implicitly assumed, exactly the right-hand size of eqn. (6.3). In words, the "universe"ends up in a quantum superposition of states which are macroscopically distinguishable. On the other hand, the whole point of correlating the state of the qubit to that of the measuring device is that we can "read off" the latter with the naked eye; in other words, whenever we inspect the final state of the device on any given run we always find it to be either in state $\mid X_{\uparrow}\rangle$ or in the macroscopically distinct state $\mid X_{\downarrow}\rangle$. For example, in the Stern-Gerlach setup (if it is properly designed) we find, on each run, that either counter 1 or counter 2 has clicked, with the other registering nothing; we never find a result which in any way reflects the linear superposition (it is indeed difficult to imagine what such a result would be like!).

This, then, is the fundamental "measurement paradox" of quantum mechanics: how come that although the ensemble of "universes" is definitely described, at the end of the measurement process, by the linear superposition (6.3) of two macroscopically distinguishable states $|\uparrow\rangle \mid X_{\uparrow}\rangle$ and $|\downarrow\rangle \mid X_{\downarrow}\rangle$, inspection on any given run always reveals one result or the other? In the succinct formulation of the late John Bell, how do you turn an "and" into an "or"?

Before turning to alleged resolutions of the paradox, let us note that the "entanglement" between the state of the qubit and that of the device represented by (6.2) is in no way essential to the formulation. Indeed, there are cases (for example, when the "qubit" is represented by the two polarization states of a single photon, and the photon is then absorbed in one or other counter depending on its polarization) when the qubit is actually physically annihilated by the measurement process, so that the final state of the "universe" is just that of the macroscopic device, $\alpha \mid X_{\uparrow}\rangle + \beta \mid X_{\downarrow}\rangle$. Of course, in such a case there is typically a high degree of entanglement internal to the device, cf. below.

The quantum measurement paradox can be dated to Schrӧdinger's famous "cat" paper of 1936, so it has now been around for more than two generations, and has generated a commensurate number of papers in the literature. It is superfluous to try to review all of them here, so I will just enumerate the main classes of alleged resolution known to me.

(1) **"Statistical" interpretations.** This class of resolution, which may be regarded as a logical extension of the classic "Copenhagen interpretation" and is nicely exemplified by Ballentine's 1970 review article [11], takes the view that the only legitimate interpretation of the formalism of quantum mechanics is as a formal calculus whose sole function is to predict (according to the standard measurement axioms) the probability of obtaining a given macroscopic result (click in counter, etc.). Thus, for example, the only "meaning" of the superposition (6.3) is that if we inspect the state of the measuring device on a particular run

of the ensemble described by it, we have a probability $|\alpha|^2$ of obtaining the result represented by $|X_\uparrow\rangle$ (e.g. "counter 1 clicked") and a corresponding probability $|\beta|^2 = 1 - |\alpha|^2$ of obtaining that represented by $|X_\downarrow\rangle$ (e.g. "counter 2 clicked"). Beyond this, the formalism of QM has no more meaning than the recipes used by the ancient Chaldeans to predict eclipses.

(2) **"Mentalistic" and "many-worlds" resolutions:** In contrast to the statistical resolution, which essentially holds that the QM formalism describes nothing in the real world, this class of resolution holds that it describes everything, or at least everything which is part of the "physical" rather than the "mental" world. In other words, the state of the physical universe at the end of the measurement process really does correspond to the linear superposition (6.3) – a quantum superposition of macroscopically distinguishable states. Then how come that we always "see" one or the other of these two states? In the "mentalistic" variant, it is the special qualities of human consciousness which effect the appropriate reduction or "realization"; these qualities, it is said, cannot be described by quantum mechanics, even in principle. The "many-worlds" (or more soberly "relative-state") resolution is even more radical, at least in the formulation of its more enthusiastic proponents: even our subjective consciousness of seeing a particular result is, it is said, an illusion - it corresponds to only one of a number (in this case two) of "parallel universes" which, even after the apparent observation, are "equally real". (If someone could explain to me what these words, ostensibly English, actually mean, I might know whether or not I believed this interpretation.)

(3) **"Calvinistic" resolution:** this idea, due to Schulman, is based on the time-reversibility of QM: for any given definite final state, we can always find an initial microscopic state, in general a superposition, which when supplemented by appropriate conditions on the initial internal degrees of freedom of the device, etc., is guaranteed to lead to this final state and only to this, i.e. to a definition outcome of the measurement. This resolution requires that it is always just such an initial state which occurs in any given experiment, which in turn would appear to imply that the experimenter's sense of having freedom to determine the initial state is an illusion.

(4) **Resolution based on decoherence:** this is what one might call the "establishment" solution to the measurement problem, and I will return to it below.

(5) **Resolutions based on modification of the QM formalism:** In contrast to resolutions (1-4), all of which assume that QM is in some sense or other a complete description of the universe (or at least of the "physical" universe), this line of resolution is based on the postulate that at some stage along the way from the atomic level up to the "everyday" world, the QM formalism becomes incomplete and has to be supplemented by other, currently unrecognized, laws of physics. Although this general idea has a long history and many variants, by far the best-developed version is that due to Ghirardi, Rimini, Weber and Pearle ("GRWP" theory) (see e.g. refs. [12,13]). In this theory there exists a sort of universal background noise, which is not itself subject to QM principles (this is crucial!) and whose effect, qualitatively speaking, on a superposition of

quantum states (in an appropriate basis) is to drive it definitely into one state or the other at random. The "preferred" basis corresponds to that in which the center-of-mass position of the object being described is diagonal, and the rate of "reduction" to one state or the other is proportional to the number N of particles involved in the object (it is an interesting and significant feature of the GRWP theory that this feature emerges naturally and does not have to be put in "by hand"). Thus, by appropriate adjustment of the characteristic parameters of the model one can ensure, for example, that a single photon ($N = 1$) emitted from the most distant quasar currently known can preserve its coherence over the whole of its trajectory across the cosmos and thus show the characteristic interference effects in a Michelson interferometer, while on the other hand a cat or other macroscopic object ($N \sim 10^{23}$) will collapse into one or other of two macroscopically distinct states (e.g. "living" or "dead") over a timescale much shorter than that accessible to direct human observation.

An exciting feature of resolutions of this class, whether the specific GRWP version or others, is that in principle they are subject to experimental test vis-à-vis the standard QM picture: I return briefly to this question below.

———

I now turn specifically to the "establishment" solution of the quantum measurement problem, namely that based on decoherence. (I call it the "establishment" solution because it is the one at which physicists who have spent long years using QM in their own research without paying much attention to the measurement problem, usually arrive once they do start paying attention). It is probably easiest to understand the general principle of this alleged resolution if we focus on the case in which the original qubit disappears, so that in the naive description used above the measurement device is described at the end of the measurement process by its own wave function, namely

$$\Psi(\infty) = \alpha \mid X_\uparrow\rangle + \beta \mid X_\downarrow\rangle \tag{6.4}$$

– a quantum superposition of macroscopically distinguishable states. The first stage of the argument is then, essentially, that this is not in fact a realistic description of the final state of the device, and that once one corrects it appropriately it is impossible to distinguish the state from a mixture of $\mid X_\uparrow\rangle$ and $\mid X_\downarrow\rangle$ with weight $\mid \alpha \mid^2$ and $\mid \beta \mid^2$ respectively. As a matter of fact, it may not even be necessary to invoke the phenomenon of decoherence to make this point, since (a) to distinguish the pure state (6.4) from the corresponding mixture it would be necessary to measure the expectation value of some operator $\hat{\Omega}$ such that $\mid \langle X_\uparrow \mid \hat{\Omega} \mid X_\downarrow\rangle \neq 0$, and if $\mid X_\uparrow\rangle$ and $\mid X_\downarrow\rangle$ represent macroscopically distinct states such an operator is very hard to find (though see below), and (b) in many realistic cases, even if the final state of the device is indeed schematically of the form (6.3), there are liable to be extra uncontrolled phase factors in the coefficients due to random classical fields, etc., which will mean that the ensemble average of $\hat{\Omega}$ is zero, thus failing to distinguish the pure state from a mixture. However, since as we shall see, there exist at least some reasonably "macroscopic" systems where considerations (a) and (b) drastically fail, recourse is usually had at this point to the idea of decoherence.

Schematically, the argument is simply that any realistic measuring device interacting with a physically realistic environment (e.g. the blackbody radiation field) the "final" state (6.3) will very rapidly evolve into an entangled state of the device and the environment:

$$(\alpha \mid X_{\uparrow}\rangle + \beta \mid X_{\downarrow}\rangle) \mid E_0\rangle \to \alpha \mid X_{\uparrow}\rangle \mid E_{\uparrow}\rangle + \beta \mid X_{\downarrow}\rangle \mid E_{\downarrow}\rangle \equiv \Psi(t) \tag{6.5}$$

where $|E_{\uparrow}\rangle$ and $|E_{\downarrow}\rangle$ are mutually orthogonal states of the environment (one of which may or may not coincide with the initial environment state $|E_0\rangle$). Then the fundamental theorem of decoherence (lecture 3) assures us that the results of any measurement at time t on the device (i.e. of any quantity which is a unit operator with respect to the environment) will be identical to those which would be obtained for a mixture of states $\mid X_{\uparrow}\rangle$ and $\mid X_{\downarrow}\rangle$ with probability $\mid \alpha \mid^2$ and $\mid \beta \mid^2$ respectively.

It should be emphasized that those physicists, including the present writer, who find the "decoherence" resolution of the paradox unsatisfactory do not object to any of the statements made so far, in particular in the statement that any realistic measuring apparatus is indeed likely to get very rapidly entangled with its environment in the way described by (6.5). What they object to is stage 2 of the argument, which I now outline.

The second step in the "decoherence" argument (which is not infrequently so much taken for granted that it is not even stated explicitly!) is that since "all predictions" for the true (entangled) state of the system are identical to those for a mixture, and since a mixture is just what we get in the case where we know that the system must definitely "be" in one state or the other, but do not know which with certainty, therefore by this time it "really is" in one or the other state, and the QM description, while formally correct, is no more than a description of our ignorance.

To my mind, there are two major objections to step 2. The first is technical and can probably be dealt with by a more sophisticated formulation, the second philosophical and to my mind quite fatal. The first objection goes back to the important distinction made in Lecture 3 between "true" and "false" decoherence: If the entanglement with the environment is adiabatic in nature, then the decoherence is "false" the fact that no measurement at time t can reveal the difference between the true state and the mixture by no means implies that no measurement in the future will be able to do so. I believe that the recent interest in quantum computing has appreciably strengthened the significance of this point (or to be more accurate, of the public perception thereof!), since the whole quantum computing scheme relies essentially on entanglement which is adiabatic in nature, and the fact that the reduced density matrix of an individual qubit is diagonal (i.e. of mixture form) throughout much or all of the computation in no way implies it is equivalent to a cbit. Morever, consideration of the quantum computing scheme reveals that whether the decoherence at a given time is "true" or "false" is not determined uniquely by the state at that time, but also depends on the nature of the control and other operations which take place at later times. Thus, in order to get even the technical aspects of the decoherence solution to the measurement problem right, it is necessary as a minimum to ensure that in the specific case considered the decoherence resulting from the entanglement with the environment is of the "true" variety. However, in many cases, involving (for example) interaction with the blackbody radiation field, it is very plausible that this is true, so I do not regard this technical point as an insuperable objection to the "decoherence" solution.

A much more serious objection is the following: Unless one believes that the formalism of QM corresponds to nothing at all in the real world at either the microscopic or the macroscopic level (in which case one may as well just embrace the full-blooded "statistical" resolution (1) from the start) then a microscopic superposition of the form $\alpha\,|{\uparrow}\rangle + \beta\,|{\downarrow}\rangle$ is telling us *something* about the state of the system in question; at the very least, we can legitimately make the "negative" statement that it is *not* true either that the system is definitely in state $|{\uparrow}\rangle$ or that it is definitely in state $|{\downarrow}\rangle$. (Whether we choose to go further and say that it is in "both", or in "neither", or to put any other verbal window-dressing on the situation, is unimportant for the present argument.) Now as we work our way up from the level of atoms to that of the everyday world, continuing to describe everything by QM, there is no point at which the quantum formalism changes in any way. Consequently, when we are presented with a *macroscopic* superposition state like (6.4), or even (6.5), however much entanglement it involves, it is to my mind *totally illegitimate* to reinterpret the formalism to infer that the universe, or any part of it, is definitely in one state or the other. Such a reinterpretation involves a fundamental confusion between the *meaning* of the formalism of QM, and the *evidence* that that meaning is correct. At the microscopic level, we have plenty of evidence, in the form of various interference experiments etc., that the meaning is (at least) as described in the above negative statement. By the time we get to the macroscopic level, the (direct) evidence has evaporated, but the meaning of the formalism cannot have changed!

To conclude this lecture, I comment briefly on the significance, in the present context, of two remarkable experiments published recently on Josephson systems (rf SQUIDs, to be precise). Clearly, the very existence of a quantum measurement paradox depends crucially on the assumption that there are physical situations where the correct quantum-mechanical description of the state of the universe is a macroscopic superposition of the type (6.4), or at worst (6.5); crudely speaking, this is equivalent to the exclusion of a solution of class (5) above. Now it is immediately obvious that if in all realistically obtainable cases the final state is of the form (6.5), and the decoherence is of the "true" variety, then it will be forever impossible to tell whether this assumption is correct. Thus, there is great interest in identifying systems – whether or not they can realistically serve as "measuring devices" is irrelevant, at least in the first instance – where one may be able to produce superpositions of the form (6.4) (or of type (6.5) with only "false" decoherence) and moreover verify that the state is indeed a superposition of macroscopically distinguishable states rather than a mixture thereof. Although various types of system have been considered in this context, one of the most promising – and certainly the one where the difference between the two basis states is by most reasonable criteria the most "macroscopic" – is the rf SQUID ring, briefly introduced in Lecture 4. As mentioned there, the two basis states correspond to currents flowing respectively in the clockwise and counterclockwise directions with a magnitude of a few μA. The most convincing evidence that this system is behaving like a genuine qubit rather than a cbit would be a measurement of the two-time correlation functions, as described for the microscopic qubit in Lecture 2; a result which conformed to the quantum-mechanical prediction (2.3), or was sufficiently close to it, would by the arguments of that lecture show that the state of the system at intermediate times could not possibly correspond to it always being definitely $|\uparrow\rangle$ or $|\downarrow\rangle$.[12] This experiment has not yet been done. However, in the last few months two groups, at SUNY Stony Brook (14) and Delft (15) respectively, have conducted a related experiment in which evidence for superposition was obtained from the "level-repulsion" phenomenon: If we consider the two-state Hamiltonian (4.3) and neglect the last two terms, so that

$$\hat{H} = -\frac{1}{2}\Delta\hat{\sigma}_x - \frac{1}{2}\epsilon\hat{\sigma}_z \tag{6.6}$$

then if we assume that QM still works at this level, we get the following results: If the state of the system is always definitely $|\uparrow\rangle$ or $|\downarrow\rangle$, the term in $\hat{\sigma}_x$ is ineffective and the energies of the states are just $\pm\frac{1}{2}\epsilon$, i.e. linear in the bias. If on the other hand we allow superpositions of $|\uparrow\rangle$ and $|\downarrow\rangle$, then the energies are of the familiar form

$$E = \pm\sqrt{\epsilon^2 + \Delta^2} \tag{6.7}$$

It is the latter result which is found, by a spectroscopic technique, in the two experiments. It should be emphasized that while the result found is entirely consistent with the hypothesis that the system is in a superposition of the macroscopically distinguishable states corresponding to $|\uparrow\rangle$ and $|\downarrow\rangle$, it does not rigorously exclude the alternative possibility that it is at all times in one or other of the two states; this possibility is excluded only if we make the explicit assumption that QM is still working at this level, which could be regarded as in some sense begging the question. Thus, it is still of very considerable (I would say fundamental) interest to do the two-time correlation experiment, and I believe there is a good chance it will be done within the next decade. If so, and if as most people (including the present writer) expect, it gives the results predicted by QM, this would bring the quantum measurement paradox considerably closer to home. By contrast, should the experiment give results which are prima facie incompatible with the QM predictions, the initial reaction of most of the community will certainly be that these predictions have not been made with sufficient care, and

[12] However, without further modification such a result would not refute the (current version of the) GRWP theory, since there is no appreciable "separation" of the center of mass in the two states superposed.

that (for example) there are “hidden” sources of decoherence which have not been adequately taken into account in the calculations. Undoubtedly, if the conventional belief that QM is the complete and ultimate truth about the physical universe is ever overthrown, it will not be without a long and bitter struggle, in which issues such as the above are subjected to minute examination; in fact, even in the best of circumstances I would estimate the chances of it happening in my lifetime to be essentially zero. But that is no reason not to pose the question!

Note added in proof: Since the above was written there have been some spectacuar advances in this area. See in particular the paper of Vion et. al., Science, **296**, 886 (2002).

References

1. L. D. Landau and E. M. Lifshitz, Quantum Mechanics, Pergamon, London 1958.
2. G. Badurek, H. Rauch and J. Summhammer, Phys. Rev. Letters **51**, 1015 (1983).
3. J. M. Raimond, M. Brune and S. Haroche, Revs. Mod. Phys. **73**, 565 (2001).
4. W. G. Unruh and W. H. Zurek, Phys. Rev. D **40**, 1071 (1989).
5. M. Grifoni, E. Paladino and U. Weiss, Eur. Phys. J. B**10**, 719 (1999).
6. A. O. Caldeira and A. J. Leggett, Ann. Phys. (NY) **149**, 374 (1983).
7. A. J. Leggett, Phys Rev. B **30**, 1208 (1984).
8. A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg and W. Zwerger, Revs. Mod. Phys. **59**, 1 (1987).
9. A. J. Leggett, in *Chance and Matter*, ed. J. Souletie et al., North-Holland, Amsterdam 1987, p. 395.
10. J. Martinis, M. H. Devoret and J. Clarke, Phys. Rev. B **35**, 4682 (1987).
11. L. E. Ballentine, Revs. Mod. Phys. **42**, 358 (1970).
12. G. C. Ghirardi, A. Rimini and T. Weber, Phys. Rev. D**34**, 470 (1986).
13. P. Pearle, Phys. Rev. A **39**, 2277 (1989).
14. J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo and J. Lukens, Nature **406**, 43 (2000).
15. C. H. van der Wal, A. C. J. ter Haar, F. K. Wilhelm, R. N. Schouten, C. J. P. M. Harmans, T. P. Orlando, S. Lloyd and J. E. Mooij, Science **290**, 773 (2000).

Introduction to Quantum Computation

Artur Ekert

Abstract. A computation is a physical process. It may be performed by a piece of electronics or on an abacus, or in your brain, but it is a process that takes place in nature and as such it is subject to the laws of physics. Quantum computers are machines that rely on characteristically quantum phenomena, such as quantum interference and quantum entanglement in order to perform computation. In this series of lectures I want to elaborate on the computational power of such machines.

1 Physical Representation of Information

Suppose you have n physical objects and each object has k distinguishable states. If you can access each object separately and put it into any of the k states then with very little effort you can prepare any of the $N = k^n$ different configurations of the combined systems. Let us put $k = 2$ and refer to each object of this type as a physical bit. We label the two states of a physical bit as 0 and 1. Any collection of n physical bits can be prepared in $N = 2^n$ different configurations which can be used to store N messages, N binary strings or N different numbers.

Suppose the two states in the physical bit are separated by the energy difference E_0 then a preparation of any particular configuration will cost not more than $E = E_0 n = E_0 \log N$ units of energy (the log is taken to the base 2). If you choose to encode N configurations into one chunk of matter, say a single harmonic oscillator with the interstate energy separation E_0, then, in the worst case, one has to use $E = E_0 N$ units of energy (e.g. to go from the ground state labelled as 0 to the most excited state labelled as N). For large N this gives an exponential gap in the energy expenditure between the binary encoding using physical bits and the so-called unary encoding using a harmonic oscillators.

One can, of course, try to switch from harmonic oscillators to objects which have a finite spread in the energy spectrum. For example, if one wants to use the energy states of the hydrogen atom to encode any number of configurations then one is guaranteed not to spend more than $E_0 = 13.6$eV (otherwise the atoms is ionised). The snag is that in this case some of the electronic states will be separated by the energy difference of the order of E_0/N and to drive the system selectively from one state to another one has to tune into the frequency $E_0/\hbar N$ which requires a sufficiently long wavepacket (so that the frequency is well defined) and consequently the interaction time of the order $N(\hbar/E_0)$. Thus we have to trade energy for time. It turns out that whichever way we try to map N configurations into a single chunk of matter we end up depleting our physical

W. Dieter Heiss (Ed.): LNP 587, pp. 47–76, 2002.
© Springer-Verlag Berlin Heidelberg 2002

resources, such as energy, time, space, at a much greater rate than in the case when we use subsystems.

This plausibility argument indicates that for an efficient processing of information the nature must be divided into subsystems. Thus in the following we will stick to the simplest subsystems namely physical bits and we will use the binary encoding, i.e. a binary alphabet $\{0,1\}$. An ordered collection of symbols from $\{0,1\}$ is called a string.

2 Qubits, Gates and Networks

Consider the two binary strings,

$$011, \tag{1}$$

$$111. \tag{2}$$

The first one can represent, for example, the number 3 (in binary) and the second one the number 7. In general, three physical bits can be prepared in $2^3 = 8$ different configurations that can represent, for example, the integers from 0 to 7. However, a register composed of three classical bits can store only one number at a given moment of time. Enter qubits and quantum registers:

A *qubit* is a quantum system in which the Boolean states 0 and 1 are represented by a prescribed pair of normalised and mutually orthogonal quantum states labeled as $\{|0\rangle, |1\rangle\}$ [1]. The two states form a 'computational basis' and any other (pure) state of the qubit can be written as a superposition $\alpha|0\rangle + \beta|1\rangle$ for some α and β such that $|\alpha|^2 + |\beta|^2 = 1$. A qubit is typically a microscopic system, such as an atom, a nuclear spin, or a polarised photon. A collection of n qubits is called a *quantum register* of size n.

We shall assume that information is stored in the registers in binary form. For example, the number 6 is represented by a register in state $|1\rangle \otimes |1\rangle \otimes |0\rangle$. In more compact notation: $|a\rangle$ stands for the tensor product $|a_{n-1}\rangle \otimes |a_{n-2}\rangle \ldots |a_1\rangle \otimes |a_0\rangle$, where $a_i \in \{0,1\}$, and it represents a quantum register prepared with the value $a = 2^0 a_0 + 2^1 a_1 + \ldots + 2^{n-1} a_{n-1}$. There are 2^n states of this kind, representing all binary strings of length n or numbers from 0 to $2^n - 1$, and they form a convenient computational basis. In the following $a \in \{0,1\}^n$ (a is a binary string of length n) implies that $\big| a \big\rangle$ belongs to the computational basis.

Thus a quantum register of size three can store individual numbers such as 3 or 7,

$$|0\rangle \otimes |1\rangle \otimes |1\rangle \equiv |011\rangle \equiv |3\rangle, \tag{3}$$

$$|1\rangle \otimes |1\rangle \otimes |1\rangle \equiv |111\rangle \equiv |7\rangle, \tag{4}$$

but, it can also store the two of them simultaneously. For if we take the first qubit and instead of setting it to $|0\rangle$ or $|1\rangle$ we prepare a superposition $1/\sqrt{2}\,(|0\rangle + |1\rangle)$

then we obtain

$$\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes |1\rangle \otimes |1\rangle \equiv \frac{1}{\sqrt{2}}\left(|011\rangle + |111\rangle\right), \tag{5}$$

$$\equiv \frac{1}{\sqrt{2}}\left(|3\rangle + |7\rangle\right). \tag{6}$$

In fact we can prepare this register in a superposition of all eight numbers – it is enough to put each qubit into the superposition $1/\sqrt{2}\left(|0\rangle + |1\rangle\right)$. This gives

$$\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \tag{7}$$

which can also be written in binary as (ignoring the normalisation constant $2^{-3/2}$),

$$|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle. \tag{8}$$

or in decimal notation as

$$|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle, \tag{9}$$

or simply as

$$\sum_{x=0}^{7} |\, x\rangle. \tag{10}$$

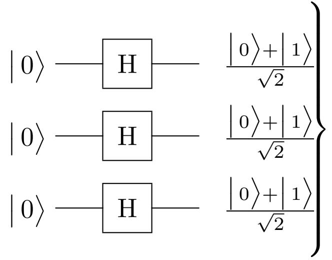
These preparations, and any other manipulations on qubits, have to be performed by unitary operations. A *quantum logic gate* is a device which performs a fixed unitary operation on selected qubits in a fixed period of time and a *quantum network* is a device consisting of quantum logic gates whose computational steps are synchronised in time [2]. The outputs of some of the gates are connected by wires to the inputs of others. The *size* of the network is the number of gates it contains.

The most common quantum gate is the Hadamard gate, a single qubit gate H performing the unitary transformation known as the Hadamard transform. It is defined as

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad |\, x\rangle \;-\!\boxed{\;H\;}\!-\; (-1)^x |\, x\rangle + |\, 1-x\rangle \quad .$$

The matrix is written in the computational basis $\{|\, 0\rangle, |\, 1\rangle\}$ and the diagram on the right provides a schematic representation of the gate H acting on a qubit in state $|\, x\rangle$, with $x = 0, 1$.

And here is a network, of size three, which affects the Hadamard transform on three qubits:



IN BINARY

$$= \frac{1}{2^{3/2}} \left\{ \begin{matrix} |\, 000\rangle + |\, 001\rangle + |\, 010\rangle + |\, 011\rangle + \\ +|\, 100\rangle + |\, 101\rangle + |\, 110\rangle + |\, 111\rangle \end{matrix} \right\}$$

$$= \frac{1}{2^{3/2}} \left\{ \begin{matrix} |\, 0\rangle + |\, 1\rangle + |\, 2\rangle + |\, 3\rangle + \\ +|\, 4\rangle + |\, 5\rangle + |\, 6\rangle + |\, 7\rangle \end{matrix} \right\}$$

IN DECIMAL .

If they are initially in state $|\,000\rangle$ then the output is the superposition of all eight numbers from 0 to 7.

If the three qubits are initially in some other state from the computational basis then the result is a superposition of all numbers from 0 to 7 but exactly half of them will appear in the superposition with the minus sign, for example,

$$|\,101\rangle \mapsto \frac{1}{2^{3/2}} \left\{ \begin{array}{c} |\,000\rangle - |\,001\rangle + |\,010\rangle - |\,011\rangle + \\ -|\,100\rangle + |\,101\rangle - |\,110\rangle + |\,111\rangle \end{array} \right\}. \tag{11}$$

In general, if we start with a register of size n in some state $y \in \{0,1\}^n$ then

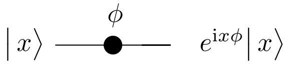
$$|\,y\rangle \mapsto 2^{-n/2} \sum_{x \in \{0,1\}^n} (-1)^{y \cdot x} |\,x\rangle, \tag{12}$$

where the product of $y = (y_{n-1}, \ldots, y_0)$ and $x = (x_{n-1}, \ldots, x_0)$ is taken bit by bit:

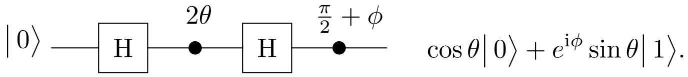
$$y \cdot x = (y_{n-1}x_{n-1} + \ldots + y_1 x_1 + y_0 x_0). \tag{13}$$

We will need another single qubit gate – the phase shift gate ϕ defined as $|\,0\rangle \mapsto |\,0\rangle$ and $|\,1\rangle \mapsto e^{\mathrm{i}\phi}\,|\,1\rangle$, or, in matrix notation,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & e^{\mathrm{i}\phi} \end{pmatrix} \tag{14}$$



The Hadamard gate and the phase gate can be combined to construct the following network (of size four), which generates the most general pure state of a single qubit (up to a global phase),


$|\,0\rangle$
H
2θ
H
$\frac{\pi}{2} + \phi$
$\cos\theta|\,0\rangle + e^{\mathrm{i}\phi}\sin\theta|\,1\rangle$.

(15)

Consequently, the Hadamard and phase gates are sufficient to construct *any* unitary operation on a single qubit.

Thus the Hadamard gates and the phase gates can be used to transform the input state $|0\rangle|0\rangle...|0\rangle$ of the n qubit register into any state of the type $|\Psi_1\rangle$ $|\Psi_2\rangle$... $|\Psi_n\rangle$, where $|\Psi_i\rangle$ is an arbitrary superposition of $|0\rangle$ and $|1\rangle$. These are rather special n-qubit states, called the product states or the separable states. In general, a quantum register of size $n > 1$ can be prepared in states which are not separable – they are known as entangled states. For example, for two qubits ($n = 2$), the state

$$\alpha\ |00\rangle + \beta\ |01\rangle = |0\rangle \otimes (\alpha\ |0\rangle + \beta\ |1\rangle) \tag{16}$$

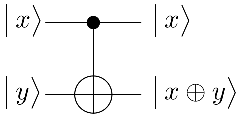
is separable, $|\,\Psi_1\rangle = |0\rangle$ and $|\,\Psi_2\rangle = \alpha\ |0\rangle + \beta\ |1\rangle$, whilst the state

$$\alpha\ |00\rangle + \beta\ |11\rangle \neq |\,\Psi_1\rangle \otimes |\,\Psi_2\rangle \tag{17}$$

is entangled ($\alpha, \beta \neq 0$), because it cannot be written as a tensor product.

In order to entangle two (or more qubits) we have to extend our repertoire of quantum gates to two-qubit gates. The most popular two-qubit gate is the controlled-NOT (C-NOT), also known as the XOR or the measurement gate. It flips the second (target) qubit if the first (control) qubit is $|\,1\rangle$ and does nothing if the control qubit is $|\,0\rangle$. The gate is represented by the unitary matrix

$$C = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \end{pmatrix}, \tag{18}$$



where $x, y = 0$ or 1 and $\oplus$ denotes XOR or addition modulo 2. If we apply the C-NOT to Boolean data in which the target qubit is $|0\rangle$ and the control is either $|0\rangle$ or $|1\rangle$ then the effect is to leave the control unchanged while the target becomes a copy of the control, i.e.

$$|x\rangle|0\rangle \mapsto |x\rangle|x\rangle \quad for \quad x = 0, 1. \tag{19}$$

One might suppose that this gate could also be used to copy superpositions such as $|\varPsi\rangle = \alpha\ |0\rangle + \beta\ |1\rangle$, so that

$$|\varPsi\rangle|0\rangle \mapsto |\varPsi\rangle|\varPsi\rangle \tag{20}$$

for any $|\varPsi\rangle$. This is not so! The unitarity of the C-NOT requires that the gate turns superpositions in the control qubit into *entanglement* of the control and the target. If the control qubit is in a superposition state $|\varPsi\rangle = \alpha|0\rangle + \beta|1\rangle$, $(\alpha, \beta \neq 0)$, and the target in $|0\rangle$ then the C-NOT generates the entangled state

$$(\alpha|0\rangle + \beta|1\rangle)\,|0\rangle \mapsto \alpha|00\rangle + \beta|11\rangle. \tag{21}$$

Let us notice in passing that it is impossible to construct a universal quantum cloning machine effecting the transformation in (20), or even the more general

$$|\varPsi\rangle|0\rangle|W\rangle \mapsto |\varPsi\rangle|\varPsi\rangle|W'\rangle \tag{22}$$

where $|W\rangle$ refers to the state of the rest of the world and $|\varPsi\rangle$ is *any* quantum state [3]. To see this take any two normalised states $|\varPsi\rangle$ and $|\varPhi\rangle$ which are non-identical ($|\langle\varPhi|\varPsi\rangle| \neq 1$) and non-orthogonal ($\langle\varPhi|\varPsi\rangle \neq 0$), and run the cloning machine,

$$|\varPsi\rangle|0\rangle|W\rangle \mapsto |\varPsi\rangle|\varPsi\rangle|W'\rangle\,, \tag{23}$$

$$|\varPhi\rangle|0\rangle|W\rangle \mapsto |\varPhi\rangle|\varPhi\rangle|W''\rangle\,. \tag{24}$$

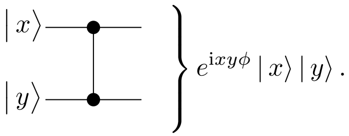
As this must be a unitary transformation which preserves the inner product hence we must require

$$\langle\varPhi|\varPsi\rangle = \langle\varPhi|\varPsi\rangle^2\langle W^{'}|W^{''}\rangle \tag{25}$$

and this can only be satisfied when $|\langle\varPhi|\varPsi\rangle| = 0$ or 1, which contradicts our assumptions. Thus states of qubits, unlike states of classical bits, cannot be faithfully cloned. This leads to interesting applications, quantum cryptography being one such.

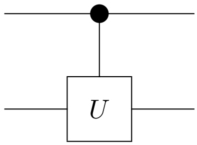
Another common two-qubit gate is the controlled phase shift gate $B(\phi)$ defined as

$$B(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{\mathrm{i}\phi} \end{pmatrix}, \qquad (26)$$



Again, the matrix is written in the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ and the diagram on the right shows the structure of the gate.

More generally, these various 2-qubit controlled gates are all of the form controlled-U, for some single-qubit unitary transformation U. The controlled-U gate applies the identity transformation to the auxiliary (lower) qubit when the control qubit is in state $|0\rangle$ and applies an arbitrary prescribed U when the control qubit is in state $|1\rangle$. The gate maps $|0\rangle|y\rangle$ to $|0\rangle|y\rangle$ and $|1\rangle|y\rangle$ to $|1\rangle(U|y\rangle)$, and is graphically represented as



The Hadamard gate, all phase gates, and the C-NOT, form an infinite *universal set of gates*, i.e. if the C-NOT gate as well as the Hadamard and all phase gates are available then any n-qubit unitary operation can be simulated exactly with $O(4^n n)$ such gates [4]. (Here and in the following we use the asymptotic notation – $O(T(n))$ means bounded above by $c\,T(n)$ for some constant $c > 0$ for sufficiently large n.) This is not the only universal set of gates. In fact, almost any gate which can entangle two qubits can be used as a universal gate [6,8]. Mathematically, an elegant choice is a pair of the Hadamard and the controlled-V (C-V) where V is described by the unitary matrix

$$V = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix}. \qquad (27)$$

The two gates form a finite universal set of gates – networks containing only a finite number of these gates can approximate any unitary transformation on two (and more) qubits. More precisely, if U is any two-qubit gate and $\varepsilon > 0$ then there exists a quantum network of size $O(\log^d(1/\varepsilon))$ (where d is a constant) consisting of only H and C-V gates which computes a unitary operation U' that is within distance ε from U. The metric is induced by the Euclidean norm – we say that U' is within distance ε from U if there exists a unit complex number λ (phase factor) such that $||U - \lambda U'|| \leq \varepsilon$. Thus if U' is substituted for U in a quantum network then the final state $\sum_x \alpha'_x\,|x\rangle$ approximates the final state of the original network $\sum_x \alpha_x\,|x\rangle$ as follows: $\sqrt{\sum_x |\lambda\alpha'_x - \alpha_x|^2} \leq \varepsilon$. The probability of any specified measurement outcome on the final state is affected by at most ε.

A *quantum computer* will be viewed here as a quantum network (or a family of quantum networks) and quantum computation is defined as a unitary evolution of the network which takes its initial state "input" into some final state "output". We have chosen the network model of computation, rather than Turing machines, because it is relatively simple and easy to work with and because it is much more relevant when it comes to physical implementation of quantum computation.

3 Quantum Arithmetic and Function Evaluations

Let us now describe how quantum computers actually compute, how they add and multiply numbers, and how they evaluate Boolean functions by means of unitary operations. Here and in the following we will often use the modular arithmetic [9]. Recall that

$$a \bmod b \tag{28}$$

denotes the remainder obtained by dividing integer b into integer a, which is always a number less than b. Basically $a = b \bmod n$ if $a = b + kn$ for some integer k. This is expressed by saying that a is *congruent* to b modulo n or that b is the *residue* of a modulo n. For example, $1 \bmod 7 = 8 \bmod 7 = 15 \bmod 7 = 50 \bmod 7 = 1$. Modular arithmetic is commutative, associative, and distributive.

$$\begin{aligned}
(a \pm b) \bmod n &= ((a \bmod n) \pm (b \bmod n)) \bmod n\,, && (29)\\
(a \times b) \bmod n &= ((a \bmod n) \times (b \bmod n)) \bmod n\,, && (30)\\
(a \times (b + c)) \bmod n &= (((ab) \bmod n + ((ac) \bmod n)) \bmod n\,. && (31)
\end{aligned}$$

Thus, if you need to calculate, say, $3^8 \bmod 7$ do not use the naive approach and perform seven multiplications and one huge modular reduction. Instead, perform three smaller multiplications and three smaller reductions,

$$((3^2 \bmod 7)^2 \bmod 7)^2 \bmod 7 = (2^2 \bmod 7)^2 \bmod 7 = 16 \bmod 7 = 2. \tag{32}$$

This kind of arithmetic is ideal for computers as it restricts the range of all intermediate results. For l-bit modulus n, the intermediate results of any addition, subtraction or multiplication will not be more than $2l$ bits long. In quantum registers of size n, addition modulo 2^n is one of the most common operations; for all $x \in \{0,1\}^n$ and for any $a \in \{0,1\}^n$,

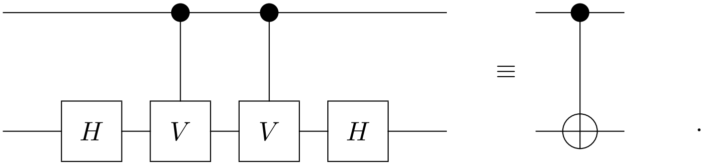
$$\big|\, x \big\rangle \mapsto \big|\, (x + a) \bmod 2^n \big\rangle \tag{33}$$

is a well defined unitary transformation.

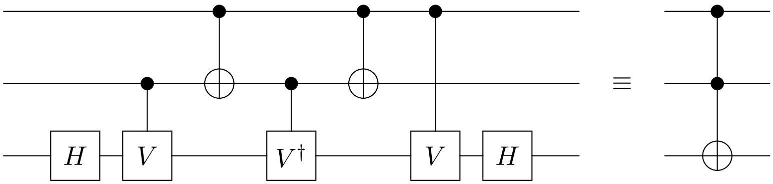
The tricky bit in the modular arithmetic is the inverse operation, and here we need some basic number theory. An integer $a \geq 2$ is said to be *prime* if it is divisible only by 1 and a (we consider only positive divisors). Otherwise, a is called *composite*. The greatest common divisor of two integers a and b is the greatest positive integer d, denoted $d = \gcd(a, b)$, that divides both a and b. Two integers a and b are said to be *coprime* or *relatively prime* if $\gcd(a, b) = 1$. Given two integers a and n that are coprime, it can be shown that there exists an unique integer $d \in \{0, \dots, n-1\}$ such that $ad = 1 \bmod n$ [9]. The integer d is called *inverse modulo* n of a, and denoted a^{-1}. For example, modulo 7 we find that $3^{-1} = 5 \bmod n$, since $3 \times 5 = 15 = 2 \times 7 + 1 = 1 \bmod 7$. This bizarre arithmetic and the notation is due to Karl Friedrich Gauss (1777-1855). It was first introduced in his *Disquistiones Arithmeticae* in 1801.

In quantum computers addition, multiplication, and any other arithmetic operation have to be embedded in unitary evolution. We will stick to the Hadamard and the controlled-V (C-V), and use them as building blocks for all other gates and eventually for quantum adders and multipliers.

If we apply C-V four times we get identity, so any three subsequent applications of C-V give the inverse of C-V, which will be called C-$V^\dagger$. Now, if we have a couple of the C-V gates and a couple of the Hadamard gates we can build the C-NOT as follows



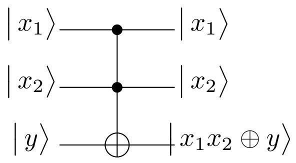
A single qubit operation NOT can be performed via a C-NOT gate if the control qubit is set to $|1\rangle$ and viewed as an auxiliary qubit. This is not to say that we want to do it in practice. The C-NOT gate is much more difficult to build than a single qubit NOT. Right now we are looking into the mathematical structure of quantum Boolean networks and do not care about practicalities. Our two elementary gates also allow us to construct a very useful gate called the controlled-controlled-NOT gate (sc^2-NOT) or the Toffoli gate [10]. The construction is given by the following network,



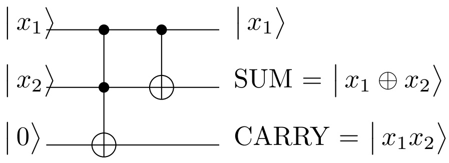
This gate has two control qubits (the top two wires on the diagram) and one target qubit which is negated only when the two controls are in the state $|1\rangle|1\rangle$. The sc^2-NOT gate gives us the logical connectives we need for arithmetic. If the target is initially set to $|0\rangle$ the gate acts as a reversible AND gate - after the gate operation the target becomes the logical AND of the two control qubits.

$$\big|\, x_1, x_2 \big\rangle \big|\, 0 \big\rangle \mapsto \big|\, x_1, x_2 \big\rangle \big|\, x_1 \wedge x_2 \big\rangle \tag{34}$$

Once we have in our repertoire operations such as NOT, AND, and C-NOT, all of them implemented as unitary operations, we can, at least in principle, evaluate any Boolean function $\{0,1\}^n \to \{0,1\}^m$ which map n bits of input into m bits of output. A simple concatenation of the Toffoli gate and the C-NOT gives a simplified quantum adder, shown below, which is a good starting point for constructing full adders, multipliers and more elaborate networks.



TOFFOLI GATE



QUANTUM ADDER

We can view the Toffoli gate and the evolution given by (34) as a quantum implementation of a Boolean function $f : \{0,1\}^2 \to \{0,1\}$ defined by $f(x_1, x_2) = x_1 \wedge x_2$. The operation AND is not reversible, so we had to embed it in the reversible operation sc^2-NOT. If the third bit is initially set to 1 rather than 0 then the value of $x_1 \wedge x_2$ is negated. In general we write the action of the Toffoli gate as the function evaluation,

$$\big|\, x_1, x_2 \big\rangle \big|\, y \big\rangle \mapsto \big|\, x_1, x_2 \big\rangle \big|\, (y + (x_1 \wedge x_2)) \bmod 2 \big\rangle. \tag{35}$$

This is how we compute any Boolean function $\{0,1\}^n \to \{0,1\}^m$ on a quantum computer. We require at least two quantum registers; the first one, of size n, to store the arguments of f and the second one, of size n, to store the values of f. The function evaluation is then a unitary evolution of the two registers,

$$|x, y\rangle \mapsto |x, (y + f(x)) \bmod 2^m\rangle\,, \tag{36}$$

for any $y \in \{0,1\}^m$. (In the following, if there is no danger of confusion, we may simplify the notation and omit the mod suffix.)

For example, a network computing $f : \{0,1\}^2 \to \{0,1\}^3$ such that $f(x) = x^2$ acts as follows

$$|00\rangle|000\rangle \mapsto |00\rangle|000\rangle, \qquad |10\rangle|000\rangle \mapsto |10\rangle|100\rangle\,, \tag{37}$$

$$|01\rangle|000\rangle \mapsto |01\rangle|001\rangle, \qquad |11\rangle|000\rangle \mapsto |11\rangle|001\rangle\,, \tag{38}$$

which can be written as

$$|x, 0\rangle \mapsto |x, x^2 \bmod 8\rangle, \tag{39}$$

e.g. $3^2 \bmod 2^2 = 1$ which explains why $|11\rangle|000\rangle \mapsto |11\rangle|001\rangle$.

In fact, for these kind of operations we also need a third register with the so-called working bits which are set to zero at the input and return to zero at the output but which can take non-zero values during the computation.

What makes quantum function evaluation really interesting is its action on a superposition of different inputs x. For example,

$$\sum_x |x, 0\rangle \mapsto \sum_x |x, f(x)\rangle \tag{40}$$

produces $f(x)$ for all x in a single run. The snag is that we cannot get them all from the entangled state $\sum_x |x, f(x)\rangle$ because any bit by bit measurement on the first register will yield one particular value $x' \in \{0,1\}^n$ and the second register will then be found with the value $f(x') \in \{0,1\}^m$.

4 Algorithms and Their Complexity

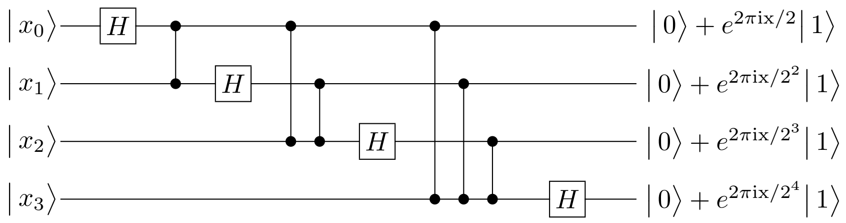
In order to solve a particular problem, computers, be it classical or quantum, follow a precise set of instructions that can be mechanically applied to yield the solution to any given instance of the problem. A specification of this set of instructions is called an algorithm. Examples of algorithms are the procedures taught in elementary schools for adding and multiplying whole numbers; when these procedures are mechanically applied, they always yield the correct result for any pair of whole numbers. Any algorithm can be represented by a family of Boolean networks $(N_1, N_2, N_3, ...)$, where the network N_n acts on all possible input instances of size n bits. Any useful algorithm should have such a family specified by an example network N_n and *a simple rule* explaining how to construct the network N_{n+1} from the network N_n. These are called *uniform* families of networks [11].[1]

The quantum Hadamard transform defined by (12) has a uniform family of networks whose size is growing as n with the number of input qubits. Another good example of a uniform family of networks is the quantum Fourier transform (QFT) [12] defined in the computational basis as the unitary operation

$$\big| y\big\rangle \mapsto 2^{-n/2} \sum_x e^{\mathrm{i}\frac{2\pi}{2^n} yx} \big| x\big\rangle, \tag{41}$$

Suppose we want to *construct* such a unitary evolution of n qubits using our repertoire of quantum logic gates. We can start with a single qubit and notice that in this case the QFT is reduced to applying a Hadamard gate. Then we can take two qubits and notice that the QFT can be implemented with two Hadamard gates and the controlled phase shift $B(\pi)$ in between. Progressing this way we can construct the three qubit QFT and the four qubit QFT, whose network looks like this:

[1] This means that the network model is not a self-contained model of computation. We need an algorithm, a Turing machine, which maps each n into an explicit description of N_n.


$|x_0\rangle$
$|x_1\rangle$
$|x_2\rangle$
$|x_3\rangle$
H
$|0\rangle + e^{2\pi \mathrm{i}x/2}|1\rangle$
$|0\rangle + e^{2\pi \mathrm{i}x/2^2}|1\rangle$
$|0\rangle + e^{2\pi \mathrm{i}x/2^3}|1\rangle$
$|0\rangle + e^{2\pi \mathrm{i}x/2^4}|1\rangle$

$$H\ B(\pi)\ H\ B(\pi/2)B(\pi)\ H\ B(\pi/4)B(\pi/2)B(\pi)\ H$$

N.B. there are three different types of the $B(\phi)$ gate in the network above: $B(\pi)$, $B(\pi/2)$ and $B(\pi/4)$.)

The general case of n qubits requires a trivial extension of the network following the same sequence pattern of gates H and B. The QFT network operating on n qubits contains n Hadamard gates H and $n(n-1)/2$ phase shifts B, in total $n(n+1)/2$ elementary gates.

The big issue in designing algorithms or their corresponding families of networks is the optimal use of physical resources required to solve a problem. Complexity theory is concerned with the inherent cost of computation in terms of some designated elementary operations, memory usage, or network size. An algorithm is said to be fast or efficient if the number of elementary operations taken to execute it increases no faster than a polynomial function of the size of the input. We generally take the input size to be the total number of bits needed to specify the input (for example, a number N requires $\log_2 N$ bits of binary storage in a computer). In the language of network complexity – an algorithm is said to be *efficient* if it has a uniform and polynomial-size network family ($O(n^d)$ for some constant d) [11]. For example, the quantum Fourier transform can be performed in an efficient way because it has a uniform family of networks whose size grows only as a quadratic function of the size of the input, i.e. $O(n^2)$. Changing from one set of gates to another, e.g. constructing the QFT out of the Hadamard and the controlled-V gates with a prescribed precision ϵ, can only affect the network size by a multiplicative constant which does not affect the quadratic scaling with n. Thus the complexity of the QFT is $O(n^2)$ no matter which set of adequate gates we use. Problems which do not have efficient algorithms are known as hard problems.

Elementary arithmetic operations taught at schools, such as long addition, multiplication or division of n bit numbers require $O(n^2)$ operations. For example, to multiply $x = (x_{n-1}...x_1x_0)$ and $y = (y_{n-1}...y_1y_0)$ we successively multiply y by x_0, x_1 and so on, shift, and then add the result. Each multiplication of y by x_k takes about n single bit operations, the addition of the n products takes of the order of n^2 bit operations, which adds to the total $O(n^2)$ operations. Knowing the complexity of elementary arithmetic one can often assess the complexity of other algorithms. For example, the greatest common divisor of two integers x and $y < x$ can be found using Euclid's algorithm; the oldest nontrivial algorithm which has been known and used since 300 BC.[2] First divide x by y obtaining remainder r_1. Then divide y by r_1 obtaining remainder r_2, then divide r_1 by r_2 obtaining remainder r_3, etc., until the remainder is zero. The last non-zero remainder is $\gcd(x, y)$ because it divides all previous remainders and hence also x and y (it is obvious from the construction that it is the *greatest* common divisor). For example, here is a sequence of remainders (r_j, r_{j+1}) when we apply Euclid's algorithm to compute $\gcd(12378, 3054) = 6$: (12378,3054), (3054,162), (162, 138), (138, 24), (24, 18), (18,6), (6,0). What is the complexity of this algorithm? It is easy to see that the largest of the two numbers is at least halved every two steps, so every two steps we need one bit less to represent the number, and so the number of steps is at most $2n$, where n is the number of bits in the two integers. Each division can be done with at most $O(n^2)$ operations hence the total number of operations is $O(n^3)$.

There are basically three different types of Boolean networks: classical deterministic, classical probabilistic, and quantum. They correspond to, respectively, deterministic, randomised, and quantum algorithms.

Classical deterministic networks are based on logical connectives such as AND, OR, and NOT and are required to always deliver correct answers. If a problem admits a deterministic uniform network family of polynomial size, we say that the problem is in the class P [11].

Probabilistic networks have additional "coin flip" gates which do not have any inputs and emit one uniformly-distributed random bit when executed during a computation. Despite the fact that probabilistic networks may generate erroneous answers they may be more powerful than deterministic ones. A good example is primality testing – given an n-bit number x decide whether or not x is prime. The smallest known uniform deterministic network family that solves this problem is of size $O(n^{d \log \log n})$, which is not polynomially bounded. However, there is a probabilistic algorithm, due to Solovay and Strassen [13], that can solve the same problem with a uniform probabilistic network family of size $O(n^3 \log(1/\epsilon))$, where ϵ is the probability of error. N.B. ϵ does not depend on n and we can choose it as small as we wish and still get an efficient algorithm.

The $\log(1/\epsilon)$ part can be explained as follows. Imagine a probabilistic network that solves a decision problem[3] and that errs with probability smaller than $\frac{1}{2} + \delta$ for fixed $\delta > 0$. If you run r of these networks in parallel (so that the size of the overall network is increased by factor r) and then use the majority voting for the final YES or NO answer your overall probability of error will bounded by $\epsilon = \exp(-\delta^2 r)$. (This follows directly from the Chernoff bound – see for instance, [14]). Hence r is of the order $\log(1/\epsilon)$. If a problem admits such a family of networks then we say the problem is in the class BPP (stands for "bounded-error probabilistic polynomial") [11].

[2] This truly 'classical' algorithm is described in Euclid's *Elements*, the oldest Greek treatise in mathematics to reach us in its entirety. Knuth (1981) provides an extensive discussion of various versions of Euclid's algorithm.

[3] A decision problem is a problem that admits only two answers: YES or NO.

Last but not least we have quantum algorithms, or families of quantum networks, which are more powerful than their probabilistic counterparts. The example here is the factoring problem – given an n-bit number x find a list of prime factors of x. The smallest known uniform probabilistic network family which solves the problem is of size $O(2^{d\sqrt{n \log n}})$. One reason why quantum computation is such a fashionable field today is the discovery, by Peter Shor, of a uniform family of quantum networks of $O(n^2(\log\log n)\log(1/\epsilon))$ in size, that solve the factoring problem [15]. If a problem admits a uniform quantum network family of polynomial size that for any input gives the right answer with probability larger than $\frac{1}{2}+\delta$ for fixed $\delta > 0$ then we say the problem is in the class BQP (stands for "bounded-error quantum probabilistic polynomial"). We have

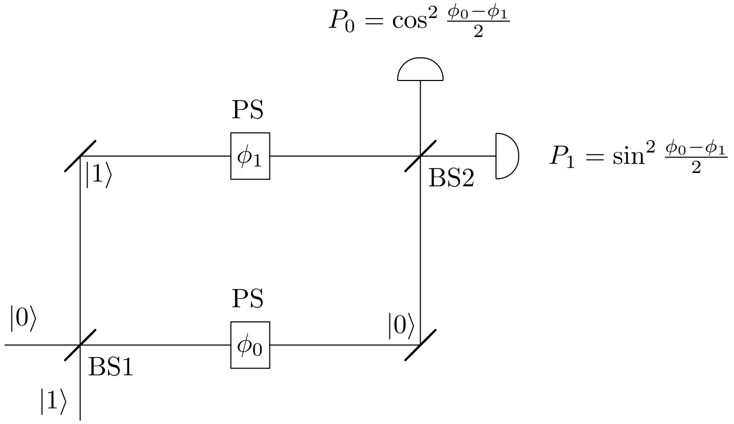
$$P \subseteq BPP \subseteq BQP\,. \tag{42}$$

Quantum networks are potentially more powerful because of multiparticle quantum interference, an inherently quantum phenomenon which makes the quantum theory radically different from any classical statistical theory.

Richard Feynman [16] was the first to anticipate the unusual power of quantum computers. He observed that it appears to be impossible to simulate a general quantum evolution on a classical probabilistic computer in an *efficient* way, i.e. any classical simulation of quantum evolution appears to involve an exponential slowdown in time as compared to the natural evolution since the amount of information required to describe the evolving quantum state in classical terms generally grows exponentially in time. However, instead of viewing this fact as an obstacle, Feynman regarded it as an opportunity. Let us then follow his lead and try to construct a computing device using inherently quantum mechanical effects.

5 From Interferometers to Computers

A single particle interference in the Mach-Zehnder interferometer works as follows. A particle, in this case a photon, impinges on a beam-splitter (BS1), and, with some probability amplitudes, propagates via two different paths to another beam-splitter (BS2) which directs the particle to one of the two detectors. Along each path between the two beam-splitters, is a phase shifter (PS).


$P_0 = \cos^2 \frac{\phi_0 - \phi_1}{2}$
PS
ϕ_1
$|1\rangle$
BS2
$P_1 = \sin^2 \frac{\phi_0 - \phi_1}{2}$
PS
$|0\rangle$
ϕ_0
$|0\rangle$
BS1
$|1\rangle$

If the lower path is labeled as state $|\,0\rangle$ and the upper one as state $|\,1\rangle$ then the particle, initially in path $|\,0\rangle$, undergoes the following sequence of transformations

$$|\,0\rangle \overset{\text{BS1}}{\mapsto} \frac{1}{\sqrt{2}}\,(|\,0\rangle + |\,1\rangle) \overset{\text{PS}}{\mapsto} \frac{1}{\sqrt{2}}(e^{\mathrm{i}\phi_0}\,|\,0\rangle + e^{\mathrm{i}\phi_1}\,|\,1\rangle) \tag{43}$$

$$= e^{\mathrm{i}\frac{\phi_0+\phi_1}{2}}\,\frac{1}{\sqrt{2}}(e^{\mathrm{i}\frac{\phi_0-\phi_1}{2}}\,|\,0\rangle + e^{\mathrm{i}\frac{-\phi_0+\phi_1}{2}}\,|\,1\rangle)$$

$$\overset{\text{BS2}}{\mapsto} e^{\mathrm{i}\frac{\phi_0+\phi_1}{2}}\,(\cos\tfrac{1}{2}(\phi_0-\phi_1)\,|\,0\rangle + \mathrm{i}\sin\tfrac{1}{2}(\phi_0-\phi_1)\,|\,1\rangle), \tag{44}$$

where ϕ_0 and ϕ_1 are the settings of the two phase shifters and the action of the beam-splitters is defined as

$$|\,0\rangle \mapsto \tfrac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle), \quad |\,1\rangle \mapsto \tfrac{1}{\sqrt{2}}(|\,0\rangle - |\,1\rangle). \tag{45}$$

(We have ignored the phase shift in the reflected beam.) The global phase shift $e^{\mathrm{i}\frac{\phi_0+\phi_1}{2}}$ is irrelevant as the interference pattern depends on the difference between the phase shifts in different arms of the interferometer. The phase shifters in the two paths can be tuned to effect any prescribed relative phase shift $\phi = \phi_0 - \phi_1$ and to direct the particle with probabilities

$$P_0 = \cos^2\left(\frac{\phi}{2}\right) = \frac{1}{2}\,(1+\cos\phi) \tag{46}$$

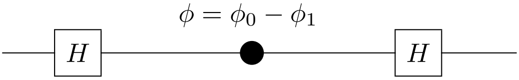
$$P_1 = \sin^2\left(\frac{\phi}{2}\right) = \frac{1}{2}\,(1-\cos\phi) \tag{47}$$

respectively to detectors “0” and “1”.

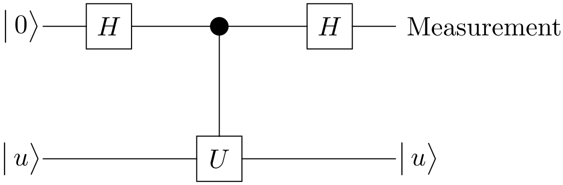
The roles of the three key ingredients in this experiment are clear. The first beam splitter prepares a superposition of possible paths, the phase shifters modify quantum phases in different paths and the second beam-splitter combines all the paths together erasing all information about which path was actually taken by the particle between the two beam-splitters. This erasure is very important as we shall see in a moment.

Needless to say, single particle interference experiments are not restricted to photons. One can go for a different "hardware" and repeat the experiment with electrons, neutrons, atoms or even molecules. When it comes to atoms and molecules both external and internal degrees of freedom can be used.

Although single-particle interference experiments are worth discussing in their own right, here we are only interested in their generic features simply because they are all "isomorphic" and once you know and understand one of them you, at least for our purposes, understand them all (modulo experimental details, of course). Let us now describe any single-particle interference experiment in more general terms. It is very convenient to view this experiment in a diagramatic way as a *quantum network* with three quantum logic gates [17]. The beam-splitters will be now called the Hadamard gates and the phase shifters the phase shift gates. In particular any single particle quantum interference can be represented by the following simple network,



In order to make a connection with a quantum function evaluation let us now describe an alternative construction which simulates the action of the phase shift gate. This construction introduces a phase factor ϕ using a controlled-U gate. The phase shift ϕ is "computed" with the help of an auxiliary qubit in a prescribed state $|\,u\rangle$ such that $U\,|\,u\rangle = e^{\mathrm{i}\phi}\,|\,u\rangle$.



In our example, shown above, we obtain the following sequence of transformations on the two qubits

$$\begin{aligned} |\,0\rangle\,|\,u\rangle \overset{H}{\mapsto} \tfrac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle)\,|\,u\rangle &\overset{sc\text{-}U}{\mapsto} \tfrac{1}{\sqrt{2}}(|\,0\rangle + e^{\mathrm{i}\phi}\,|\,1\rangle)\,|\,u\rangle \\ &\overset{H}{\mapsto} (\cos\tfrac{\phi}{2}\,|\,0\rangle + \mathrm{i}\sin\tfrac{\phi}{2}\,|\,1\rangle)\,|\,u\rangle\,. \end{aligned} \tag{48}$$

We note that the state of the auxiliary qubit $|\,u\rangle$, being an eigenstate of U, is not altered along this network, but its eigenvalue $e^{\mathrm{i}\phi}$ is "kicked back" in front of the $|\,1\rangle$ component in the first qubit. The sequence (48) is the exact simulation of the Mach-Zehnder interferometer and, as we shall see later on, the kernel of quantum algorithms.

Some of the controlled-U operations are special – they represent quantum function evaluations! Indeed, a unitary evolution which computes $f : \{0,1\}^n \mapsto \{0,1\}^m$,

$$\big|\, x\big\rangle\big|\, y\big\rangle \mapsto \big|\, x\big\rangle\big|\, (y + f(x)) \bmod 2^m\big\rangle, \tag{49}$$

is of the controlled-U type. The unitary transformation of the second register, specified by

$$\big|\, y\big\rangle \mapsto \big|\, (y + f(x)) \bmod 2^m\big\rangle, \tag{50}$$

depends on x – the state of the first register. If the initial state of the second register is set to

$$|\, u\rangle = \frac{1}{2^{m/2}} \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} y\right) |y\rangle, \tag{51}$$

by applying the QFT to the state $\big|\, 111...1\big\rangle$, then the function evaluation generates

$$\begin{aligned}
|x\rangle\,|\, u\rangle &= \frac{1}{2^{m/2}}|x\rangle \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} y\right) |y\rangle && (52)\\
&\mapsto \frac{1}{2^{m/2}}|x\rangle \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} y\right) |f(x)+y\rangle && (53)\\
&= \frac{e^{\frac{2\pi \mathrm{i}}{2^m} f(x)}}{2^{m/2}}|x\rangle \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} (f(x)+y)\right) |f(x)+y\rangle && (54)\\
&= \frac{e^{\frac{2\pi \mathrm{i}}{2^m} f(x)}}{2^{m/2}}|x\rangle \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} y\right) |y\rangle && (55)\\
&= e^{\frac{2\pi \mathrm{i}}{2^m} f(x)}|x\rangle\,|\, u\rangle\,, && (56)
\end{aligned}$$

where we have relabelled the summation index in the sum containing 2^m terms

$$\sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} (f(x)+y)\right) |f(x)+y\rangle = \sum_{y=0}^{2^m-1} \exp\left(-\frac{2\pi \mathrm{i}}{2^m} y\right) |y\rangle. \tag{57}$$

Again, the function evaluation effectively introduces the phase factors in front of the $|x\rangle$ terms in the first register.

$$|x\rangle\,|\, u\rangle \mapsto \exp\left(\frac{2\pi \mathrm{i}}{2^m} f(x)\right) |x\rangle\,|\, u\rangle \tag{58}$$

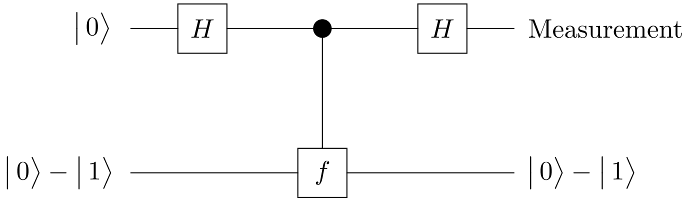
Please notice that the resolution in $\phi(x) = \frac{2\pi}{2^m} f(x)$ is determined by the size m of the second register. For $m = 1$ we obtain $\phi(x) = \pi f(x)$, i.e. the phase factors are $(-1)^{f(x)}$. Let us see how this approach explains the internal working of quantum algorithms.

6 The First Quantum Algorithms

The first quantum algorithms showed advantages of quantum computation without referring to computational complexity measured by the scaling properties of network sizes. The computational power of quantum interference was discovered by counting how many times certain Boolean functions have to be evaluated in order to find the answer to a given problem. Imagine a “black box” (also called an *oracle*) computing a Boolean function and a scenario in which one wants to learn about a given property of the Boolean function but has to pay for each use of the “black box” (often referred to as a *query*). The objective is to minimise number of queries.

Consider, for example, a “black box” computing a Boolean function $f : \{0,1\} \mapsto \{0,1\}$. There are exactly four such functions: two constant functions ($f(0) = f(1) = 0$ and $f(0) = f(1) = 1$) and two “balanced” functions ($f(0) = 0, f(1) = 1$ and $f(0) = 1, f(1) = 0$). The task is to deduce, by queries to the “black box”, whether f is constant or balanced (in other words, whether $f(0)$ and $f(1)$ are the same or different).

Classical intuition tells us that we have to evaluate both $f(0)$ and $f(1)$, which involves evaluating f twice (two queries). We shall see that this is not so in the setting of quantum information, where we can solve this problem with a single function evaluation (one query), by employing an algorithm that has the same mathematical structure as the Mach-Zehnder interferometer. The quantum algorithm that accomplishes this is best represented as the quantum network shown below, where the middle operation is the “black box” representing the function evaluation [17].



The initial state of the qubits in the quantum network is $|\,0\rangle\,(|\,0\rangle - |\,1\rangle)$ (apart from a normalization factor, which will be omitted in the following). After the first Hadamard transform, the state of the two qubits has the form $(|\,0\rangle + |\,1\rangle)(|\,0\rangle - |\,1\rangle)$. To determine the effect of the function evaluation on this state, first recall that, for each $x \in \{0,1\}$,

$$|\,x\rangle\,(|\,0\rangle - |\,1\rangle) \overset{f}{\mapsto} (-1)^{f(x)}\,|\,x\rangle\,(|\,0\rangle - |\,1\rangle). \tag{59}$$

Therefore, the state after the function evaluation is

$$[(-1)^{f(0)}\,|\,0\rangle + (-1)^{f(1)}\,|\,1\rangle](|\,0\rangle - |\,1\rangle)\ . \tag{60}$$

That is, for each x, the $|\,x\rangle$ term acquires a phase factor of $(-1)^{f(x)}$, which corresponds to the eigenvalue of the state of the auxiliary qubit under the action of the operator that sends $|\,y\rangle$ to $|\,y+f(x)\rangle$. The second qubit is of no interest to us any more but the state of the first qubit

$$(-1)^{f(0)}\,|\,0\rangle+(-1)^{f(1)}\,|\,1\rangle \tag{61}$$

is equal either to

$$\pm\left(|\,0\rangle+|\,1\rangle\right), \tag{62}$$

when $f(0)=f(1)$, or

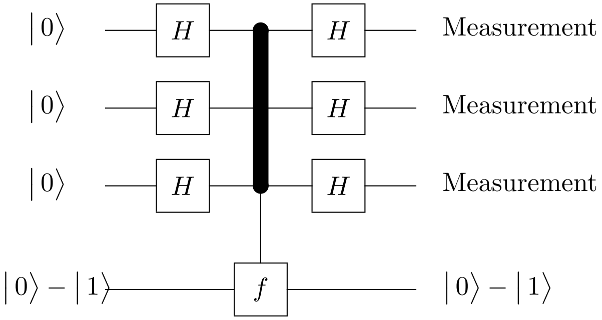
$$\pm\left(|\,0\rangle-|\,1\rangle\right), \tag{63}$$

when $f(0)\neq f(1)$. Hence, after applying the second Hadamard gate the state of the first qubit becomes $|\,0\rangle$ if the function f is constant and $|\,1\rangle$ if the function is balanced! A bit-value measurement on this qubit distinguishes these cases with certainty.

This example [17] is an improved version of the first quantum algorithm proposed by Deutsch [18]. (The original Deutsch algorithm provides the correct answer with probability 50%.) Deutsch's result laid the foundation for the new field of quantum computation, and was followed by several other quantum algorithms.

Deutsch's original problem was subsequently generalised to cover "black boxes" computing Boolean functions $f:\{0,1\}^n\mapsto\{0,1\}$. Assume that, for one of these functions, it is "promised" that it is either constant or balanced (i.e. has an equal number of 0's outputs as 1's), and the goal is to determine which of the two properties the function actually has. How many queries to f are required to do this? Any classical algorithm for this problem would, in the worst-case, require $2^{n-1}+1$ queries before determining the answer with certainty. There is a quantum algorithm that solves this problem with a single evaluation of f.

The algorithm is illustrated by a simple extension of the network which solves Deutsch's problem.



The control register, now composed out of n qubits ($n=3$ in the diagram above), is initially in state $\big|\,00\cdots0\big\rangle$ and an auxiliary qubit in the second register starts and remains in the state $\big|\,0\big\rangle-\big|\,1\big\rangle$.

Stepping through the execution of the network, the state after the first n-qubit Hadamard transform is applied is

$$\sum_x \left| x \right\rangle (\left| 0 \right\rangle - \left| 1 \right\rangle) \,, \tag{64}$$

which, after the function evaluation, is

$$\sum_x (-1)^{f(x)} \left| x \right\rangle (\left| 0 \right\rangle - \left| 1 \right\rangle). \tag{65}$$

Finally, after the last Hadamard transform, the state is

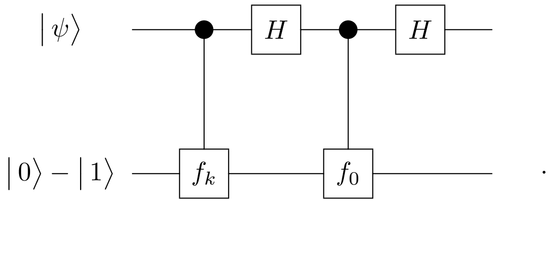
$$\sum_{x,y} (-1)^{f(x)+(x \cdot y)} \left| y \right\rangle (\left| 0 \right\rangle - \left| 1 \right\rangle). \tag{66}$$

Note that the amplitude of $\left| 00 \cdots 0 \right\rangle$ is $\sum_x \frac{(-1)^{f(x)}}{2^n}$ which is $(-1)^{f(0)}$ when f is constant and 0 when f is balanced. Therefore, by measuring the first n qubits, it can be determined with certainty whether f is constant or balanced. The algorithm follows the same pattern as Deutsch's algorithm: the Hadamard transform, a function evaluation, the Hadamard transform (the H-f-H sequence). We recognize it as a generic interference pattern.

7 Quantum Search

The generic H-f-H sequence may be repeated several times. This can be illustrated, for example, with Grover's data base search algorithm [19]. Suppose we are given, as an oracle, a Boolean function f_k which maps $\{0,1\}^n$ to $\{0,1\}$ such that $f_k(x) = \delta_{xk}$ for some k. Our task is to find k. Thus in a set of numbers from 0 to $2^n - 1$ one element has been "tagged" and by evaluating f_k we have to find which one. In order to find k with probability of 50% any classical algorithm, be it deterministic or randomised, will need to evaluate f_k a minimum of 2^{n-1} times. In contrast, a quantum algorithm needs only $O(2^{n/2})$ evaluations.

Unlike the algorithms studied so far, Grover's algorithm consists of *repeated* applications of the *same* unitary transformation many ($O(2^{n/2})$) times. The initial state is chosen to be the one that has equal overlap with each of the computational basis states: $\left| S \right\rangle = 2^{-n/2} \sum_{i=0}^{2^n-1} \left| i \right\rangle$. The operation applied at each individual iteration, referred to as the Grover iterate, can be best represented by the following network:


$\left| \psi \right\rangle$
H
H
$\left| 0 \right\rangle - \left| 1 \right\rangle$
f_k
f_0
.

The components of the network are by now familiar: Hadamard transforms (H) and controlled-f gates. It is important to notice that in drawing the network we have used a shorthand notation: the first register (with the $|\,\psi\rangle$ input) actually consists of n qubits. The Hadamard transform is applied to each of those qubits and the controlled-f gates act on all of them simultaneously. Also, the input to the second register is always $|\,0\rangle - |\,1\rangle$ but the input to the first register, denoted $|\,\psi\rangle$ changes from iteration from iteration, as the calculation proceeds. As usual, the second register will be ignored since it remains constant throughout the computation.

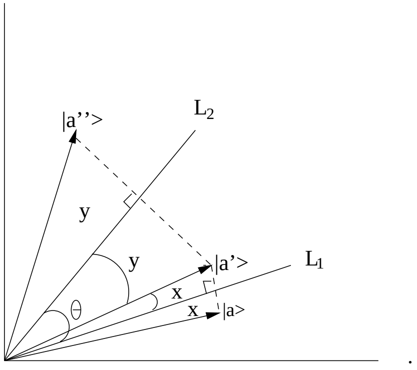
To begin, consider only the controlled-f_k gate. This is just the phase-kickback construction that was introduced in Section 4 but for the specific function f_k. In particular, the transformation does nothing to any basis elements except for $|\,k\rangle$, which goes to $-|\,k\rangle$. Geometrically, this is simply a reflection in the hyperplane perpendicular to $|\,k\rangle$ so let us call it R_k.

Similarly, with respect to the first register only, the controlled-f_0 operation sends $|\,0\rangle$ to $-|\,0\rangle$ and fixes all other basis elements, so it can be written R_0. Now consider the sequence of operations HR_0H. Since $H^2 = I$, we can rewrite the triple as HR_0H^{-1} which is simply R_0 performed in a different basis. More specifically, it is reflection about the hyperplane perpendicular to

$$H|\,0\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |\,x\rangle = |\,S\rangle\,, \tag{67}$$

so we will simply write the triple as R_S.

We can therefore rewrite the Grover iterate in the simple form $G = R_S R_k$. Now, since each reflection is an orthogonal transformation with negative determinant, their composition must be an orthogonal transformation with unit determinant, in other words, a rotation. The question, of course, is which rotation. To find the answer it suffices to consider rotations in the plane spanned by $|\,k\rangle$ and $|\,S\rangle$ since all other vectors are fixed by the Grover iterate. The generic geometrical situation is then illustrated in this diagram:



If the vector $\big|\,a\big\rangle$ is reflected through the line L_1 to produce the vector $\big|\,a'\big\rangle$ and then reflected a second time through line L_2 to produce the vector $\big|\,a''\big\rangle$, then the net effect is a rotation by the total subtended angle between $\big|\,a\big\rangle$ and $\big|\,a''\big\rangle$, which is $2x + 2y = 2(x + y) = 2\theta$.

Therefore, writing $\big|\,k^{\perp}\big\rangle$ and $\big|\,S^{\perp}\big\rangle$ for plane vectors perpendicular to $\big|\,k\big\rangle$ and $\big|\,S\big\rangle$ respectively, the Grover iterate performs a rotation of twice the angle from $\big|\,k^{\perp}\big\rangle$ to $\big|\,S^{\perp}\big\rangle$. Setting, $\sin\phi = \frac{1}{2^{n/2}}$, this is easily seen to be a rotation by

$$2(3\frac{\pi}{2} - \phi) = \pi - 2\phi \bmod 2\pi. \tag{68}$$

Thus, up to phases, the Grover iterate rotates the state vector by an angle 2ϕ towards the desired solution $\big|\,k\big\rangle$. Normally, the initial state for the first register is chosen to be $\big|\,S\big\rangle$. Since this initial state $\big|\,S\big\rangle$ is already at an angle ϕ to $\big|\,k\big\rangle$, the iterate should be repeated m times, where

$$(2m + 1)\phi \approx \frac{\pi}{2}, \tag{69}$$

giving

$$m \approx \frac{\pi}{4\phi} - \frac{1}{4} \tag{70}$$

to get a probability of success bounded below by $\cos^2(2\phi)$, which goes to 1 as $n \mapsto \infty$. For large n, $\frac{1}{2^{n/2}} = \sin\phi \approx \phi$, so

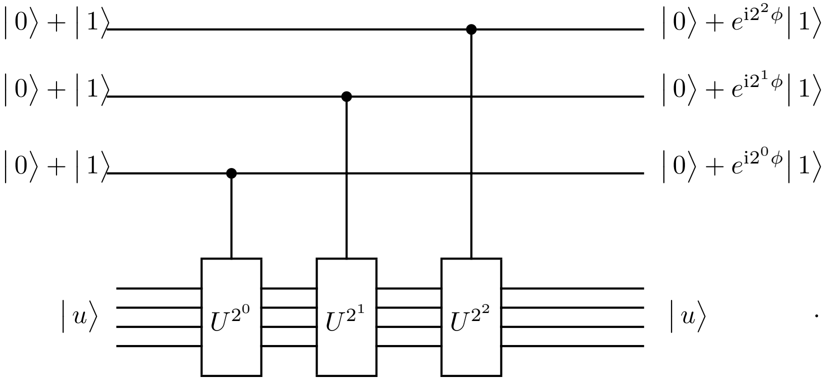
$$m \approx \frac{\pi}{4}\frac{1}{2^{n/2}}. \tag{71}$$

This is an astounding result: any search of an unstructured database can be performed in time proportional to the square-root of the number of entries in the database. Subsequent work extended the result to searches for multiple items [20], searches of structured databases [21], and many other situations. Also, Zalka [22], Boyer et. al. [20] and others have demonstrated that Grover's algorithm is optimal, in the sense that any other quantum algorithm for searching an unstructured database must take time at least $O(2^{n/2})$.

8 Optimal Phase Estimation

Query models of quantum computation provided a natural setting for subsequent discoveries of "real quantum algorithms". The most notable example is Shor's quantum factoring algorithm [15] which evolved from the order-finding problem, which was originally formulated in the language of quantum queries. Following our "interferometric approach" we will describe this algorithm in the terms of multiparticle quantum interferometry. We start with a simple eigenvalue or phase estimation problem.

Suppose that U is any unitary transformation on m qubits and $|\,u\rangle$ is an eigenvector of U with eigenvalue $e^{\mathrm{i}\phi}$ and consider the following scenario. We do not explicitly know U or $|u\rangle$ or $e^{\mathrm{i}\phi}$, but instead we are given devices that perform controlled-U, controlled-U^{2^1}, controlled-U^{2^2} and so on until we reach controlled-$U^{2^{n-1}}$. Also, assume that we are given a single preparation of the state $|u\rangle$. Our goal is to obtain an n-bit estimator of ϕ. We start by constructing the following network,


$|0\rangle + |1\rangle$
$|0\rangle + |1\rangle$
$|0\rangle + |1\rangle$
$|u\rangle$
U^{2^0}
U^{2^1}
U^{2^2}
$|0\rangle + e^{\mathrm{i}2^2\phi}|1\rangle$
$|0\rangle + e^{\mathrm{i}2^1\phi}|1\rangle$
$|0\rangle + e^{\mathrm{i}2^0\phi}|1\rangle$
$|u\rangle$

.

The second register of m qubits is initially prepared in state $|u\rangle$ and remains in this state after the computation, whereas the first register of n qubits evolves into the state,

$$(|0\rangle + e^{\mathrm{i}2^{n-1}\phi}|1\rangle)(|0\rangle + e^{\mathrm{i}2^{n-2}\phi}|1\rangle)\cdots(|0\rangle + e^{\mathrm{i}\phi}|1\rangle) = \sum_{y=0}^{2^n-1} e^{2\pi\mathrm{i}\frac{\phi y}{2^n}}|y\rangle. \quad (72)$$

Consider the special case where $\phi = 2\pi x/2^n$ for $x = \sum_{i=0}^{n-1} 2^i x_i$, and recall the quantum Fourier transform (QFT) introduced in Section 2. The state which gives the binary representation of x, namely, $|x_{n-1}\cdots x_0\rangle$ (and hence ϕ) can be obtained by applying the inverse of the QFT, that is by running the network for the QFT in the backwards direction (consult the diagram of the QFT). If x is an n-bit number this will produce the exact value ϕ.

However, ϕ does not have to be a fraction of a power of two (and may not even be a rational number). For such a ϕ, it turns out that applying the inverse of the QFT produces the best n-bit approximation of ϕ with probability at least $4/\pi^2 \approx 0.405$.

To see why this is so, let us write $\phi = 2\pi(a/2^n + \delta)$, where $a = (a_{n-1}\dots a_0)$ is the best n-bit estimate of $\frac{\phi}{2\pi}$ and $0 < |\delta| \le 1/2^{n+1}$. Applying the inverse QFT to the state in (72) now yields the state

$$\frac{1}{2^n}\sum_{x=0}^{2^n-1}\sum_{y=0}^{2^n-1} e^{\frac{2\pi\mathrm{i}}{2^n}(a-x)y} e^{2\pi\mathrm{i}\delta y}|x\rangle \quad (73)$$

and the coefficient in front of $|x = a\rangle$ in the above is the geometric series

$$\frac{1}{2^n}\sum_{y=0}^{2^n-1}(e^{2\pi\mathrm{i}\delta})^y = \frac{1}{2^n}\left(\frac{1-(e^{2\pi\mathrm{i}\delta})^{2^n}}{1-e^{2\pi\mathrm{i}\delta}}\right). \quad (74)$$

Since $|\delta| \leq \frac{1}{2^{n+1}}$, it follows that $2^n|\delta| \leq 1/2$, and using the inequality $2z \leq \sin \pi z \leq \pi z$ holding for any $z \in [0, 1/2]$, we get $|1 - e^{2\pi \mathrm{i}\delta 2^n}| = 2|\sin(\pi\delta 2^n)| \geq 4|\delta|2^n$. Also, $|1-e^{2\pi\mathrm{i}\delta}| = 2|\sin \pi\delta| \leq 2\pi\delta$. Therefore, the probability of observing $a_{n-1} \cdots a_0$ when measuring the state is

$$\left|\frac{1}{2^n}\left(\frac{1-(e^{2\pi\mathrm{i}\delta})^{2^n}}{1-e^{2\pi\mathrm{i}\delta}}\right)\right|^2 \geq \left(\frac{1}{2^n}\left(\frac{4\delta 2^n}{2\pi\delta}\right)\right)^2 = \frac{4}{\pi^2}, \tag{75}$$

which proves our assertion. In fact, the probability of obtaining the best estimate can be made $1 - \delta$ for any $0 < \delta < 1$, by creating the state in (72) but with $n + O(\log(1/\delta))$ qubits and rounding the answer off to the nearest n bits [17].

9 Periodicity and Quantum Factoring

Amazingly, the application of optimal phase estimation to a very particular unitary operator will allow us to factor integers efficiently. In fact, it will allow us to solve a more general class of problems related to the periodicity of certain integer functions.

Let N be an m-bit integer, and let a be an integer smaller than N, and coprime to N. Define a unitary operator U_a acting on m qubits such that for all $y < N$

$$\left|\, y\right\rangle \mapsto U_a \left|\, y\right\rangle = \left|\, ay \bmod N\right\rangle. \tag{76}$$

This unitary operation can be called multiplication by a modulo N. Since a is coprime to N, as discussed in Section 2, there exists a least strictly positive r such that $a^r = 1 \bmod N$. This r is called the *order* of a modulo N. Equivalently, r is the period of the function $f(x) = a^x \bmod N$, i.e. the least $r > 0$ such that $f(x) = f(x+r)$ for all x. We are after the optimal n-bit estimate of this period, given some specified precision n.

Now let the vectors $\left|\, u_k\right\rangle$ ($k \in \{1, \ldots, r\}$) be defined by

$$\left|\, u_k\right\rangle = r^{-1/2} \sum_{j=0}^{r-1} e^{-\frac{2\pi\mathrm{i}kj}{r}} \left|\, a^j \bmod N\right\rangle. \tag{77}$$

It is easy to check [23] that for each $k \in \{1, \ldots, r\}$, $\left|\, u_k\right\rangle$ is an eigenvector with eigenvalue $e^{2\pi\mathrm{i}\frac{k}{r}}$ of the modular multiplication operator U_a defined above.

It is important to observe that one can efficiently construct a quantum network for controlled multiplication modulo some number N. Moreover, for any j, it is possible to efficiently implement a controlled-$U_a^{2^j}$ gate [24,25]. Therefore, we can apply the techniques for optimal phase estimation discussed in Section 7. For any $k \in \{1, \ldots, r\}$, given the state $\left|\, u_k\right\rangle$ we can obtain the best n-bit approximation to $\frac{k}{r}$. This is tantamount to determining r itself. Unfortunately, there is a complication.

Our task is: given an m bit long number N and randomly chosen $a < N$ coprime with N, find the order of a modulo N. The problem with the above method is that we are not aware of a straightforward efficient way to prepare any of the states $|\, u_k\rangle$. However, the state

$$|\, 1\rangle = r^{-1/2} \sum_{k=1}^{r} |\, u_k\rangle \tag{78}$$

is most definitely an easy state to prepare.

If we start with $\big|\, 1\big\rangle$ in place of the eigenvector $\big|\, u_k\big\rangle$, apply the phase estimation network and measure the first register bit by bit we will obtain n binary digits of x such that, with probability exceeding $4/\pi^2$, $\frac{x}{2^n}$ is the best n-bit estimate of $\frac{k}{r}$ for a randomly chosen k from $\{1,\ldots,r\}$. The question is: given x how to compute r? Let us make few observations:

- *k/r is unique, given x.*
  Value $x/2^n$, being the n-bit estimate, differs by at most $1/2^n$ from k/r. Hence, as long as $n > 2m$, the n bit estimate x determines a unique value of $\frac{k}{r}$ since r is an m-bit number.
- *Candidate values for k/r are all convergents to $x/2^m$.*
  For any real number θ, there is a unique sequence of special rationals $(\frac{p_n}{q_n})_{n\in\mathbf{N}}$ ($\gcd(p_n, q_n) = 1$) called the *convergents* to θ that tend to θ as n grows. A theorem [9] states that if p and q are integers with $\left|\theta - \frac{p}{q}\right| < \frac{1}{2q^2}$ then p/q is a convergent to θ. Since we have $\frac{1}{2^n} \leq \frac{1}{2(2^m)^2} \leq \frac{1}{2r^2}$, this implies $\left|\frac{x}{2^n} - \frac{k}{r}\right| < \frac{1}{2r^2}$ and k/r is a convergent to $x/2^n$.
- *Only one convergent is eligible.*
  It is easy to show that there is at most one fraction a/b satisfying both $b \leq r$ and $\left|\frac{x}{2^n} - \frac{a}{b}\right| < \frac{1}{2r^2}$.

Convergents can be found efficiently using the well-known *continued fraction* method [9]. Thus we employ continued fractions and our observations above to find a fraction a/b such that $b \leq 2^m$ and $\left|\frac{x}{2^n} - \frac{a}{b}\right| < \frac{1}{2^n}$. We get the rational k/r, and $k = a, r = b$, provided k and r are coprime. For randomly chosen k, this happens with probability greater than or equal to $1/\ln r$ [26].

Finally, we show how order-finding can be used to factor a composite number N. Let a be a randomly chosen positive integer smaller than N such that $\gcd(a, N) = 1$. Then the order of a modulo N is defined, and we can find it efficiently using the above algorithm. If r is even, then we have:

$$a^r \qquad = 1 \bmod N\,, \tag{79}$$

$$\Leftrightarrow \quad (a^{r/2})^2 - 1^2 = 0 \bmod N\,, \tag{80}$$

$$\Leftrightarrow \quad (a^{r/2} - 1)(a^{r/2} + 1) = 0 \bmod N\,. \tag{81}$$

The product $(a^{r/2} - 1)(a^{r/2} + 1)$ must be some multiple of N, so unless $a^{r/2} = \pm 1 \bmod N$ at least one of terms must have a nontrivial factor in common with N. By computing the greatest common divisor of this term and N, one gets a non-trivial factor of N.

Furthermore, if N is odd with prime factorisation

$$N = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}\,, \tag{82}$$

then it can be shown [26] that if $a < N$ is chosen at random such that $\gcd(a, N) = 1$ then the probability that its order modulo N is even and that $a^{r/2} \neq \pm 1 \bmod N$ is:

$$\Pr(r \text{ is even } \textsc{and } a^{r/2} \neq \pm 1 \bmod N) \geq 1 - \frac{1}{2^{s-1}}. \tag{83}$$

Thus, combining our estimates of success at each step, with probability greater than or equal to

$$\frac{4}{\pi^2}\frac{1}{\ln r}\left(1 - \frac{1}{2^{s-1}}\right) \geq \frac{2}{\pi^2}\frac{1}{\ln N} \tag{84}$$

we find a factor of N[4]. (Here we have used that N is composite and $r < N$.) If N is $\log N = n$ bits long then by repeating the whole process $O(n)$ times, or by a running $O(n)$ computations in parallel by a suitable extension of a quantum factoring network, we can then guarantee that we will find a factor of N with a fixed probability greater than $\frac{1}{2}$. This, and the fact that the quantum network family for controlled multiplication modulo some number is uniform and of size $O(n^2)$, tells us that factoring is in the complexity class BQP.

But why should anybody care about efficient factorisation?

10 Cryptography

Human desire to communicate secretly is at least as old as writing itself and goes back to the beginnings of our civilisation. Methods of secret communication were developed by many ancient societies, including those of Mesopotamia, Egypt, India, and China, but details regarding the origins of cryptology[5] remain unknown [27].

Originally the security of a cryptosystem or a cipher depended on the secrecy of the entire encrypting and decrypting procedures; however, today we use ciphers for which the algorithm for encrypting and decrypting could be revealed to anybody without compromising their security. In such ciphers a set of specific parameters, called a *key*, is supplied together with the plaintext as an input to the encrypting algorithm, and together with the cryptogram as an input to the decrypting algorithm [28]. This can be written as

$$\hat{E}_k(P) = C, \text{ and conversely, } \hat{D}_k(C) = P, \tag{85}$$

where P stands for plaintext, C for cryptotext or cryptogram, k for cryptographic key, and $\hat{E}$ and $\hat{D}$ denote an encryption and a decryption operation respectively.

The encrypting and decrypting algorithms are publicly known; the security of the cryptosystem depends entirely on the secrecy of the key, and this key must consist of a *randomly chosen*, sufficiently long string of bits. Probably the best

[4] N.B. by (83), the method fails if N is a prime power, $N = p^\alpha$, but prime powers can be efficiently recognised and factored by classical means.

[5] The science of secure communication is called cryptology from Greek *kryptos* hidden and *logos* word. Cryptology embodies cryptography, the art of code-making, and cryptanalysis, the art of code-breaking.

way to explain this procedure is to have a quick look at the Vernam cipher, also known as the one-time pad [29].

If we choose a very simple digital alphabet in which we use only capital letters and some punctuation marks such as

A	B	C	D	E	...	...	X	Y	Z		?	,	.
00	01	02	03	04	...	...	23	24	25	26	27	28	29

we can illustrate the secret-key encrypting procedure by the following simple example (we refer to the dietary requirements of 007):

S	H	A	K	E	N		N	O	T		S	T	I	R	R	E	D
18	07	00	10	04	13	26	13	14	19	26	18	19	08	17	17	04	03
15	04	28	13	14	06	21	11	23	18	09	11	14	01	19	05	22	07
03	11	28	23	18	19	17	24	07	07	05	29	03	09	06	22	26	10

In order to obtain the cryptogram (sequence of digits in the bottom row) we add the plaintext numbers (the top row of digits) to the key numbers (the middle row), which are randomly selected from between 0 and 29, and take the remainder after division of the sum by 30, that is we perform addition modulo 30. For example, the first letter of the message "S" becomes a number "18"in the plaintext, then we add $18 + 15 = 33$; $33 = 1 \times 30 + 3$, therefore we get 03 in the cryptogram. The encryption and decryption can be written as $P_i + k_i \pmod{30} = C_i$ and $C_i - k_i \pmod{30} = P_i$ respectively for the symbol at position i.

The cipher was invented in 1917 by the American AT&T engineer Gilbert Vernam. It was later shown, by Claude Shannon [30], that as long as the key is truly random, has the same length as the message, and is never reused then the one-time pad is perfectly secure. So, if we have a truly unbreakable system, what is wrong with classical cryptography?

There is a snag. It is called *key distribution.* Once the key is established, subsequent communication involves sending cryptograms over a channel, even one which is vulnerable to total passive eavesdropping (e.g. public announcement in mass-media). This stage is indeed secure. However in order to establish the key, two users, who share no secret information initially, must at a certain stage of communication use a reliable and a very secure channel. Since the interception is a set of measurements performed by an eavesdropper on this channel, however difficult this might be from a technological point of view, *in principle* any classical key distribution can always be passively monitored, without the legitimate users being aware that any eavesdropping has taken place.

In the late 1970s Whitfield Diffie and Martin Hellman [31] proposed an interesting solution to the key distribution problem. It involved two keys, one public key π for encryption and one private key κ for decryption:

$$\hat{E}_\pi(P) = C, \text{ and } \hat{D}_\kappa(C) = P. \tag{86}$$

In these systems users do not need to share any private key before they start sending messages to each other. Every user has his own two keys; the public key is publicly announced and the private key is kept secret. Several public-key cryptosystems have been proposed since 1976; here we concentrate our attention on the most popular one namely the RSA [32]. In fact the techniques were first discovered at CESG in the early 1970s by James Ellis, who called them "Non-Secret Encryption" [33]. In 1973, building on Ellis' idea, C. Cocks designed what we now call RSA [34], and in 1974 M. Williamson proposed what is essentially known today as the Diffie-Hellman key exchange protocol.

Suppose that Alice wants to send an RSA encrypted message to Bob. The RSA encryption scheme works as follows:

Key generation. Bob picks randomly two distinct and large prime numbers p and q. We denote $n = pq$ and $\phi = (p-1)(q-1)$. Bob then picks a random integer $1 < e < \phi$ that is coprime with $\phi(\gcd(e, \phi) = 1)$, and computes the inverse d of e modulo ϕ. This inversion can be achieved efficiently using for instance the extended Euclidean algorithm for the greatest common divisor [9]. Bob's private key is $\kappa = d$ and his public key is $\pi = (e, n)$

Encryption. Alice obtains Bob's public key $\pi = (e, n)$ from some sort of yellow pages or an RSA public key directory. Alice then writes her message as a sequence of numbers using, for example, our digital alphabet. This string of numbers is subsequently divided into blocks such that each block when viewed as a number P satisfies $P \leq n$. Alice encrypts each P as

$$C = \hat{E}_\pi(P) = P^e \bmod n \tag{87}$$

and sends the resulting cryptogram to Bob.

Decryption. Receiving the cryptogram C, Bob decrypts it by calculating

$$\hat{D}_\kappa(C) = C^d \bmod n = P \tag{88}$$

where the last equality will be proved shortly.

The mathematics behind the RSA is a lovely piece of number theory which goes back to the XVI century when a French lawyer Pierre de Fermat discovered that if a prime p and a positive integer a are coprime, then

$$a^{p-1} = 1 \bmod p. \tag{89}$$

A century later, Leonhard Euler found the more general relation

$$a^{\phi(n)} = 1 \bmod n, \tag{90}$$

for relatively prime integers a and n $(a < n)$. Here $\phi(n)$ is Euler's ϕ function [9] which counts the number of positive integers smaller than n and coprime to n. Clearly, for any prime integer p, $\phi(p) = p-1$ (any strictly positive integer smaller than p is coprime to p). It can also be shown than for any positive integers s and t that are coprime to each others, $\phi(st) = \phi(s)\phi(t)$. In our case we obtain $\phi(n) = (p-1)(q-1) = \phi$ (p and q are distinct primes). Thus the cryptogram $C = P^e \bmod n$ can indeed be decrypted by $C^d \bmod n = P^{ed} \bmod n$ because $ed = 1 \bmod \phi(n)$, implying the existence of an integer k such that $ed = k\phi(n)+1$, and

$$P^{ed} \bmod n = P^{k\phi(n)+1} \bmod n = P. \tag{91}$$

For example, let us suppose that Bob's public key is $\pi = (e, n) = (179, 571247)$.[6] He generated it following the prescription above choosing $p = 773$, $q = 739$ and $e = 179$. The private key d was obtained by solving $179d = 1 \bmod 772 \times 738$ using the extended Euclidean algorithm which yields $d = 515627$. Now if we want to send Bob encrypted "SHAKEN NOT STIRRED" we first use our digital alphabet to obtain the plaintext which can be written as the following sequence of six digit numbers

$$180700 \quad 100413 \quad 261314 \quad 192618 \quad 190817 \quad 170403$$

Then we encipher each block P_i by computing $C_i = P_i^e \bmod n$; e.g. the first block $P_1 = 180700$ will be eciphered as

$$P_1^e \bmod n = 180700^{179} \bmod 571247 = 141072 = C_1, \tag{92}$$

and the whole message is enciphered as:

$$141072 \quad 253510 \quad 459477 \quad 266170 \quad 286377 \quad 087175$$

The cryptogram C composed of blocks C_i can be send over to Bob. He can then decrypt each block using his private key $d = 515627$, e.g. the first block is decrypted as

$$141072^{515627} \bmod 571247 = 180700 = P_1. \tag{93}$$

In order to recover plaintext P from cryptogram C, an outsider, who knows C, n, and e, would have to solve the congruence

$$P^e \bmod n = C, \tag{94}$$

for example, in our case,

$$P_1^{179} \bmod 571247 = 141072. \tag{95}$$

Solving such an equation is believed to be a hard computational task for classical computers. So far, no classical algorithm has been found that computes the solution efficiently when n is a large integer (say 200 decimal digits long or more). However, if we know the prime decomposition of n it is a piece of cake to figure out the private key d: we simply follow the key generation procedure and solve the congruence $ed = 1 \bmod (p-1)(q-1)$. This can be done efficiently even when p and q are very large. Thus, in principle, anybody who knows n can find d by factoring n. The security of RSA therefore relies among others on the assumption that factoring large numbers is computationally difficult. In the context of classical computation, such difficulty has never been proved. Worse still, we have seen in Section 8 that there is a quantum algorithm that factors large number efficiently. This means that the security of the RSA cryptosystem will be completely compromised if large-scale quantum computation becomes one

[6] Needless to say, number n in this example is too small to guarantee security; do not try this public key with Bob.

day practical. This way, the advent of quantum computation rules out public cryptographic schemes commonly used today that are based on the “difficulty” of factoring or the “difficulty” of another mathematical operation called discrete logarithm [9].

On the other hand, quantum computation provides novel techniques to generate a shared private key with perfect confidentiality, regardless the computational power (classical or quantum) of the adversaries. Such techniques are referred to as *quantum key distribution* protocols. Discussion on quantum key distribution is outside the scope of this lecture. Interested readers are referred to [35,36,37].

Acknowledgments

This work was supported in part by the European TMR Research Network ERP-4061PL95-1412, The Royal Society London, and Elsag plc.

References

1. The term was coined by B. Schumacher. See, for example, *Phys. Rev. A* **51** 2738 (1995).
2. D. Deutsch, *Proc. R. Soc. Lond. A* **425** 73 (1989).
3. W. K. Wootters and W. H. Zurek, *Nature* **299** 802 (1982).
4. A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. W. Shor, T. Sleator, J. Smolin and H.Weinfurter, *Phys. Rev. A* **52** 3457 (1995).
5. D. Deutsch, A. Barenco and A. Ekert, *Proc. R. Soc. Lond. A* **449** 669 (1995).
6. A. Barenco, D. Deutsch, A. Ekert and R. Jozsa, *Phys. Rev. Lett.* **74** 4083 (1995).
7. D. P. DiVincenzo, *Phys. Rev. A* **51** 1015 (1995).
8. S. Lloyd, *Phys. Rev. Lett.* **75** 346 (1995).
9. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford University Press, Oxford, 1979).
10. T. Toffoli, *Mathematical Systems Theory* **14** 13 (1981).
11. C. H. Papadimitriou, *Computational Complexity* (Addison-Wesley, 1994).
12. D. Coppersmith, *IBM Research report* (1994).
13. R. Solovay and V. Strassen, *SIAM J. Comp.* **6** 84 (1977).
14. R. Motwani and P. Raghavan, *Randomised Algorithms* (Cambridge University Press, 1995).
15. P. W. Shor, “Algorithms for quantum computation: Discrete logarithms and factoring” *Proc. 35th Annual Symposium on the Foundations of Computer Science*, p. 124 edited by S. Goldwasser (IEEE Computer Society Press, Los Alamitos, CA 1994). Expanded version of this paper is available at LANL quant-ph archive.
16. R. P. Feynman, *Int. J. of Theor. Phys.* **21** 467 (1982).
17. R. Cleve, A. Ekert, C. Macchiavello and M. Mosca, *Proc. R. Soc. Lond. A* **454** 339 (1998).
18. D. Deutsch, *Proc. R. Soc. Lond. A* **400** 97 (1985).
19. L. K. Grover, “A fast quantum mechanical algorithm for database search”, *Proc. 28th Annual ACM Symposium on the Theory of Computing (STOC’96)*, p. 212 (ACM, Philadelphia, Pennsylvania, 1996).

20. M. Boyer, G. Brassard, P. Hoyer, A. Tapp, *Proc. of the Workshop on Physics and Computation (PhysComp96)* 36 (1996).
21. T. Hogg, *Physica* **D120** 102 (1998).
22. C. Zalka, *Physical Review* **A60** 2746 (1999).
23. A. Y. Kitaev, LANL quant-ph archive, quant-ph/9511026 (1995).
24. V. Vedral, A. Barenco and A. Ekert, *Phys. Rev. A* **54** 147 (1996).
25. D. Beckman, A. Chari, S. Devabhaktuni and J. Preskill, *Phys. Rev. A* **54** 1034 (1996).
26. A. Ekert and R. Jozsa, *Rev. Mod. Phys.* **68** 733 (1996).
27. D. Kahn, *The Codebreakers: The Story of Secret Writing*, (Macmillan, New York,1967).
28. D. Stinson, *Cryptography: Theory and Practice* (CRC Press, 1995).
29. G. S. Vernam, *J. AIEE* **45** 109 (1926).
30. C. E. Shannon, *Bell Syst. Tech. J.* **28** 657 (1949).
31. W. Diffie and M. E. Hellman, *IEEE Transactions on Information Theory* **22** 644 (1976).
32. R. L. Rivest, A. Shamir and L. M. Adleman, *Communication of the ACM* **21** 120 (1978).
33. J. H. Ellis, *Tech. report* Communications-Electronics Security Group, United Kingdom (1970).
34. C. Cocks, *Tech. report* Communications-Electronics Security Group, United Kingdom (1973).
35. C. H. Bennett and G. Brassard, *Proc. IEEE Int. Conference on Computers, Systems and Signal Processing* (IEEE, New York, 1984).
36. C. H. Bennett, *Phys. Rev. Lett.* **68** 3121 (1992).
37. A. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).

Environment-Induced Decoherence and the Transition from Quantum to Classical

Juan Pablo Paz and Wojciech Hubert Zurek

Abstract. We study dynamics of quantum open systems, paying special attention to these aspects of their evolution which are relevant to the transition from quantum to classical. We begin with a discussion of the conditional dynamics of simple systems. The resulting models are straightforward but suffice to illustrate basic physical ideas behind quantum measurements and decoherence. To discuss decoherence and *environment-induced superselection* (*einselection*) in a more general setting, we sketch perturbative as well as exact derivations of several master equations valid for various systems. Using these equations we study einselection employing the general strategy of the predictability sieve. Assumptions that are usually made in the discussion of decoherence are critically reexamined along with the "standard lore" to which they lead. Restoration of quantum-classical correspondence in systems that are classically chaotic is discussed. The dynamical second law – it is shown – can be traced to the same phenomena that allow for the restoration of the correspondence principle in decohering chaotic systems (where it is otherwise lost on a very short time-scale). Quantum error correction is discussed as an example of an anti-decoherence strategy. Implications of decoherence and einselection for the interpretation of quantum theory are briefly pointed out.

1 Introduction and Overview

The quantum origin of the classical world was so difficult to imagine for the forefathers of quantum theory that they were often willing to either postulate its independent existence (Bohr), or even to give up quantum theory and look for something with more fundamental classical underpinnings (de Broglie and, to a lesser extent, also Einstein). The source of the problem is the quantum principle of superposition, which, in effect, exponentially expands the set of available states to all of the conceivable superpositions. Thus, coherent superpositions of dead and live cats have – in the light of the quantum theory – the same right to exist as either of the two classical alternatives. Within the Hilbert space describing a given system "classically legal" states are exceptional. The set of all states in the Hilbert space is enormous as compared with the size of the set of states where one finds classical systems. Yet, it is a fact of life that classical objects are only found in a very small subset of all possible (and in principle, allowed) states. So, one has to explain the origin of this apparent "superselection" rule that prevents the existence of most states in the Hilbert space of some physical systems. Decoherence and its principal consequence – environment-induced super-selection or *einselection* – account for this experimental fact of life.

W. Dieter Heiss (Ed.): LNP 587, pp. 77–148, 2002.
© Springer-Verlag Berlin Heidelberg 2002

Decoherence is caused by the interaction between the system and its environment. Under a variety of conditions, which are particularly easy to satisfy for macroscopic objects, it leads to the einselection of a small subset of quasi-classical states from within the enormous Hilbert space. The classicality is then an emergent property, induced in the system by its interaction with the environment. Arbitrary superpositions are dismissed, and a preferred set of "pointer states" emerges. These preferred states are the candidate classical states. They correspond to the definite readings of the apparatus pointer in quantum measurements, as well as to the points in the phase space of a classical dynamical system.

The role of the process of decoherence in inducing classicality has become clear only relatively recently – within the past two decades. The key idea is relatively simple: An environment of a quantum system can, in effect, monitor its states through continuous interaction. The imprint of the system left on the environment will contain information about selected states of the system. The states that leave the imprint without getting perturbed in the process are the preferred states. Thus, the key property of quasi-classical pointer states is their insensitivity to monitoring by – and consequently their resistance to the entanglement caused by – interaction with the environment: states that entangle least are most stable. They are also, almost by definition, the only states that remain an accurate description of the the system alone: All other states evolve into joint system-environment states, preserving their purity (and, consequently, the information the observer has about them) only when both the system and the environment are included in a larger "supersystem".

The fact that the interaction between quantum systems produces entanglement was well known almost since the beginning of quantum theory. Indeed, because the ideas of decoherence and einselection rely on quantum theory, and on quantum theory alone, it may be useful to ask why it took so long to arrive at a natural explanation of the quantum origins of classicality. There are several possible explanations for this delay. We shall return to them later in the paper. But, for the moment, it is useful to note that the ability of environment-induced decoherence to result in the same set of preferred states, essentially independently of the initial state of the system and the environment, is crucial. This was not appreciated until relatively recently [1,2]. It is precisely this stability of the set of preferred states that allows them to be regarded as good candidates for the quantum counterparts of classical reality. Indeed, only still more recent research on the predictability sieve has allowed for more fundamental and general understanding of the emergent classicality (see [3,4] and also [5]).

The prejudice that seems to have delayed serious study of the role of the "openness" of a quantum system in the emergence of classicality is itself rooted in the classical way of thinking about the Universe. Within the context of classical physics, all fundamental questions were always settled in the context of closed systems. The standard strategy to ensure isolation involved enlarging a system – i.e., by including the immediate environment. The expectation was that in this manner one can always reduce any open system to a larger closed system. This strategy does indeed work in classical physics, where the enlargement can help in satisfying conservation laws for quantities such as energy or momentum. It fails in the quantum case under discussion, because now it is the information (about the state of the system) that must be prevented from spreading. Information is much harder to contain when the system in question becomes larger. Thus, in the end, the only truly isolated macroscopic system is the Universe as a whole. And we, the observers, are certainly not in a position to study it from the outside.

In what sense is the preferred set of states preferred? It is clear that generic superpositions of the members of this preferred set will decay into mixtures. On the other hand, if the initial state is just one of the members of the preferred set, the temporal evolution will minimally affect the state, which will resist becoming entangled with the environment. Einselection can thus be thought of as a process by which a “record” of the state of the system is created dynamically (through interaction) in the state of the environment. It is this ongoing process by which the system is being continuously monitored by the environment that leads to the emergence of a natural set of preferred states that are the least affected by the interaction.

As sketched above, the physical principles of decoherence and einselection appear, in retrospect, rather straightforward. How much can be accomplished by exploring their consequences? There are several interesting and important questions that naturally arise in this context and that have been asked (and answered, in most cases) over the last two decades. First, one naturally asks how much can we explain with these ideas (i.e., is it consistent to think that all objects that are known to behave classically are doing so because of decoherence?). A closely related question is the one concerning natural time-scales associated with decoherence. How fast does decoherence take place? This is a very important question because a first look at the decoherence process may leave us wondering if decoherence may be consistent with the existence of a “reversible” classical world. Thus, if one believes that classicality is really an emergent property of quantum open systems one may be tempted to conclude that the existence of emergent classicality will always be accompanied by other manifestations of openness such as dissipation of energy into the environment (this would be a problem because, as we know, there are many systems that behave classically while conserving energy). Second, one also wonders how, in detail, is the preferred set of states dynamically selected through the interaction with the environment. In particular, it is interesting to know how this pointer basis is determined by the structure of the interaction Hamiltonian between system and environment and/or to the other details of the physics involved. Third, a related question arises in this context: are there observable manifestations of decoherence other than einselection?

A remarkable characteristic of the current debates on the nature of the quantum to classical transition and on the problem of quantum measurement is that for the first time in history there have been actual experiments probing the boundary between the quantum and the classical domains in a controlled way [6,8,7,9,10]. Controlled decoherence experiments (which are very difficult because nature provides us with classical or quantum systems but not with objects whose interaction with the environment can be controlled at will) were recently carried on for the first time and help us in understanding the nature of this process. Some of the most notable experiments in this area were performed at the Ecole Normale Superieure in Paris.

Our lectures start with an introduction to quantum conditional dynamics using two-state systems. Conditional dynamics is responsible both for setting up the problem of measurement, and for the decoherence and einselection that solve it. The resulting models are straightforward and can serve in the idealized studies of the measurement process. However, they are clearly too simple to be realistic – classicality is, after all, a property of essentially every sufficiently macroscopic object. To discuss decoherence and einselection in this more general setting, we shall therefore study dynamics of quantum open systems. Section 3 is devoted to the derivation of the key tool – a master equation for the reduced density matrix. This basic tool is immediately used in Sect. 4, where environment-induced superselection is studied, including, in particular, the predictability sieve. Section 5 analyzes some “loose ends” – that is, essentially technical issues that are usually omitted in the derivations of the master equations. We show there that although the qualitative conclusions arrived at on the basis of the “naive” master equation approach are essentially unaffected by the detailed examination of some of the idealized assumptions that go into its derivation, quantitative estimates can change quite significantly when a more realistic approach is adopted. Section 6 is devoted to the study of the effect of decoherence on the quantum-classical correspondence in systems that are classically chaotic. We show there that decoherence not only explains the origin of classical dynamics, but that it may be responsible for the loss of information that accounts for the second law of thermodynamics as well. Section 7 is devoted to quantum error correction – to the strategies which can be used to suppress decoherence. The summary and conclusions are briefly stated in Sect. 8.

2 Quantum Measurements

In this section we shall introduce the measurement problem – the issue that has dominated the discussion of the relation between quantum and classical for a very long time. This will afford us the opportunity to study conditional dynamics that will be employed in one form or another throughout this review. Such interactions are necessary to achieve entanglement between quantum systems that set up the measurement problem. They are necessary for accomplishing decoherence, which leads to environment-induced superselection (or einselection), and thus resolves many of the problems arising on the border between quantum and classical. Last, but not least, quantum conditional dynamics and entanglement underlie quantum logic and will be of importance in the latter part of the paper devoted to quantum error correction.

Predictability is rightly regarded as one of the key attributes of classical dynamics. On the other hand, the defining feature of quantum mechanics is thought to be its probabilistic nature, which manifests itself in measurements. This discord between classical determinism and quantum randomness is often blamed for the difficulties with interpretation of quantum theory. Yet, the fundamental equations of either classical or quantum theory allow them – indeed, demand of them – to be perfectly predictable: It is just that what can be predicted with certainty, especially in the quantum case, cannot be often accessed by measurements. And, conversely, what can be measured in an evolving quantum system cannot usually be predicted, except in the probabilistic sense.

The Schrödinger equation allows one to predict the state of an isolated system at any subsequent moment of time. In an isolated quantum system, dynamical evolution is strictly deterministic. This perfect quantum predictability could be of use only if one were to measure observables that have the resulting evolving state as one of its eigenstates. These observables are generally inaccessible to reasonable measuring devices, and therefore are of no interest.

Quantum determinism is of little use for an observer who is only a part of the whole system. The overall quantum determinism could have predictive power only for someone who is (i) monitoring quantum systems from the outside. Moreover, it would help if the observer was endowed with (ii) enough memory to measure and store data, and (iii) sufficient ability to compute and to model deterministic evolution of the system of interest. For an observer trapped inside of the quantum universe, this is obviously not the case.

The universe is all there is. Therefore, by definition, it is a closed quantum system. Given the deterministic nature of the Schrödinger evolution, one may be surprised that there is a problem with the interpretation of quantum theory. After all, the interpretational ideal often mentioned in such discussions is deterministic Newtonian dynamics. However, the interpretation problem stems from the fact that deterministic unitary evolution of quantum theory is incompatible with classical determinism. Indeed, as the studies of chaotic systems demonstrate, classical dynamics has more room for randomness than quantum physics.

States of the quantum systems are perturbed by the very act of monitoring them. The elemental unpredictability associated with the act of observation cannot be avoided unless the observer knows in advance which observables can be measured with impunity. This feature of quantum information is essential to guarantee the security of quantum cryptography (see lectures of A. Ekert; also [11]) – the state of a quantum system cannot be found out by the eavesdropper unless the observation is carried out on the same basis as the one used by the intended recipient of the message. The “no cloning theorem” [12,13] prevents duplication of quantum information – amplification is associated with breaking the symmetry associated with the superposition principle.

Environment-induced superselection rules allow the observer to be a successfull eavesdropper, and to extract useful information from the quantum systems without the environment getting in the way because (in contrast to the strategies employed in quantum cryptography) the measurements carried out by the environment are restricted to few observables. The state of the system is therefore of necessity "precollapsed" and commutes with these observables. Further measurements carried out by the observer will only reveal (rather than perturb) the pre-existing state of affairs. Thus, environment-induced decoherence supplies a justification for the persistent impression of "reality". In contrast to the observables encountered in the microscopic realm, macroscopic quantum systems can appear only in one of the preselected (pointer basis) set of quantum states. The "collapse of the wave packet" viewed in this way is just a familiar classical process of finding out which of the possible outcomes has actually occurred. The danger of interference between the alternatives was eradicated by decoherence long before the observer became involved.

How can one ever hope for a resolution that would allow for the familiar combination of classical determinism and classical randomness to emerge? At the risk of anticipating results that will be justified in detail only later, we note that quantum determinism may be relevant only for an observer who knows the initial state of an isolated quantum system. For a quantum observer immersed in a quantum universe this is a very rare exception, attainable only in carefully controlled laboratory experiments, and only for rather small quantum systems. The information capacity, memory, and information processing abilities of an observer that is a (macroscopic, yet comparatively small) subsystem of the universe are miniscule compared to the task of simulating even a small quantum system, let alone the universe as a whole. And as soon as the idea of the observer knowing the entire state of the universe is recognized as not feasible, "environmental monitoring" of both the state of the observer and of the observables he recorded begins to matter. An observer with decohering memory can keep reliable records only in the einselected states of his/her memory bits [14,3,16,17,15]. Records will have predictive power only when they correlate with the einselected observables in the rest of the universe.

2.1 Bit-by-Bit Measurement and Quantum Entanglement

This problem of transition from quantum determinism to classical definiteness is illustrated most vividly by the analysis of quantum measurements. An answer to a "generic" question about the state of a quantum system (and the outcome of a measurement of the corresponding observable) is *not* deterministic. In the usual textbook discussions, this random element is blamed on the "collapse of the wave packet" that is invoked whenever a quantum system comes into contact with a classical apparatus. In a fully quantum discussion of the problem, this issue still arises, in spite (or rather because) of the overall deterministic quantum evolution of the state vector of the universe. Indeed, as carefully pointed out by von Neumann [18] in his quantum analysis of measurements, there seems to be no room for "real collapse" in purely unitary models of measurements.

To illustrate the ensuing difficulties, we consider with von Neumann a quantum system $\mathcal{S}$ initially in a state $|\psi\rangle$ interacting with a quantum apparatus $\mathcal{A}$ initially in a state $|A_0\rangle$. The interaction will generally result in an entangled final state,

$$|\Psi_0\rangle = |\psi\rangle|A_0\rangle = (\sum_i a_i|s_i\rangle)|A_0\rangle \longrightarrow \sum_i a_i|s_i\rangle|A_i\rangle = |\Psi_t\rangle. \tag{1}$$

Here $\{|A_i\rangle\}$ and $\{|s_i\rangle\}$ are states in the Hilbert spaces of the apparatus and of the system, respectively, and a_i are complex coefficients. This transition can be accomplished by means of a unitary Schrödinger evolution. It leads to an uncomfortable conclusion. All that an appropriate interaction between $\mathcal{A}$ and $\mathcal{S}$ can achieve is putting the measuring apparatus (or an observer) in an EPR-like *entangled state* of all the possible outcomes consistent with the initial state [1]. Operationally, this EPR-like nature of the state emerging from the premeasurement (as the step achieved by (1) is often called) can be made more explicit by rewriting the sum in a different basis

$$|\Psi_t\rangle = \sum_i a_i|A_i\rangle|s_i\rangle = \sum_i b_i|B_i\rangle|r_i\rangle = |\Psi_t\rangle\ . \tag{2}$$

All we have done is use an alternative basis for both the apparatus and the system, exploiting the freedom of choice guaranteed by the quantum principle of superposition. Therefore, if one were to associate the state of the apparatus (observer) with a state in the decomposition of $|\Psi_t\rangle$, then even before one could start enquiring about the specific outcome of the measurement one would have to decide what decomposition of $|\Psi_t\rangle$ is to be used, because the change of the basis corresponds to a redefinition of the measured quantity.

One could make the clash between quantum and classical even more dramatic by making an additional measurement on the same quantum system after the premeasurement correlation is established. In accord with (2), such an additional measurement would have a power to select an arbitrary observable of the system $\mathcal{S}$ and would single out the corresponding states of the apparatus $\mathcal{A}$. Yet, given the freedom to rewrite $|\Psi_t\rangle$ in an infinite number of ways, this state of $\mathcal{A}$ would be for almost any choice of the decomposition of the sum of (2) completely “nonclassical” in any reasonable sense, and it would depend on the initial state of the quantum system.

In a quantum domain, such an entanglement must be commonplace, along with its disturbing consequences. Indeed, a “Schrödinger kitten” state recently implemented by means of an atomic physics experiment ([19] is an excellent illustration of the distinction between the quantum entanglement and the classical correlation in the context of quantum measurements). The NIST group in Boulder has managed – manipulating a single ion inside a trap with lasers – to establish a correlation between its internal state (designated here by $\{|\uparrow\rangle, |\downarrow\rangle\}$, respectively, for “excited” and “ground”) and its location ($|L\rangle$ or $|R\rangle$ for “left” or “right”). The final correlated wavefunction has a premeasurement, EPR-like form,

$$|\Psi_A\rangle \ = \ (|+\rangle|L\rangle \ + \ |-\rangle|R\rangle)/\sqrt{2}\ , \tag{3}$$

where

$$|\pm\rangle \ = \ (|\uparrow\rangle \ \pm \ |\downarrow\rangle)/\sqrt{2}\ \ , \tag{4}$$

are superpositions of the ground and excited states. This very same $|\Psi_A\rangle$ can be written therefore as

$$|\Psi_A\rangle = \{|\uparrow\rangle(|L\rangle + |R\rangle) + |\downarrow\rangle(|L\rangle - |R\rangle)\}/\sqrt{2}\ . \tag{5}$$

Thus, the same correlated state of the "atom cat" can be expressed in two very different – looking ways, implying the potential for still more kinds of ambiguous correlations. Expressed in the first way (see (3)), the atom can be in one of the two alternative locations, depending on its internal state that is defined as a superposition of ground and excited states. In the second way (given by (5)) the natural internal states of the atom are correlated with a very nonclassical state – a superposition of an atom in two locations. Monroe et al. [19] measure the internal state of the atom in the basis corresponding to the decomposition of (5) and verify that it is indeed in a superposition of $|L\rangle$ and $|R\rangle$ with either a positive or a negative sign (an "even" or and "odd" Schrödinger cat).

Given the atomic size of this "kitten", its ability to appear in a superposition of two different widely separated locations may or may not be a surprise. But the point this recent experiment allows us to make is at the heart of the interpretation problem. If the quantum laws are universally valid, very nonclassical Schrödinger cat–like states should be commonplace for an apparatus that measures a quantum system and, indeed, for run-of-the-mill macroscopic systems in general. One should be able to prepare such nonclassical states at will, by entangling arbitrarily large objects with quantum states of microscopic systems and then measuring these quantum objects in some arbitrary basis. If such sequences of events were common, classical objects would almost always be in very nonlocal superposition states.

Quantum theory mandates this pandemonium. Yet, we never seem to encounter it, least of all in the course of measurements. The task of the interpretation of quantum theory is to understand why. In the Copenhagen interpretation, this problem never arises, because the apparatus is by definition classical. However, if one insists on the universality of quantum theory, the difficulty described above is inevitable. It arises, for instance, in Everett's Many Worlds Interpretation, which was in fact originally called "the Relative State Interpretation" [20]. Everett and other followers of the MWI philosophy tried to occasionally bypass this question by insisting that one should only discuss correlations. Correlations are indeed at the heart of the problem, but it is not enough to explain how to compute them; for that, quantum formalism is straightforward enough. What is needed instead is an explanation of why some states retain correlations, but most of them do not, in spite of the arbitrariness in basis selection that is implied by (2). Or, equivalently, what is needed is an explanation of the loss of general quantum entanglement, but a selective retention of classical correlations – correlations that are also quantum in their origin, but which consistently single out the same basis of the quantum states violating the spirit of the superposition principle.

2.2 Interactions and the Information Transfer in Quantum Measurements

The interaction required to accomplish the correlation between the measured system and the apparatus, (1), can be regarded as a generalization of the basic logical operation known as a "controlled not" or a `c-not`. Classically, `c-not` changes the state of the target bit when the control bit is in a state 1, and does nothing otherwise:

$$\begin{aligned} 0_{\mathrm{c}} \begin{matrix} 0_{\mathrm{t}} \\ 1_{\mathrm{t}} \end{matrix} &\longrightarrow 0_{\mathrm{c}} \begin{matrix} 0_{\mathrm{t}} \\ 1_{\mathrm{t}} \end{matrix} \\ 1_{\mathrm{c}} \begin{matrix} 0_{\mathrm{t}} \\ 1_{\mathrm{t}} \end{matrix} &\longrightarrow 1_{\mathrm{c}} \begin{matrix} 1_{\mathrm{t}} \\ 0_{\mathrm{t}}. \end{matrix} \end{aligned} \tag{6}$$

Quantum `c-not` is a straightforward quantum version of (6). It differs from the classical case only because arbitrary superpositions of the control bit and of the target bit are allowed

$$(\alpha|0_{\mathrm{c}}\rangle \ + \ \beta|1_{\mathrm{c}}\rangle)|a_{\mathrm{t}}\rangle \longrightarrow \ \alpha|0_{\mathrm{c}}\rangle|a_{\mathrm{t}}\rangle \ + \ \beta|1_{\mathrm{c}}\rangle|\neg a_{\mathrm{t}}\rangle.$$

Above a "negation" of a state $|\neg a_{\mathrm{t}}\rangle$ is a basis–dependent operation defined by

$$\neg(\gamma|0_{\mathrm{t}}\rangle + \delta|1_{\mathrm{t}}\rangle) = \gamma|1_{\mathrm{t}}\rangle + \delta|0_{\mathrm{t}}\rangle.$$

It suffices to identify $|A_0\rangle = |0_{\mathrm{t}}\rangle$, and $|A_1\rangle = |1_{\mathrm{t}}\rangle$ to have an obvious correspondence between the `c-not` and a premeasurement.

In the classical `c-not` the direction of the information transfer is always consistent with the designations of the two participating bits. The state of the control bit remains unchanged while it controls the state of the target bit, (6). Written in terms of the logical $\{|0\rangle, |1\rangle\}$ basis, the truth table of the quantum `c-not` is essentially – that is, save for the possibility of superpositions – the same as (6). One might therefore anticipate that the direction of information transfer and the designations ("control/system" and "target/apparatus") of the two qubits will also be unambiguous, as they are in the classical case. This expectation however is incorrect, as can be readily demonstrated by expressing the process in the conjugate basis $\{|+\rangle, |-\rangle\}$ that, for either control or target bit, is obtained through the Hadamard transform:

$$|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}. \tag{9}$$

The truth table of (6) in conjunction with the principle of superposition (which allows one to write down (9)) leads to a new complementary truth table

$$\begin{aligned} |\pm\rangle|+\rangle &\longrightarrow |\pm\rangle|+\rangle \\ |\pm\rangle|-\rangle &\longrightarrow |\mp\rangle|-\rangle. \end{aligned} \tag{10}$$

That is, in the complementary basis $\{|+\rangle, |-\rangle\}$ the roles of the control and the target bit are reversed. The state of the former target – represented by the second ket in (10) – remains unaffected in the new basis, and the state of the former control is conditionally "flipped".

In the above c-not (or bit-by-bit measurement), the appropriate interaction Hamiltonian is

$$\begin{aligned} H_{\text{int}} = g|1\rangle\langle 1|_{\mathcal{S}}|-\rangle\langle -|_{\mathcal{A}} &= \frac{g}{2}|1\rangle\langle 1|_{\mathcal{S}}(\mathbf{1} - (|0\rangle\langle 1| + |1\rangle\langle 0|))_{\mathcal{A}} \\ &= g(\frac{1-\sigma_z}{2})_{\mathcal{A}}(\frac{1-\sigma_x}{2})_{\mathcal{S}}. \end{aligned} \tag{11}$$

Above, g is a coupling constant, σ_i are Pauli matrices, and the two operators refer to the system (i.e., to the former control), and to the apparatus pointer (the former target), respectively. It is easy to see that the states $\{|0\rangle, |1\rangle\}_{\mathcal{S}}$ of the system are unaffected by H_{int}, because

$$[H_{\text{int}},\ e_0|0\rangle\langle 0| + e_1|1\rangle\langle 1|] = 0. \tag{12}$$

Thus, the measured (control) observable $\hat{\epsilon} = e_0|0\rangle\langle 0| + e_1|1\rangle\langle 1|$ is a constant of motion under the evolution generated by H_{int}.

The states $\{|+\rangle, |-\rangle\}_{\mathcal{A}}$ of the apparatus (which encode the information about the phase between the logical states) have exactly the same "immunity"

$$[H_{\text{int}},\ f_+|+\rangle\langle +| + f_-|-\rangle\langle -|] = 0. \tag{13}$$

Hence, when the apparatus is prepared in a definite phase state (rather than in a definite pointer/logical state), it will pass its phase onto the system, as the truth table, (10), shows. Indeed, H_{int} can be rewritten in the Hadamard transformed basis

$$\begin{aligned} H_{\text{int}} &= g|1\rangle\langle 1|_{\mathcal{S}}|-\rangle\langle -|_{\mathcal{A}} \\ &= \frac{g}{2}(\mathbf{1} - (|-\rangle\langle +| + |+\rangle\langle -|))_{\mathcal{S}}|-\rangle\langle -|_{\mathcal{A}}, \end{aligned} \tag{14}$$

which, in comparison with (11), makes this "immunity" obvious.

This basis-dependent direction of the information flow in a quantum c-not (or in a premeasurement) is a direct consequence of complementarity. It can be summed up by stating that although the information about the observable with the eigenstates $\{|0\rangle, |1\rangle\}$ travels from the measured system to the apparatus, in the complementary $\{|+\rangle, |-\rangle\}$ basis it seems to be the apparatus that is being measured by the system. This observation also clarifies the sense in which phases are inevitably "disturbed" in measurements. They are not really destroyed, but, rather, as the apparatus measures a certain observable of the system, the system "measures" the phases between the possible outcome states of the apparatus. These phases in a macroscopic apparatus coupled to the environment are fluctuating rapidly and uncontrollably, thus leading to the destruction of phase coherence. However, even if this consequence of decoherence were somehow prevented (i.e., by perfectly isolating the apparatus pointer from the environment), preexisting phases between the outcome states of the apparatus would have to be known while, simultaneously, $\mathcal{A}$ is in $|A_0\rangle$, the "ready–to–measure–state". This would require a simultaneous knowledge of the two non-commuting observables, and is therefore impossible because of Heisenberg indeterminacy.

It appears that even the question “which of the two interacting systems is a measuring device?” (which should be decided by the direction of the flow of information) depends on the initial states. In “classical practice” this ambiguity does not arise because the initial state of the apparatus can never be selected at the whim of the observer. Einselection limits the set of possible states of the apparatus to a small subset of all the states available in Hilbert space.

2.3 Monitoring by the Environment and Decoherence

In this section, we shall see how the quantum-classical correspondence can be reestablished by decoherence and einselection, caused by the monitoring of the to-be-classical observables by the environment. The environment is defined as any set of degrees of freedom that are coupled to the system of interest, and which can therefore ‘monitor’ – become entangled with – its states. Environments can be external (such as particles of air or photons that scatter off, say, the apparatus pointer) or internal (e.g., collections of phonons or other excitations in the materials from which an apparatus is constructed). Often, environmental degrees of freedom emerge from the split of the original set of degrees of freedom into the “system of interest” that is some collective observable (order parameter in a phase transition), and the “microscopic remainder”.

The superposition principle applies only when the quantum system is closed. When the system is open, interaction with the environment will inevitably result in an incessant “monitoring” of some of the observables by the environmental degrees of freedom. This will result in the degradation of the pure states into mixtures. These mixtures will often – remarkably often – turn out to be diagonal in the same set of “preferred states” that are nearly independent of the initial state of the system and of the environment, but which are selected with the crucial help of the interaction Hamiltonian. This decoherence process determines the relative “fitness” of all the possible superpositions that exist in the Hilbert space. The resulting “natural selection” is responsible for the emergence of classical reality. Its consequence is known as environment-induced superselection [2], or einselection.

The set of habitually decohering states is often called “the pointer basis”, in recognition of its role in the measurement problem. The criterion for the selection of pointer states goes well beyond the often–repeated characterizations based solely on the instantaneous eigenstates of the density matrix. What is of the essence is the ability of the einselected states to survive monitoring by the external degrees of freedom. This heuristic criterion can be made rigorous by quantifying the predictability of the evolution of the candidate classical states or of the associated pointer observables. To put it succinctly, measurement of the pointer observables yields an optimal initial condition. In spite of the openness of the system, its results can be employed for the purpose of prediction better than the other Hilbert space alternatives.

The contrast between the resilience of the states associated with the preferred (pointer) observables and the fragility of their superpositions can be analyzed in terms of Heisenberg's principle of indeterminacy. The environment monitors observables with the accuracy dictated by the interaction Hamiltonian. Thus, only a measurement that happens to commute with the observables monitored by the environment will result in a useful record that can be successfully employed for the purpose of prediction. In contrast, a system prepared by the measurement in an arbitrary superposition will also be monitored by the environment, which will tend to correlate with the pointer observable. When the initial superposition prepared by the observer does not commute with the observables monitored by the environment, Heisenberg's indeterminacy implies that the records of the observer are of no use for the purpose of prediction. The monitoring continuously carried out by the environment on the pointer observables makes anything except for the pointer states a poor choice.

Three quantum systems – the measured system $\mathcal{S}$, the apparatus pointer or the memory of the observer $\mathcal{A}$, and the environment $\mathcal{E}$ – and the correlations between them will be the subject of the discussion below. In quantum measurements, $\mathcal{S}$ and $\mathcal{A}$ will be coupled. Their quantum entanglement will be converted into an effectively classical correlation as a result of the interaction between $\mathcal{A}$ and $\mathcal{E}$. In measurements of classical systems, both $\mathcal{S}$ and $\mathcal{A}$ will interact with $\mathcal{E}$ and decohere. In either case, states einselected by the environment will be the focus of attention. In $\mathcal{A}$, they will be the repository of information, serving as pointer states of the apparatus or memory states of the observer. The system $\mathcal{S}$ can also look effectively classical when it is subject to einselection, and when $\mathcal{A}$ keeps records of its einselected states.

This $\mathcal{SAE}$ triangle (or a triangle much like it) is necessary for careful study of decoherence and its consequences. By keeping all three corners of this triangle in mind, one can avoid the confusion about the relation of the instantaneous eigenstates of the density matrix (see, for example, the discussion following [21]). This three-system context is necessary to keep track of the correlations between the memory of the observer and the state of the measured system. The evolution from a quantum entanglement to the classical correlation may be the easiest relevant theme to define operationally. In spite of this focus on the correlation, I shall often suppress one of the corners of the above triangle to simplify the equations. All three parts of the triangle will however play a role in the formulation of the questions we shall pose and in motivating the criteria for classicality that we shall devise.

2.4 One-Bit Environment for a Bit-by-Bit Measurement

The simplest discussion of a single act of decoherence involves just three one-bit systems [1,22]. They are denoted by $\mathcal{S}$, $\mathcal{A}$, and $\mathcal{E}$ in an obvious reference to their designated roles. The measurement starts with the interaction of a measured system with the apparatus,

$$\begin{aligned} |\uparrow\rangle|A_0\rangle &\longrightarrow\ |\uparrow\rangle|A_1\rangle \\ |\downarrow\rangle|A_0\rangle &\longrightarrow\ |\downarrow\rangle|A_0\rangle, \end{aligned} \tag{15}$$

where $\langle A_0|A_1\rangle = 0$. For a general state,

$$(\alpha|\uparrow\rangle\ +\ \beta|\downarrow\rangle)|A_0\rangle\ \longrightarrow \alpha|\uparrow\rangle|A_1\rangle\ +\ \beta|\downarrow\rangle|A_0\rangle = |\Phi\rangle. \tag{16}$$

These formulae are an example of a `c-not`–like premeasurement that has already been discussed. As was noted previously, a correlated state of this form is not enough to claim that a measurement has taken place. The biggest problem with $|\Phi\rangle$ is the basis ambiguity. Equation (16) represents only an entanglement, the same as in Bohm's version of the EPR state [23]. The ambiguity of the basis selection in this simple example can be settled with the help of one additional system, $\mathcal{E}$, which performs a premeasurement on the apparatus. As a result,

$$\begin{aligned} |\Psi(0)\rangle_{\mathcal{SAE}} &= (\alpha|\uparrow\rangle|A_1\rangle + \beta|\downarrow\rangle|A_0\rangle)|\varepsilon_0\rangle \\ &\longrightarrow \alpha|\uparrow\rangle|A_1\rangle|\varepsilon_1\rangle + \beta|\downarrow\rangle|A_0\rangle|\varepsilon_0\rangle = |\Psi\rangle. \end{aligned} \tag{17}$$

It may seem that very little can be accomplished by repeating the step that has led to the $\mathcal{S}$-$\mathcal{A}$ correlation and the associated problems. But this is not the case. A collection of three correlated quantum systems is no longer subject to the basis ambiguity we have pointed out in connection with the EPR-like state $|\Phi\rangle$, (16). This is especially true when the states Of the environment are correlated with the simple products of the states of the apparatus–system combination [1]. In (17) above, this can be guaranteed (irrespective of the value of α and β) providing that:

$$\langle\varepsilon_0|\varepsilon_1\rangle = 0\ . \tag{18}$$

When this condition is satisfied, the description of the $\mathcal{A}$-$\mathcal{S}$ pair can be readily obtained in terms of a reduced density matrix:

$$\begin{aligned} \rho_{\mathcal{AS}} &= \mathrm{Tr}_{\mathcal{E}}|\Psi\rangle\langle\Psi| \\ &= |\alpha|^2|\uparrow\rangle\langle\uparrow||A_1\rangle\langle A_1|\ +\ |\beta|^2|\downarrow\rangle\langle\downarrow||A_0\rangle\langle A_0|\ . \end{aligned} \tag{19}$$

This reduced density matrix contains only terms corresponding to classical correlations.

If the condition of (18) did not hold – that is, if the orthogonal states of the environment were not correlated with the apparatus in the basis in which the original premeasurement was carried out – then the terms on the diagonal of the reduced density matrix $\rho_{\mathcal{AS}}$ would be the sum of products rather than simply products of states of $\mathcal{S}$ and $\mathcal{A}$. An extreme example of that situation is the pre-decoherence density matrix of the pure state:

$$\begin{aligned} |\Phi\rangle\langle\Phi| = |\alpha|^2|\uparrow\rangle\langle\uparrow||A_1\rangle\langle A_1| + \alpha\beta^*|\uparrow\rangle\langle\downarrow||A_1\rangle\langle A_0| \\ + \alpha^*\beta|\downarrow\rangle\langle\uparrow||A_0\rangle\langle A_1| + |\beta|^2|\downarrow\rangle\langle\downarrow||A_0\rangle\langle A_0|\ . \end{aligned} \tag{20}$$

Its eigenstate is simply $|\Phi\rangle$. When expanded, $|\Phi\rangle\langle\Phi|$ contains terms that are off–diagonal when expressed in the natural basis consisting of tensor products of states in the two subspaces. Their disappearance as a result of tracing over the environment signals the disappearance of the basis ambiguity. There is of course a conceptual difference with the classical case. In classical mechanics, it was in principle possible to imagine that the outcome was predetermined. In quantum mechanics this is usually impossible even in principle. However, that distinction can be made only with a more complete knowledge than the one typically available to the observer.

The pointer observable that emerges from this simple case is easy to characterize. The interaction Hamiltonian between the apparatus and the environment, $H_{\mathcal{AE}}$, should have the same structure as for the `c-not`. It should be a function of the pointer observable

$$\hat{A} \;=\; a_1|A_1\rangle\langle A_1| + a_0|A_0\rangle\langle A_0| \tag{21}$$

of the apparatus. Consequently, the states of the environment will bear an imprint of the pointer states $\{|A_1\rangle, |A_0\rangle\}$. As was also noted in the discussion of `c-nots`, $[H_{\mathcal{AE}}, \hat{A}] = 0$ immediately implies that $\hat{A}$ is a control, and its eigenstates will be preserved.

Disappearance of quantum coherence because of a “one–bit” measurement has been verified experimentally in neutron and, more recently, in atomic interferometry [24,25,26]. A single act of quantum measurement we have discussed here should be regarded as an elementary discrete instance of continuous monitoring, which is required to bring about the appearance of classicality.

2.5 Decoherence of a Single (Qu)Bit

Another example of decoherence is afforded by a two-state apparatus $\mathcal{A}$ interacting with an environment of N other spins [2]. We can think of it as just another two–state system, and, in that spirit, we shall identify in this section the two apparatus states as $\{|\Uparrow\rangle, |\Downarrow\rangle\}$. The process of decoherence is definitely not limited to states of the apparatus pointers, so these two generic candidate pointer states can belong to any system.

The simplest, yet already quite illustrative example of this situation occurs when the self-Hamiltonian of the apparatus disappears, $H_{\mathcal{A}} = 0$, and the interaction Hamiltonian has the form:

$$H_{\mathcal{AE}} = (|\Uparrow\rangle\langle\Uparrow| - |\Downarrow\rangle\langle\Downarrow|) \otimes \sum_k g_k (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|)_k . \tag{22}$$

Under the influence of this Hamiltonian, the initial state

$$|\Phi(0)\rangle = (a|\Uparrow\rangle + b|\Downarrow\rangle) \;\prod_{k=1}^{N} (\alpha_k|\uparrow\rangle_k + \beta_k|\downarrow\rangle_k) \tag{23}$$

evolves into

$$|\Phi(t)\rangle = a|\Uparrow\rangle \otimes |\mathcal{E}_{\Uparrow}(t)\rangle + b|\Downarrow\rangle \otimes |\mathcal{E}_{\Downarrow}(t)\rangle \ . \tag{24}$$

Here:

$$|\mathcal{E}_{\Uparrow}(t)\rangle = \prod_{k=1}^{N} (\alpha_k \exp(\mathrm{i} g_k t)| \uparrow\rangle_k + \beta_k \exp(-\mathrm{i} g_k t)| \downarrow\rangle_k) = |\mathcal{E}_{\Downarrow}(-t)\rangle \ . \tag{25}$$

The reduced density matrix is then

$$\rho_{\mathcal{A}} = |a|^2 | \Uparrow\rangle\langle\Uparrow | + ab^* r(t) | \Uparrow\rangle\langle\Downarrow | + a^* b r^*(t) | \Downarrow\rangle\langle\Uparrow | + |b|^2 | \Downarrow\rangle\langle\Downarrow |. \tag{26}$$

The coefficient $r(t)$ determines the relative size of the off-diagonal terms. It is given by

$$r(t) = \langle \mathcal{E}_{\Uparrow} | \mathcal{E}_{\Downarrow} \rangle = \prod_{k=1}^{N} [\cos(2 g_k t) + \mathrm{i}(|\alpha_k|^2 - |\beta_k|^2) \sin(2 g_k t)] \ . \tag{27}$$

Unless the k-th spin of the environment is initially in an eigenstate of the interaction Hamiltonian, its contribution to the product will be less than unity. Consequently, for large environments consisting of many (N) spins and at large times the off-diagonal terms are typically small,

$$|r(t)|^2 \simeq 2^{-N} \prod_{k=1}^{N} [1 + (|\alpha_k|^2 - |\beta_k|^2)^2]. \tag{28}$$

This effect can be illustrated with the help of the Bloch sphere. The density matrix of any two-state system can be represented by a point in the 3-D space. In terms of the coefficients a, b, and $r(t)$ that we have previously used, the coordinates of the point representing $\rho(t)$ are: $z = (|a|^2 - |b|^2)$, $x = \mathrm{Re}(ab^* r)$, and $y = \mathrm{Im}(ab^* r)$. When the state is pure, $x^2 + y^2 + z^2 = 1$ – pure states lie on the surface of the Bloch sphere (Fig. 1). When the state is mixed, the point representing it lies inside that sphere. Any conceivable (unitary or non unitary) quantum evolution of the two–state system can be thought of as a transformation of the surface of the pure states into an ellipsoid contained inside the Bloch sphere. Deformation of the Bloch sphere caused by decoherence is a special case of such general evolutions. The decoherence process does not affect a or b. Hence, evolution caused by decoherence alone occurs in a constant–z plane. Such a "slice" through the Bloch sphere would show the point representing the state at a fraction $|r(t)|$ of its maximum distance. The complex number $r(t)$ can be expressed as the sum of the complex phase factors rotating with the frequencies given by differences $\Delta\omega_j$ between the energy eigenvalues of the interaction Hamiltonian, weighted with the probabilities of finding these energy eigenstates in the initial state,

$$r(t) = \sum_{j=1}^{2^N} p_j \exp(-\mathrm{i}\Delta\omega_j t) \ . \tag{29}$$

The index j now denotes partial energy eigenstates of the environment of the interaction Hamiltonian (tensor products of $\uparrow$ and $\downarrow$ states of the environmental spins). The corresponding eigenvalue differences between the two complete energy eigenstates $|\Uparrow\rangle|j\rangle$ and $|\Downarrow\rangle|j\rangle$ are

$$\Delta\omega_j = \langle\Uparrow|\langle j|H_{\mathcal{AE}}|j\rangle|\Downarrow\rangle\ . \tag{30}$$

There are 2^N distinct states $|j\rangle$, and, barring degeneracies, the same number of different $\Delta\omega_j$'s. The probabilities p_j are given by

$$p_j = |\langle j|\mathcal{E}(t=0)\rangle|^2\ , \tag{31}$$

which, in turn, is easily expressed in terms of the appropriate squares of the products of α_k and β_k.

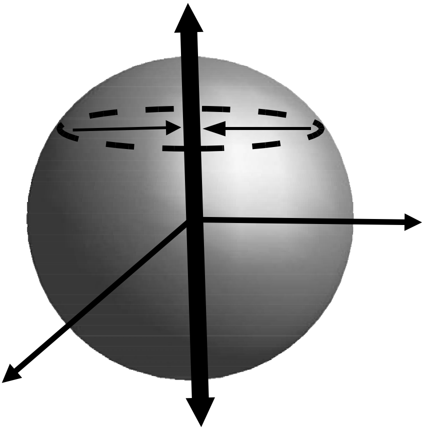


Fig. 1. Decoherence can be seen in the Bloch sphere as the process that induces the states to "move towards the vertical axis", which is defined by the two pointer states on the poles. The classical domain consists of just two pointer states. The classical core is the set of all mixtures of pointer states.

The evolution of $r(t)$ given by (29) is a consequence of the rotations of the complex vectors $p_k \exp(-\mathrm{i}\Delta\omega_j t)$ with different frequencies. The resultant $r(t)$ will then start with the amplitude 1 and quickly "crumble" to a value approximately equal to

$$\langle|r(t)|^2\rangle = \sum_{j=1}^{2^N} p_j^2 \sim 2^{-N}.$$

In this sense, decoherence is exponentially effective – the expected magnitude of the off-diagonal terms decreases exponentially fast with the physical size N of the environment – with the number of systems (spins in our example). In effect, any initial state asymptotically approaches the z axis as a result of decoherence.

We note that the effectiveness of einselection depends on the initial state of the environment: When $\mathcal{E}$ is in the k-th eigenstate of $H_{\mathcal{AE}}$, and $p_j = \delta_{jk}$, the coherence in the system will be retained because the environment is now in an eigenstate of the "control". This situation is, however, unlikely in realistic circumstances because the self-Hamiltonian of the environment $H_{\mathcal{E}}$ will not commute, in general, with $H_{\mathcal{AE}}$. Moreover, even when $H_{\mathcal{E}} = 0$, finding an environment in an energy eigenstate of the Hamiltonian seems extremely unlikely –the eigenvalues of such eigenstates are bound to be dense in large systems, and therefore they will be easily perturbed by the interaction with their environments. Furthermore, the 2^N partial eigenstates of the interaction Hamiltonian are exponentially rare among arbitrary superpositions.

The geometry of the flows induced by decoherence inside the Bloch sphere exhibits characteristics that are encountered in more general physical situations, involving decoherence in bigger Hilbert spaces as follows:

(i) Domain of pure quasi–classical states consisting of all the einselected pointer states ($\{| \Uparrow\rangle, | \Downarrow\rangle\}$ in our case). Pointer states are the pure states least affected (here, unaffected) by decoherence.

(ii) Classical core of probability distributions, i.e., all the mixtures of pointer states. In Fig. 1 it corresponds to the section [-1,+1] of the z axis.

(iii) The rest of the space – the rest of the volume of the Bloch sphere – consists of more general density matrices. As a result of decoherence, that part of the Hilbert space is "ruled out" by einselection.

Visualization of this decoherence-induced decomposition of the Hilbert space is still possible in the simple two-dimensional case studied here, but the existence of the elements (i)–(iii) is a general feature. It characterizes the emergence of classicality under all circumstances. We shall therefore appeal to the intuitions developed in the course of this discussion later. However, it may be useful to anticipate a few of the phenomena that can take place when decoherence combines with the evolution induced by the self-Hamiltonian of the system or when it is caused by more complicated couplings to the environment.

(a) Approach to equilibrium would affect elements diagonal in the pointer basis, so that the density matrix would asymptotically approach a time-independent distribution (such as $\rho \sim 1$ for a thermal equilibrium at infinite temperature or $\rho \sim | \Downarrow\rangle\langle\Downarrow |$ for decay). This corresponds to a flow towards some specific point (i.e., the center or the "south pole" in the above two examples) within the Bloch sphere. However, when decoherence dominates, the flow would start somewhere within the Bloch sphere, and quickly (on the decoherence time scale) converge towards a point on the z axis (the classical core). This would be followed by a much slower relaxation, a flow more or less along the z axis (and therefore essentially within the classical core) on a *relaxation time scale.*

(b) Approximately reversible classical dynamics can coexist with decoherence when the self-Hamiltonian of the system can generate motions within the surfaces of constant entropy inside the classical core. In the case considered here, the core is one-dimensional and the subspaces of constant entropy within it are zero-dimensional. Therefore, it is impossible to generate continuous isentropic motion within them. In multidimensional Hilbert spaces with richer dynamics that are nearly isentropic, approximately reversible evolution is often possible and allows for the idealization of trajectories in the classical limit.

(c) A sharp distinction between the classical core and the rest of the Hilbert space is possible only in idealized situations (or in an even more idealized "mathematical classical limit", in which $\hbar \to 0$, mass$\to \infty$, etc.). In realistic situations, all that will be required is a clear contrast between the rates of the entropy production between the inside and the outside of the classical core. We shall refine such criteria in the discussion of the *predictability sieve* – a criterion for the selection of the preferred pointer states, which in effect demands that the entropy production rate should be minimized for the einselected states. In the case discussed here, pointer states obviously satisfy this criterion, and the entropy production vanishes in the classical domain. In more general situations, we shall not be equally lucky. For instance, in the case of chaotic systems, entropy will also be produced in a classical core, but at a rate set by the classical dynamics (i.e., by the self-Hamiltonian rather than by the coupling with the environment) and much more slowly than outside of the classical core.

2.6 Decoherence, Einselection, and Controlled Shifts

The above discussion of decoherence can be straightforwardly generalized to the situation where the system, the apparatus, and the environment have many states, and where the interactions between them are much more complicated. Here we assume that the system is isolated and that it interacts with the apparatus only briefly. As a result of that interaction, the state of the apparatus becomes entangled with the state of the system, $(\sum_i \alpha_i|s_i\rangle)|A_0\rangle \;\; \to \sum_i \alpha_i|s_i\rangle|A_i\rangle$. By analogy with a `c-not`, we shall refer to this conditional operation as a `c-shift`. This quantum correlation suffers from the basis ambiguity we have discussed previously: The $\mathcal{S}$-$\mathcal{A}$ entanglement implies that for any state of either of the two systems there exists a corresponding pure state of the other. Indeed, when the initial state of $\mathcal{S}$ is chosen to be one of the eigenstates of the conjugate basis, $|r_l\rangle = N^{-\frac{1}{2}} \sum_{k=0}^{N-1} \exp(2\pi i k l/N)|s_k\rangle$, this `c-shift` would equally well represent a measurement of the apparatus state (in the basis conjugate to $\{|A_k\rangle\}$) by the system [27]. Thus, it is not just the basis that is ambiguous, but also the roles of the control (system) and of the target (apparatus) can be reversed when the conjugate basis is selected. These ambiguities that exist for the $\mathcal{SA}$ pair can be removed by recognizing the role of the environment.

Decoherence is represented schematically in Fig. 2 by a sequence of `c-nots` (or `c-shifts`) which, in some fixed basis, 'measure' the state of the apparatus and record the outcome of the measurement in the environment. The requirement for a good apparatus is to retain correlations between the measured observable of the system and some "pointer observable". This will happen when the `c-shift` between $\mathcal{S}$ and $\mathcal{A}$ correlates the state of the system with the observable of the apparatus that is itself monitored (but not perturbed) by the environment. That is, in an idealized measurement, the measured observable of the system is playing the role of the control with respect to the $\mathcal{S}$-$\mathcal{A}$ `c-shift`. In a well–designed apparatus, the pointer observable is a target of the $\mathcal{S}$-$\mathcal{A}$ `c-shift`, but a control of the $\mathcal{A}$-$\mathcal{E}$ `c-shifts`. Eigenstates of the pointer observable of the apparatus play the role of an alphabet of a communication channel. They encode a state

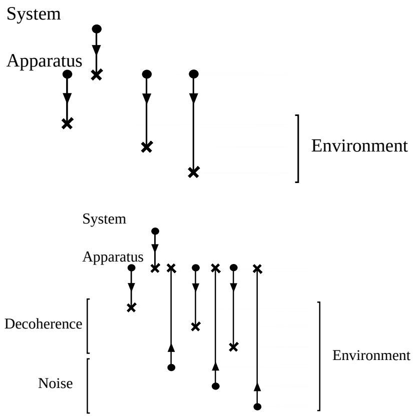


Fig. 2. (a) Decoherence can be viewed as the consequence of the monitoring of the state of the system by the environment. This is symbolically represented here by a sequence of `c-not` gates where the pointer states of the apparatus act as the control and the environment is the target. (b) The distinction between decoherence and noise depends on the direction of the information flow in the preferred basis. Preferred states minimize the number of `c-nots` directed from the environment.

of the system and retain the correlation in spite of the interaction with the environment.

The graph in Fig. 2 captures the essence of the idealized decoherence process, which yields – in spite of the interaction with the environment – a noiseless classical communication channel [28,29]. This is possible because in the pointer basis, the $\mathcal{A}$-$\mathcal{E}$ `c-shifts` operate without disturbing the pointer observable, which is the constant of motion of the $\mathcal{A}$-$\mathcal{E}$ interaction Hamiltonian.

The advantage of the graphical representation of the decoherence process as a sequence of `c-shifts` lies in its simplicity and suggestiveness. However, the actual process of decoherence is usually caused by a continuous interaction (so that it can be only approximately broken up into discrete `c-shifts`). Moreover, in contrast to the `c-nots` used in quantum logic circuits, the record inscribed in the environment is more often than not distributed over many degrees of freedom. Last but not least, the observable of the apparatus (or any other open system) may be a subject to noise (and not just decoherence) or it may evolve in a manner that will rotate pointer states into their superpositions.

The basic physics of decoherence is a simple premeasurement–like process carried out by the environment $\mathcal{E}$ as a result of the interaction with the apparatus,

$$\begin{aligned} |\Psi_{\mathcal{SA}}\rangle|\varepsilon_0\rangle &= (\sum_j \alpha_j |s_j\rangle|A_j\rangle)|\varepsilon_0\rangle \\ &\longrightarrow \sum_j \alpha_j |s_j\rangle|A_j\rangle|\varepsilon_j\rangle \;=\; |\Phi_{\mathcal{SAE}}\rangle . \end{aligned} \tag{33}$$

Decoherence leads to the einselection when the states of the environment $|\varepsilon_j\rangle$ corresponding to different pointer states become orthogonal,

$$\langle\varepsilon_i|\varepsilon_j\rangle = \delta_{ij}. \tag{34}$$

When this orthogonality condition is satisfied, Schmidt decomposition of the state vector $|\Phi_{\mathcal{SAE}}\rangle$ into a composite subsystem $\mathcal{SA}$ and $\mathcal{E}$ yields product states $|s_j\rangle|A_j\rangle$ as partners of the orthogonal environment states. The density matrix describing the correlated but decohered $\mathcal{SA}$ pair is then:

$$\rho^D_{\mathcal{SA}} = \sum_j |\alpha_j|^2 |s_j\rangle\langle s_j||A_j\rangle\langle A_j| = \mathrm{Tr}_{\mathcal{E}}|\Phi_{\mathcal{SAE}}\rangle\langle\Phi_{\mathcal{SAE}}| \ . \tag{35}$$

The reduced density matrix of the $\mathcal{SA}$ pair is diagonal in the product states.

For notational simplicity, we shall often discard reference to the object that does not interact with the environment (here, the system $\mathcal{S}$). Nevertheless, it is useful to keep in mind that the preservation of the $\mathcal{SA}$ correlations is the criterion used to define the pointer basis. The density matrix of a single object evolving in contact with the environment will be always diagonal in the same (instantaneous) Schmidt basis. This instantaneous diagonality should not be used as a sole criterion for classicality (although see [31,32]; as well as [33,34]). Rather, the ability of certain sets of states to retain correlations in spite of the coupling to the environment is decisive in the emergence of "classical reality". This is especially obvious in quantum measurements.

When the interaction with the apparatus has the form

$$H_{\mathcal{AE}} = \sum_{k,l,m} g^{\mathcal{AE}}_{klm} |A_k\rangle\langle A_k||\varepsilon_l\rangle\langle\varepsilon_m| + h.c. \ , \tag{36}$$

the basis $\{|A_k\rangle\}$ is left unperturbed. Then, any correlation with the states $\{|A_k\rangle\}$ will be preserved. And, by definition, the states that preserve correlations will be the pointer states. Any observable A co-diagonal with the interaction Hamiltonian will be an effective pointer observable. For, when the Hamiltonian depends on A, it will commute with A,

$$[H_{\mathcal{AE}}(A), A] = 0. \tag{37}$$

Moreover, the dependence of the interaction Hamiltonian on the observable is an obvious precondition for the monitoring of that observable by the environment.

3 Dynamics of Quantum Open Systems: Master Equations

One of the most practical tools for analyzing the dynamics of a quantum open system is the evolution equation for the reduced density matrix, known as the "master equation". In this section we will review some of the most common techniques for obtaining such an equation. As usual, we divide our universe into a system of interest $\mathcal{S}$ that interacts with an environment $\mathcal{E}$. The reduced density matrix of the system is the operator that allows us to answer all physical questions that concern the system $\mathcal{S}$ only. We will denote the reduced density matrix as ρ, which is obtained from the total density matrix of the universe by tracing over the environment Hilbert space. Thus,

$$\rho = \mathrm{Tr}_{\mathcal{E}} \rho_{\mathcal{T}},$$

where the total density matrix is denoted as $\rho_{\mathcal{T}}$.

In principle, the evolution equation for ρ could be obtained by solving the Schrödinger (or von Neumann) equation for the total density matrix and then taking the trace. However, this task can be analytically completed in very few cases, and the study of the evolution of the reduced density matrix should be done by using some approximations.

This section is divided in two parts. First we review some of the standard techniques used for obtaining approximate master equations. Our plan is not to give a complete review of master equation techniques but to present some useful tools to be applied later in studying decoherence. We do this not only to ensure that the paper is self–contained but also because we think it might be useful to present some simple and helpful results that are not so well known. We focus on the simplest approximation scheme, obtaining master equations valid to a second order in a perturbative expansion in the system–environment coupling strength. We first review the general perturbative scheme and apply it to two physically interesting examples: (1) The Brownian motion of a particle coupled to an environment of independent oscillators, and (2) a quantum particle locally coupled to an environment formed by a quantum scalar field. As a further illustration of the way in which perturbative master equations can be obtained, we find the corresponding equations for a two–level system coupled to a bosonic bath in two physically relevant cases (the decay of a two–level atom and the spin-boson model).

In the second part of this section we review the properties of an important model that is amenable to an exact solution. Thus, we concentrate on the linear quantum Brownian motion model analyzing the properties of its exact master equation. In particular, we stress the fact that in this simple but physically relevant model, the exact master equation has the same functional form as the one obtained using perturbation theory and can always be cast in terms of a local differential equation with time-dependent coefficients.

3.1 Master Equation: Perturbative Evaluation

Here we present the general procedure that can be used to derive the master equation, assuming that the system–environment coupling is small. Thus, we sketch a textbook derivation of the master equation using perturbation theory. We think it is convenient to present this derivation just to stress the fact that perturbative master equations can always be shown to be local in time. The calculation we follow is closely related to the one presented, for example, in [35] and can be seen to be a variant of the time-convolutionless method discussed in [36].

Let us consider the total Hamiltonian to be

$$H = H_{\mathcal{S}} + H_{\mathcal{E}} + V,$$

where $H_{\mathcal{S}}$ and $H_{\mathcal{E}}$ are respectively the self–Hamiltonian of the system and the environment and V is the interaction term. The equation for the complete density matrix $\rho_{\mathcal{T}}$, in the interaction picture, reads (we use a tilde to denote operators in the interaction picture),

$$\mathrm{i}\hbar\dot{\tilde{\rho}}_{\mathcal{T}} = [\tilde{V}(t), \tilde{\rho}_{\mathcal{T}}], \tag{38}$$

where the interaction potential and density matrix are $\tilde{V}(t) = U_0^{\dagger} V U_0$ and $\tilde{\rho}_{\mathcal{T}} = U_0^{\dagger} \rho_{\mathcal{T}} U_0$, with $U_0 = \exp(-\mathrm{i}(H_{\mathcal{S}} + H_{\mathcal{E}})t/\hbar)$. Solving (38) perturbatively is rather straightforward and leads to the Dyson series,

$$\tilde{\rho}_{\mathcal{T}}(t) = \sum_{n \geq 0} \int_0^t \mathrm{d}t_1 \ldots \int_0^{t_{n-1}} \mathrm{d}t_n (\frac{1}{\mathrm{i}\hbar})^n [\tilde{V}(t_1), \ldots, [\tilde{V}(t_n), \tilde{\rho}_{\mathcal{T}}(0)]].$$

We can use this to compute the reduced density matrix to second order. To obtain the master equation we compute the time derivative of the resulting expression and perform the trace over the environment. We get

$$\dot{\tilde{\rho}} = \frac{1}{\mathrm{i}\hbar} \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), \rho_{\mathcal{T}}(0)] - \frac{1}{\hbar^2} \int_0^t \mathrm{d}t_1 \, \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), [\tilde{V}(t_1), \rho_{\mathcal{T}}(0)]]. \tag{40}$$

So far, the only assumption we made was the validity of a perturbative expansion up to second order. Now we will assume that the initial state is not entangled, i.e., that the total density matrix is a tensor product of the form $\rho_{\mathcal{T}}(0) = \rho(0) \otimes \rho_{\mathcal{E}}(0)$. Substituting this into (40) we find,

$$\begin{aligned} \dot{\tilde{\rho}} = & \frac{1}{\mathrm{i}\hbar} \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), \rho(0) \otimes \rho_{\mathcal{E}}(0)] \\ & - \frac{1}{\hbar^2} \int_0^t \mathrm{d}t_1 \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), [\tilde{V}(t_1), \rho(0) \otimes \rho_{\mathcal{E}}(0)]]. \end{aligned} \tag{41}$$

To finish the derivation, we make a rather trivial observation that enables us to rewrite the master equation in a very simple way: The initial state $\rho(0)$ that appears in the right-hand-side of (41) could again be expressed in terms of $\tilde{\rho}(t)$ using the same perturbative expansion that enabled us to obtain (41). By doing this we can rewrite the right-hand-side of the master equation entirely in terms of the reduced density matrix evaluated at time t. The resulting equation is

$$\begin{aligned} \dot{\tilde{\rho}} = & \frac{1}{\mathrm{i}\hbar} \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), \tilde{\rho} \otimes \rho_{\mathcal{E}}(0)] - \frac{1}{\hbar^2} \int_0^t \mathrm{d}t_1 \mathrm{Tr}_{\mathcal{E}} [\tilde{V}(t), [\tilde{V}(t_1), \tilde{\rho} \otimes \rho_{\mathcal{E}}]] \\ & + \frac{1}{\hbar^2} \int_0^t \mathrm{d}t_1 \mathrm{Tr}_{\mathcal{E}} \Big([\tilde{V}(t), \mathrm{Tr}_{\mathcal{E}} \big([\tilde{V}(t_1), \tilde{\rho} \otimes \rho_{\mathcal{E}}] \big) \otimes \rho_{\mathcal{E}}] \Big). \end{aligned} \tag{42}$$

This, when rewritten in the Schrödinger picture, is the basic master equation we will use in this section. It is important to keep in mind that to derive it, we only made two important assumptions: (a) we used a perturbative expansion up to second order in the system–environment coupling constant and (b) we assumed uncorrelated initial conditions.

Below, we will apply this equation to study three interesting examples. Before doing that, let us stress that the master equation is local in time even though to obtain it no Markovian assumption was made (see below). Moreover, this rather simple form can be simplified further by assuming that the system–environment coupling is of the form

$$V = \sum_n \big(S_n E_n + S_n^\dagger E_n^\dagger\big), \tag{43}$$

where S_n (E_n) are operators acting on the Hilbert space of the system (environment) only. In such case, the master equation in the Schrödinger picture can be written as

$$\begin{aligned}\dot{\rho} &= \frac{1}{\mathrm{i}\hbar}[H_{\mathcal{S}},\rho] + \frac{1}{\mathrm{i}\hbar}\sum_n [\langle E_n\rangle S_n + \langle E_n^\dagger\rangle S_n^\dagger, \rho] \\ &\quad - \frac{1}{2\hbar^2}\sum_{nm}\int_0^t \mathrm{d}t_1 \Big(K_{nm}^{(1)}(t,t_1)[S_n,[S_m^\dagger(t_1-t),\rho]] \\ &\quad + K_{nm}^{(2)}(t,t_1)[S_n,\{S_m^\dagger(t_1-t),\rho\}] + K_{nm}^{(3)}(t,t_1)[S_n,[S_m(t_1-t),\rho]] \\ &\quad + K_{nm}^{(4)}(t,t_1)[S_n,\{S_m(t_1-t),\rho\}] + h.c.\Big),\end{aligned} \tag{44}$$

where the bracket notation indicates the expectation value over the initial state of the environment and the kernels $K_{nm}^{(i)}$ are simply determined by the two time correlation functions of the environment as follows:

$$\begin{aligned}K_{nm}^{(1)}(t,t_1) &= \frac{1}{2}\langle\{E_n(t),E_m^\dagger(t_1)\}\rangle - \langle E_n\rangle\langle E_m^\dagger\rangle, \\ K_{nm}^{(2)}(t,t_1) &= \frac{1}{2}\langle[E_n(t),E_m^\dagger(t_1)]\rangle, \\ K_{nm}^{(3)}(t,t_1) &= \frac{1}{2}\langle\{E_n(t),E_m(t_1)\}\rangle - \langle E_n\rangle\langle E_m\rangle, \\ K_{nm}^{(2)}(t,t_1) &= \frac{1}{2}\langle[E_n(t),E_m(t_1)]\rangle.\end{aligned} \tag{45}$$

At this point, it is interesting to consider another important approximation that is usually employed in this context, i.e., the Markovian approximation that we have refrained from using so far. The Markovian approximation corresponds to considering cases for which the kernels $K^{(i)}$ are strongly peaked about $t = t_1$. When this is the case, i.e. when the environment has a very short correlation time, one can transform the temporal integrals into integrals over the variable $\tau = t - t_1$, which can then be extended over the entire interval $[0,\infty)$. As we mentioned above, so far, we have not used the Markovian assumption and therefore the above equations are valid even if the environment has a long correlation time and the kernels $K^{(i)}$ are not strongly peaked. In the examples below, we will mention some cases where this happens and use the above equation to study decoherence produced by a non Markovian environment.

It is also worth mentioning that to go one step beyond equation (44), one needs to know the temporal dependence of the free Heisenberg operators of the system (i.e., $S_n(t)$) which obviously depend on the Hamiltonian H_S that we have not specified so far. We will do so in some concrete examples below.

3.2 Example 1: Perturbative Master Equation in Quantum Brownian Motion

The system of interest is a quantum particle, which moves in a one dimensional space (generalization to higher dimensions is immediate). The environment is an ensemble of harmonic oscillators interacting bilinearly through position with the system. Thus, the complete Hamiltonian is $H = H_{\mathcal{S}} + H_{\mathcal{E}} + V$ where

$$H_{\mathcal{E}} = \sum_n (\frac{1}{2m_n} p_n^2 + \frac{1}{2} m_n \omega_n^2 q_n^2)$$

and $V = \sum_n \lambda_n q_n x$. The Hamiltonian of the system will be left unspecified for the moment (we will concentrate later on the case of a harmonic oscillator). The initial state of the environment will be assumed to be a thermal equilibrium state at temperature $T = 1/k_{\mathrm{B}}\beta$. Under these assumptions the first-order term in the master equation disappears because $\mathrm{Tr}_{\mathcal{E}}(\tilde{V}(t)\rho_{\mathcal{E}}) = 0$. Therefore, the master equation in the Schrödinger picture is

$$\dot{\rho} = \frac{1}{\mathrm{i}\hbar}[H_{\mathcal{S}}, \rho] - \frac{1}{\hbar}\int_0^t \mathrm{d}t_1 \Big(\nu(t_1)[x,[x(-t_1),\rho]] - \mathrm{i}\eta(t_1)[x,\{x(-t_1),\rho\}]\Big). \quad (47)$$

The two kernels appearing here are respectively called the noise and the dissipation kernel and are defined as

$$\begin{aligned} \nu(t) &= \frac{1}{2\hbar}\sum_n \lambda_n^2 \langle\{q_n(t), q_n(0)\}\rangle = \int_0^\infty \mathrm{d}\omega\, J(\omega)\cos(\omega t)(1+2N(\omega)), \\ \eta(t) &= \frac{\mathrm{i}}{2\hbar}\sum_n \lambda_n^2 \langle[q_n(t), q_n(0)]\rangle = \int_0^\infty \mathrm{d}\omega\, J(\omega)\sin(\omega t), \end{aligned} \quad (48)$$

where $J(\omega) = \sum_n \lambda_n^2 \delta(\omega - \omega_n)/2m_n\omega_n$ is the spectral density of the environment and $N(\omega)$ is the mean occupation number of the environmental oscillators (i.e., $1 + 2N(\omega) = \coth(\beta\hbar\omega/2)$).

Equation (47) is already very simple but it can be further simplified if one assumes that the system is a harmonic oscillator. Thus, if we consider the Hamiltonian of the system to be $H_{\mathcal{S}} = p^2/2M + M\Omega^2 x^2/2$, we can explicitly solve the Heisenberg equations for the system and determine the operator $x(t)$ to be $x(t) = x\cos(\Omega t) + \frac{1}{M\Omega} p \sin(\Omega t)$. Inserting this into (47), we get the final expression for the master equation,

$$\begin{aligned}\dot{\rho} = &-\frac{\mathrm{i}}{\hbar}\big[H_{\mathcal{S}} + \frac{1}{2}M\tilde{\Omega}^2(t)x^2, \rho\big] - \frac{\mathrm{i}}{\hbar}\gamma(t)\big[x, \{p, \rho\}\big] \\ &- D(t)\big[x, [x, \rho]\big] - \frac{1}{\hbar}f(t)\big[x, [p, \rho]\big].\end{aligned} \tag{49}$$

Here the time-dependent coefficients (the frequency renormalization $\tilde{\Omega}(t)$, the damping coefficient $\gamma(t)$, and the two diffusion coefficients $D(t)$ and $f(t)$) are

$$\begin{aligned}\tilde{\Omega}^2(t) &= -\frac{2}{M}\int_0^t \mathrm{d}t' \cos(\Omega t')\eta(t'),\ \ \gamma(t) = \frac{1}{M\Omega}\int_0^t \mathrm{d}t' \sin(\Omega t')\eta(t'), \\ D(t) &= \frac{1}{\hbar}\int_0^t \mathrm{d}t' \cos(\Omega t')\nu(t'),\ \ f(t) = -\frac{1}{M\Omega}\int_0^t \mathrm{d}t' \sin(\Omega t')\nu(t').\end{aligned} \tag{50}$$

From this equation it is possible to have a qualitative idea of the effects the environment produces on the system. First we observe that there is a frequency renormalization. Thus, the “bare” frequency of the oscillator is renormalized to $\tilde{\Omega}^2$. This term does not affect the unitarity of the evolution. The terms proportional to $\gamma(t)$, $D(t)$ and $f(t)$ bring about non unitary effects. Thus, one can easily see that the second term is responsible for producing friction ($\gamma(t)$ plays the role of a time-dependent relaxation rate). The last two are diffusion terms. The one proportional to $D(t)$ is the main cause for decoherence.

Of course, the explicit time dependence of the coefficients can only be computed once we specify the spectral density of the environment. To illustrate their qualitative behavior, we will consider a typical ohmic environment characterized by a spectral density of the form

$$J(\omega) = 2M\gamma_0 \frac{\omega}{\pi}\frac{\Lambda^2}{\Lambda^2 + \omega^2}, \tag{51}$$

where Λ plays the role of a high–frequency cutoff and γ_0 is a constant characterizing the strength of the interaction. For this environment, it is rather straightforward to find the following exact expressions for the coefficients $\tilde{\Omega}(t)$ and $\gamma(t)$:

$$\gamma(t) = \gamma_0 \frac{\Lambda^2}{\Lambda^2 + \Omega^2}\left(1 - \left(\cos(\Omega t) + \frac{\Lambda}{\Omega}\sin(\Omega t)\right)\exp(-\Lambda t)\right) \tag{52}$$

$$\tilde{\Omega}^2(t) = -2\gamma_0\Lambda \frac{\Lambda^2}{\Lambda^2 + \Omega^2}\left(1 - \left(\cos(\Omega t) - \frac{\Omega}{\Lambda}\sin(\Omega t)\right)\exp(-\Lambda t)\right). \tag{53}$$

From these equations we see that these coefficients are initially zero and grow to asymptotic values on a time scale that is fixed by the high–frequency cutoff Λ. Thus, we see the relation between this result and the one we would obtain by using a Markovian approximation simply corresponds to taking the limit $\Lambda \to \infty$. In such a case both coefficients are not continuous at $t = 0$ and jump to constant values (the frequency renormalization diverges as it is proportional to the product $\gamma_0\Lambda$).

The time dependence of the diffusion coefficients can also be studied for the above environment. However, the form of the coefficients for arbitrary temperature is quite complicated. To analyze the qualitative behavior, it is convenient to evaluate them numerically. In Fig. 3 one can see the dependence of the coefficients (for both the long and short time scales) for several temperatures (high and low). We observe that both coefficients have an initial transient where they exhibit a behavior that is essentially temperature independent (over periods of time comparable with the one fixed by the cutoff). The direct diffusion coefficient $D(t)$ after the initial transient rapidly settles into the asymptotic value given by $D_\infty = M\gamma_0 \Omega \coth(\beta\hbar\Omega/2)\Lambda^2/\hbar(\Lambda^2+\Omega^2)$. The anomalous diffusion coefficient $f(t)$ also approaches an asymptotic value (which for high temperatures is suppressed with respect to D_∞ by a factor of Λ), but the approach is algebraic rather than exponential. More general environments can be studied using our equation. In fact, the behavior of the coefficients is rather different for environments with different spectral content. This has been analyzed in the literature, in particular in relation to decoherence [37].

It is interesting to mention that the master equation (49) (although it has been derived perturbatively) can be shown to be very similar to its exact counterpart whose derivation we will discuss later in this section.

3.3 Example 2: Perturbative Master Equation for a Two-Level System Coupled to a Bosonic Heat Bath

As a second example we obtain the perturbative master equation for a two–level system coupled to an oscillator environment. We consider two different models characterized by different interaction Hamiltonians. First, we discuss the model describing the physics of the decay of a two–level atom (in the rotating wave approximation),

$$H = \frac{1}{2}\hbar\Delta\sigma_z + \sum_n \lambda_n \left(a_n\sigma_+ + a_n^\dagger\sigma_-\right) + \sum_n \hbar\omega_n a_n^\dagger a_n, \tag{54}$$

where a_n and $a_n^\dagger$ are annihillation and creation operators of the environment oscillators, and $\sigma_\pm$ are the raising and lowering operators of the two–level system. The perturbative master equation obtained following the procedure described above is

$$\begin{aligned}\dot\rho &= \frac{1}{\mathrm{i}\hbar}[H_{\mathcal{S}},\rho] \\ &\quad - \frac{1}{2\hbar^2}\int_0^t \mathrm{d}t_1 k(t_1)\left([\sigma_+,[\sigma_-(-t_1),\rho]] + [\sigma_+,\{\sigma_-(-t_1),\rho\}]\right) + h.c.),\end{aligned}$$

where the kernel $k(t)$ is defined as

$$k(t) = \sum_n \lambda_n^2 \langle [a_n(t), a_n^\dagger]\rangle = \sum_n \lambda_n^2 \exp(-\mathrm{i}\omega_n t).$$

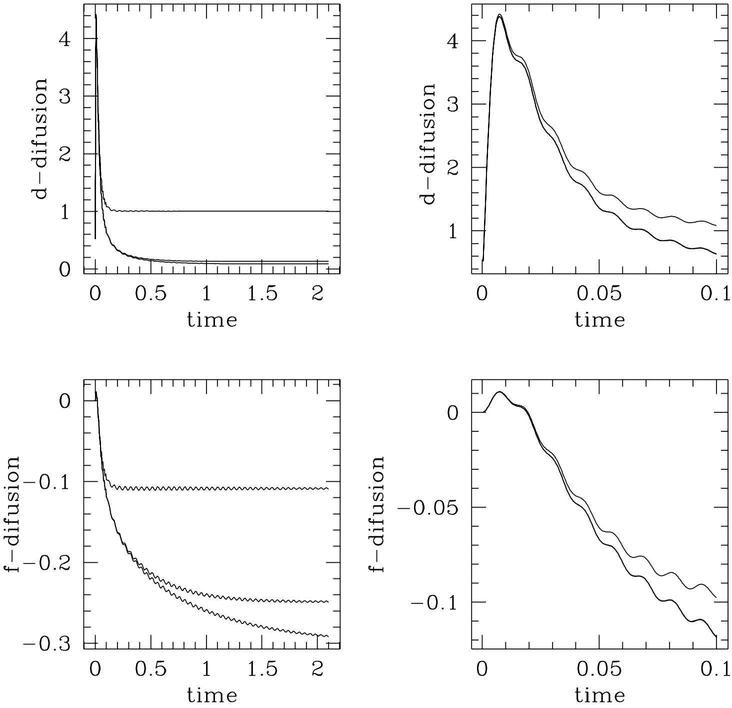


Fig. 3. Time dependence of the diffusion coefficients of the perturbative master equation for quantum Brownian motion. Plots on the right show that the initial transient is temperature independent (different curves correspond to different temperatures, higher temperatures produce higher final values of the coefficients). Plots on the left show that the final values of the coefficients are strongly dependent on the temperature of the environment. The parameters used in the plot (where time is measured in units of $1/\Omega$) are $\gamma/\Omega = 0.05$, $\Lambda/\Omega = 100$, $k_{\mathrm{B}}T/\hbar\Omega = 10, 1, 0.1$.

Using the solution of the free Heisenberg equations for the spin operator (i.e., $\sigma_\pm(t) = \sigma_\pm \exp(\pm \mathrm{i}\Delta t)$), we can deduce that the master equation is

$$\dot\rho = \frac{1}{\mathrm{i}\hbar}[\hbar\left(\frac{\Delta}{2} - c(t)\right), \rho] + a(t)\left(\sigma_+\sigma_-\rho + \rho\sigma_+\sigma_- - 2\sigma_-\rho\sigma_+\right),$$

where the time-dependent coefficients are

$$a(t) = 2\mathrm{Re}f(t), \qquad c(t) = \mathrm{Im}(f(t)),$$

with

$$f(t) = \frac{1}{2\hbar^2}\int_0^t dsk(s)\exp(\mathrm{i}\Delta s). \tag{58}$$

We recognize in this equation similar features to those present in the one for quantum Brownian motion (QBM). The interaction with the environment on the one hand renormalizes the Hamiltonian of the particle through the term $c(t)$

(including thermal fluctuations, we could verify that $c(t)$ is generally temperature dependent, as opposed to the QBM case). The non-Hermitian part has a zero temperature contribution that is responsible for the spontaneous decay of the two–level system. The decay rate is determined by $b(t)$ and has a time dependence that is essentially the same as the one found for the diffusion coefficient in the zero temperature QBM case analyzed above. The finite temperature contributions can be shown to be responsible not only for the changes in the value of the decay rate $b(t)$ (which in that case would account also for the induced decay) but also for adding new terms to the master equation that take into account the induced absorption.

Finally we obtain the perturbative master equation for the spin boson Hamiltonian, which is also widely used in various condensed–matter physics problems (and was thoroughly studied in the nonperturbative regime in [38])

$$H = \frac{1}{2}\hbar\Delta\sigma_x + \sigma_z\sum_n \lambda_n q_n + \sum_n \hbar\omega_n a_n^\dagger a_n, \tag{59}$$

where q_n are the coordinates of the environmental oscillators. The master equation can be shown to be

$$\dot{\rho} = \frac{1}{\mathrm{i}\hbar}[H_{\mathcal{S}},\rho] - \frac{1}{\hbar}\int_0^t \mathrm{d}t_1\Big(\nu(t_1)[\sigma_z,[\sigma_z(-t_1),\rho]] - \mathrm{i}\eta(t_1)[\sigma_z,\{\sigma_z(-t_1),\rho\}]\Big),$$

where the two kernels are the same as defined above in the QBM case (48). Using the free Heisenberg operator $\sigma_z(t) = \sigma_z\cos(\Delta t) + \sigma_y\sin(\Delta t)$ we obtain the master equation,

$$\dot{\rho} = \frac{1}{\mathrm{i}\hbar}[H_{\mathrm{eff}},\rho] - \tilde{D}(t)[\sigma_z,[\sigma_z,\rho]] + z(t)\sigma_z\rho\sigma_y + z^*(t)\sigma_y\rho\sigma_z \ ,$$

where the effective Hamiltonian and the time-dependent coefficients are now given by

$$\begin{aligned} H_{\mathrm{eff}} &= \hbar\left(\frac{\Delta}{2} - z^*(t)\right)\sigma_x, \\ \tilde{D}(t) &= \int_0^t ds\nu(s)\cos(\Delta s), \qquad z(t) = \int_0^t ds\,(\nu(s) - \mathrm{i}\eta(s))\sin(\Delta s). \end{aligned}$$

As before, the interpretation is quite straightforward. The effect of the environment is to renormalize the frequency as well as to introduce the decay of the system. This effect takes place only if the bare frequency Δ is nonzero (otherwise $z(t)$ vanishes). The other effect of the environment is to destroy the nondiagonal terms in the density matrix, a task that is carried out by the term proportional to $\tilde{D}$, which is present even when the bare driving vanishes. As before, the expression for the time-dependent coefficients is qualitatively similar to the one observed in the QBM model.

3.4 Example 3: Perturbative Master Equation for a Particle Interacting with a Quantum Field

We consider the following simple model: The system is a particle with position $\boldsymbol{x}$ (moving in a 3-dimensional space) and the environment is a quantum scalar field ϕ. The interaction between them is local as described by the Hamiltonian $V = e\phi(\boldsymbol{x})$, where e is the coupling constant (the "charge" of the particle). Expanding the scalar field in normal modes, the Hamiltonian can be written as $V = \int d\boldsymbol{k}(h_{\boldsymbol{k}} \exp(\mathrm{i}\boldsymbol{kx}) + h.c.)$ where the Fourier components $h_{\boldsymbol{k}}$ are proportional to annihilation operators of the quantum field (i.e., $h_{\boldsymbol{k}} = e\ a_{\boldsymbol{k}}/(2\pi)^{/2}(2\omega_k)^{1/2}$). More generally, we could consider models in which the particle–field interaction is slightly nonlocal taking into account the finite extent of the particle (thus, a nonrelativistic treatment of the quantum particle would only give consistent results if we do not attempt to localize it beyond its Compton wavelength). In this case, the interaction Hamiltonian $\tilde{H}_{\mathrm{int}} = e\int d\boldsymbol{y}W(\boldsymbol{x}-\boldsymbol{y})\phi(\boldsymbol{y})$ depends upon the window function $W(\boldsymbol{r})$ whose support lies inside a sphere of radius R (the Compton radius of the particle) centered around the origin. This nonlocal interaction corresponds to a Hamiltonian whose Fourier components $h_{\boldsymbol{k}}$ are multiplied by $\hat{W}(\boldsymbol{k})$ (the Fourier transform of $W(\boldsymbol{r})$). As we can see, the net effect of taking into account the finite size of the particle is to introduce an ultraviolet cutoff in the scalar field (the particle does not interact with the field modes with frequencies higher than corresponding to its rest mass).

It is interesting to note that for this class of models we can also derive a master equation for the reduced density matrix of the particle. Thus, using the perturbative approach described above, we simply obtain (assuming the initial state of the quantum field is thermal equilibrium) the master equation as follows:

$$\begin{aligned}\dot{\rho} = -\frac{\mathrm{i}}{\hbar}[H,\rho] - \frac{e^2}{\hbar^2}\int d\boldsymbol{k}\int_0^t \mathrm{d}t_1 \Big(& G_H(\boldsymbol{k},t_1)\big[e^{\mathrm{i}\boldsymbol{kx}}, \big[e^{-\mathrm{i}\boldsymbol{kx}(-t_1)},\rho\big]\big] \\ & - \mathrm{i}G_R(\boldsymbol{k},t_1)\big[e^{\mathrm{i}\boldsymbol{kx}}, \big\{e^{-\mathrm{i}\boldsymbol{kx}(-t_1)},\rho\big\}\big]\Big). \end{aligned} \tag{62}$$

Here, $\boldsymbol{x}(t)$ is the Heisenberg position operator for the particle (evolved with the free Hamiltonian H) and $G_{R,H}(\boldsymbol{k},t)$ are the Fourier transforms of the retarded and symmetric two–point functions of the scalar field (multiplied by the appropriate window function if the interaction is nonlocal). When the environment is a free field, we have

$$\begin{aligned} G_R(\boldsymbol{k},t) &= W(\boldsymbol{k})\sin(\omega_{\boldsymbol{k}}t)/2\omega_{\boldsymbol{k}}\ , \\ G_H(\boldsymbol{k},t) &= W(\boldsymbol{k})\cos(\omega_{\boldsymbol{k}}t)(1+2N_k)/2\omega_{\boldsymbol{k}}\ , \end{aligned} \tag{63}$$

where N_k is the number density of particles in the initial state of the quantum field (the above result is valid if the field is not free, in which case the propagators are appropriately dressed). This master equation is extremely rich. Here, we will use it for two main purposes. On the one hand, we can see that the Quantum Brownian Motion case is a special limit of this particle–field model that arises in the so-called dipole approximation. This is the most widely used approximation in this context and is valid whenever the dominant wavelengths in the environment are much larger than the length scale over which the position of the particle varies. If this is the case, we can expand the exponentials up to second order ($\boldsymbol{kx} \ll 1$) and obtain:

$$\dot{\rho} = -\frac{\mathrm{i}}{\hbar}[H,\rho] - \frac{e^2}{\hbar^2}\int_0^t \mathrm{d}t_1 \Big(F_H(t_1)\big[\boldsymbol{x},\,[\boldsymbol{x}(-t_1),\rho]\big] \\ -\; iF_R(t_1)\big[\boldsymbol{x},\{\boldsymbol{x}(-t_1),\rho\}\big]\Big),$$

where $F_{R,H}(t_1) = \int d\boldsymbol{k}\boldsymbol{k}^2 G_{R,H}(\boldsymbol{k},t_1)/N(2\pi)^{3/2}$. Thus, our first example of a linear Brownian particle coupled to an oscillator environment arises as the dipole approximation of the particle field model. With this in mind, we will use the particle field model as an example to show that some of the results obtained in the QBM case are just artifacts of the dipole approximation. In particular, this will be the case with the dependence of the decoherence rate on distance. Using the master equation of our particle–field model we will easily show that the decoherence rate does not indefinitely grow with distance but exhibits saturation.

3.5 Exact Master Equation for Quantum Brownian Motion

After presenting some simple perturbative master equations one may wonder under what circumstances are they a reasonable approximation. To partially address this issue, it is interesting to compare these equations with the ones that can be obtained for exactly solvable problems. In particular, we describe the master equation for a model that has been thoroughly studied in connection with decoherence, i.e., the linear quantum Brownian motion. Thus, because the Hamiltonian is quadratic both in the coordinates of the system and the environment, it is not surprising that it can be exactly solved. In this subsection, we will describe a simple derivation of the exact master equation, discuss its main features, and show that its functional form is the same as the one obtained by using perturbation theory. Indeed, the exact master equation has the same functional form as (49), the only difference being that the time dependence of the coefficients is different in general, as expected.

It is interesting to note that the exact master equation for QBM has only been found recently in spite of the simplicity of the model (in particular, the fact that it can always be written as an equation that is local in time was not appreciated until very recently [37]). Unfortunately, the derivation of the exact master equation is not so simple and, to say the least, the original one presented in [37] is indeed rather complicated. Here we will present the simplest derivation of the exact master equation that we know of, which is done following the method proposed first in [39]. Previous studies of the master equation for QBM, obtained under various approximations, include the celebrated paper by Caldeira and Legget [40] among others (see also [42,41]).

The derivation will focus on properties of the evolution operator for the reduced density matrix. This operator will be denoted as J and is defined as the one that enables us to find the reduced density matrix at some arbitrary time from the initial one. Thus, by definition, this operator satisfies:

$$\rho(x,x,t) = \int dx_0 \int dx_0' J(x,x',t;x_0,x_0',t_0)\rho(x_0,x_0',t_0) \; . \tag{64}$$

The derivation of the exact master equation has two essential steps. The first step is to find an explicit form for the evolution operator of the reduced density matrix. The second step is to use this explicit form to obtain the master equation satisfied by the reduced density matrix. To make our presentation simpler, we postpone the proof of the first step, which will be done below using path integral techniques. Here, we first want to demonstrate how to obtain the master equation once we know the explicit form of the evolution operator. So, let us show what the evolution operator for the reduced density matrix looks like. For linear QBM we will show later that it can always be written as

$$\begin{aligned} J(X,Y,t;X_0,Y_0,t_0) = \frac{b_3}{2\pi} &\exp(\mathrm{i}\,(b_1XY + b_2X_0Y - b_3XY_0 - b_4X_0Y_0)) \\ &\times \exp\left(-a_{11}Y^2 - a_{12}YY_0 - a_{22}Y_0^2\right) , \end{aligned} \tag{65}$$

where for notational convenience we are using sum and difference coordinates (i.e., $X = x + x'$, $Y = x - x'$, etc) and the coefficients b_i and a_{jl} are time-dependent functions whose explicit form will be given below (and depend on the properties of the environment). Thus, the evolution operator (65) is simply a Gaussian function of its arguments with time-dependent coefficients. This comes as no surprise because the problem is linear.

Knowing the propagator for the reduced density matrix, it is easy to obtain the master equation following the simple method described in [39]. This is the second step of the derivation of the master equation and is done as follows. We compute the temporal derivative of the propagator J noting that the only time dependence is through the coefficients b_i and a_{jl}. Thus, we obtain

$$\begin{aligned} \dot{J} = \Big(&\frac{\dot{b}_3}{b_3} + \mathrm{i}(\dot{b}_1XY + \dot{b}_2X_0Y + \dot{b}_3XY_0 + \dot{b}_4X_0Y_0) \\ &- \dot{a}_{11}Y^2 - \dot{a}_{12}YY_0 - \dot{a}_{22}Y_0^2\Big)J \; . \end{aligned} \tag{66}$$

Using this equation, we can try to find the master equation through multiplying by the initial density matrix and integrating this over the initial coordinates. The master equation would be trivially obtained in this way if, after multiplying by the initial density matrix, we could integrate over all the initial coordinates. This is straightforward, with some of the terms appearing in (66) but it is not so obvious how to handle terms that explicitly depend upon the initial coordinates X_0 and Y_0. Fortunately, there is a simple trick that we can use: because we know that the propagator (65) is Gaussian, we can make use of this fact to obtain the following simple relations:

$$Y_0 J = \left(\frac{b_1}{b_3} Y + \frac{\mathrm{i}}{b_3} \partial_X \right) J,$$

$$\text{and} \qquad X_0 J = \left(-\frac{b_1}{b_2} X - \frac{\mathrm{i}}{b_2} \partial_Y - \mathrm{i}(\frac{2a_{11}}{b_2} + \frac{a_{12} b_1}{b_2 b_3} Y + \frac{a_{12}}{b_2 b_3}) \partial_X \right) J \ . \quad (67)$$

These two equations can be used in (66) and in this way we can express the right hand side of this equation entirely in terms of the reduced density matrix. The resulting master equation is

$$\begin{aligned} \dot\rho(x,x') = & \frac{1}{\mathrm{i}\hbar} \langle x | [H_R(t), \rho] | x \rangle - \gamma(t)(x-x')(\partial_x - \partial'_x)\rho(x,x') \\ & - D(t)(x-x')^2 \rho(x,x') + \mathrm{i} f(t)(x-x')(\partial_x + \partial'_x)\rho(x,x') \ . \end{aligned} \quad (68)$$

The coefficients appearing in this equation are determined by b_i and a_{jl} as follows:

$$\begin{aligned} \Omega^2(t) &= 2(\dot b_2 b_1/b_2 - \dot b_1) \qquad \gamma(t) = -\dot b_2/2b_2 - b_1, \\ D(t) &= \dot a_{11} - 4a_{11} b_1 + \dot a_{12} b_1/b_2 - \dot b_2(2a_{11} + a_{12} b_1/b_3)/b_2, \\ 2f(t) &= \dot a_{12}/b3 - \dot b_2 a_{12}/b_2 b_3 - 4a_{11} \ . \end{aligned} \quad (69)$$

Thus, we showed that the exact master equation is a simple consequence of the Gaussian form of the evolution operator (65). To complete our derivation of this equation we need to explicitly show how to obtain equation (65) and also find the explicit form of the time-dependent coefficients (which is also required to simplify the expressions leading to the master equation (68)).

To obtain the explicit form of the evolution operator we will follow a derivation based on the use of path integral techniques (see [44,37,43,39,45]). To understand it, very little previous knowledge of path integrals is required. The main ingredient is the path integral expression for the evolution operator of the complete wave function. Thus, if the action of the combined system is $S_T[x,q]$, the matrix elements of the evolution operator U can be written as

$$U(x,q,t;x_0,q_0,t_0) = \int DxDq \ \mathrm{e}^{\mathrm{i}S_T[x,q]} \ , \quad (70)$$

where the integration is over all paths that satisfy the boundary conditions,

$$x(0) = x_0, \ x(t) = x, \ q(0) = q_0, \ q(t) = q \ . \quad (71)$$

In the above and following equations, to avoid the proliferation of sub-indices we use q to collectively denote all the coordinates of the oscillators q_n (we will not write the subscript n that should be implicitly assumed). Using this equation, one can obtain a path integral representation of the evolution operator of the complete density matrix and, after taking the final trace over the environment, we find a path integral representation of the propagator for the reduced density matrix. It is clear that the resulting expression will involve a double path integral (one to evolve kets and another one to evolve bras). For a generic initial state $\rho_{\mathcal{T}}$, the propagator is a somewhat complicated–looking expression. To simplify our presentation, we will only consider here factorizable initial states (and refer the reader to [45] for the most general situation where initial correlations are included). Thus, if the initial state can be factored we can express the reduced density matrix at arbitrary times as a function of the reduced density matrix at initial time using a (state–independent) propagator that has the following path integral representation:

$$J(x, x', t; x_0, x'_0, t_0) = \int Dx \int Dx' \exp(\mathrm{i}S[x] - \mathrm{i}S[x'])F[x, x'] \ . \tag{72}$$

where the integral is over paths satisfying the above boundary conditions, $S[x]$ is the action for the system only, and $F[x, x']$ is the so–called "Influence Functional" first introduced by Feynman and Vernon [46]. This functional is responsible for carrying all the physical effects produced by the environment on the evolution of the system. In fact, if there is no coupling between the system and the environment, the Influence Functional is equal to the identity, and the above expression reduces to the one corresponding to the free Schrödinger evolution for the isolated system. The Influence Functional is defined as

$$F[x, x'] = \int dq dq_0 dq'_0 \rho_{\mathcal{E}}(q_0, q'_0) \int DqDq' \exp(\mathrm{i}(S_{\mathcal{SE}}[x, q] - S_{\mathcal{SE}}[q', x'])), \tag{73}$$

where $\rho_{\mathcal{E}}$ is the initial state of the environment and $S_{\mathcal{SE}}[q, x]$ is the action of the environment (including the interaction term with the system). It is easy to see that if there is no interaction (or if the two systems trajectories are the same, i.e., $x = x'$), then the influence functional is equal to one.

Calculating the Influence Functional for an environment formed by a set of independent oscillators coupled linearly to the system is a rather straightforward task (and, to our knowledge, was first done by Feynman and Vernon in [46]). Assuming the initial state of the environment is thermal equilibrium at temperature $T = 1/k_{\mathrm{B}}\beta$, the result is

$$\begin{aligned} F[x, x'] = \exp(-\mathrm{i} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 Y(t_1)\eta(t_1 - t_2)X(t_2) \\ - \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 Y(t_1)\nu(t_1 - t_2)Y(t_2)) \ , \end{aligned} \tag{74}$$

where $X = x + x'$, $Y = x - x'$, and the two kernels $\nu(s)$ and $\eta(s)$ are the so–called noise and dissipation kernels that were defined above in (48). Thus, all the influence of the environment on the evolution of the system is encoded in the noise and dissipation kernels (two different environments that produce the same kernels would be equivalent as to the impact they have on the system). To obtain the above expression is a simple exercise in path integrals. However, the calculation can also be done by a more straightforward procedure that makes no reference to path integrals. Indeed, one can notice that the influence functional can always be expressed in operator language as

$$F[x,x'] = \mathrm{Tr}_{\mathcal{E}}\ (\ T(e^{-\mathrm{i}\int_0^t \mathrm{d}t_1 V_{\mathrm{int}}[x'(t_1),q(t_1)]})\rho_{\mathcal{E}} \times \\ \tilde{T}(e^{\mathrm{i}\int_0^t \mathrm{d}t_1 V_{\mathrm{int}}[x(t_1),q(t_1)]}))\ ,$$

where T ($\tilde{T}$) denotes the time ordered (antitime ordered) product of the corresponding Heisenberg operators, and V_{int} is the interaction term between the system and the environment. If the interaction is bilinear and the initial state of the environment is thermal, one can easily realize that the result should be a Gaussian functional of both x and x'. Therefore, one can just write down such most general Gaussian functionals in terms of unknown kernels. These kernels could be identified by using the above expression, taking functional derivatives with respect to x and x' and evaluating the result when $x = x'$. In this way, one realizes that the result is given by (74), where the noise and dissipation kernels are given by expectation of symmetric and antisymmetric two–time correlation functions of the environment oscillators, exactly as in (48).

Knowing the Influence Functional enables us to compute the exact expression for the evolution operator of the reduced density matrix. In fact, all we need is to perform the path integral in (72). If the system is linear we see that the integrand is Gaussian and, therefore, the integral can also be explicitly computed. To perform this integral is not so trivial because the integrand is not separable into a product of functions of x and x'. However, the integral can be calculated simply by changing variables. First we should integrate over sum and difference coordinates X and Y. Then, we should change variables writing $X = X_{\mathrm{c}} + \tilde{X}$ and $Y = Y_{\mathrm{c}} + \tilde{Y}$ where X_{c} and Y_{c} satisfy the equations obtained by varying the phase of the integrand and imposing the corresponding boundary conditions. In this way, we show that the result of the path integral is simply the integrand evaluated in the trajectories X_{c}, Y_{c}, multiplied by a time-dependent function that can be determined by normalization. The only nontrivial part of this derivation is to realize that the trajectories X_{c} and Y_{c} can be chosen as the ones extremizing only the phase of the integrand, (and not the entire exponent that, as we saw, has a real part coming from the noise). For more details on this derivation the interested reader can look in [37,44,45]. Therefore, the final result is given in (65) where the coefficients b_i and a_{jl} are time-dependent functions that are determined in the following way. Let the functions $u_{\frac{1}{2}}$ be two solutions of the equation,

$$\ddot{u}(s) + \Omega^2 u(s) + 2\int_0^s ds' \eta(s-s')u(s') = 0\ , \tag{75}$$

satisfying the boundary conditions $u_1(0) = u_2(t) = 1$ and $u_1(t) = u_2(0) = 0$. Then, the coefficients appearing in (65) are simply given by

$$b_{\frac{1}{2}} = \frac{1}{2}\dot{u}_{\frac{2}{1}}(t), \qquad b_{\frac{3}{4}} = \frac{1}{2}\dot{u}_{\frac{2}{1}}(0)$$
$$a_{jl} = (1+\delta_{jl})^{-1}\int_0^t ds \int_0^t ds' u_j(s)u_k(s')\nu(s-s')\ . \tag{76}$$

The time dependence of the coefficients of the master equation can be investigated after specifying the spectral density and the temperature of the environment. This has been done in great detail in a series of papers [37,43,39,47]. We will not review these results in detail but would just like to mention that for the case that is most interesting for studying decoherence, which is the underdamped (i.e., weakly coupled) harmonic oscillator, the time dependence of the exact coefficients is very similar to the one obtained by analyzing the coefficients appearing in the perturbative master equation. Indeed, the perturbative coefficients obtained above can be recovered by solving the equation for the functions u_1 perturbatively and replacing these equations inside (76) and (69). Thus, to get a qualitative idea about the behavior of the coefficients, we restrict ourselves to the analysis already made for the perturbative ones (see Fig. 3).

It will be useful to analyze decoherence not only using the reduced density matrix but also the Wigner function that is the phase space distribution function that can be obtained from the density matrix as [48]

$$W(x,p) = \int_{-\infty}^{+\infty} \frac{dz}{2\pi\hbar} \mathrm{e}^{\mathrm{i}pz/\hbar} \rho(x - z/2, x + z/2). \tag{77}$$

It is simple to show that for the case of the harmonic oscillator, the evolution equation for the Wigner function can be obtained from the master equation and has the form of a Fokker Planck equation

$$\dot{W} = -\left\{H_{\mathrm{ren}}(t), W\right\}_{\mathrm{PB}} + \gamma(t)\partial_p(pW) + D(t)\partial^2_{pp}W - f(t)\partial^2_{px}W. \tag{78}$$

The form of the evolution equation for the Wigner function for more general (nonlinear) systems will be discussed in Sect. 6.

As a final remark, it is worth pointing out that the exact master equation does not have the so–called "Lindblad form". A master equation is of the Lindblad form [49] if it can be written as

$$\dot{\rho} = \frac{1}{\mathrm{i}\hbar}[H, \rho] - \sum_n \gamma_n (L_n^\dagger L_n \rho + \rho L_n^\dagger L_n - 2 L_n \rho L_n^\dagger) \ , \tag{79}$$

for some operators L_n and some (positive) constants γ_n. As shown by Lindblad, this is the most general master equation with the property of being Markovian and preserving the positivity of the density matrix. The fact that the exact master equation does not have the Lindblad form may be puzzling but after some thinking becomes natural. Of course, the exact evolution also preserves positivity of the density matrix, but it does so in a more subtle way than through a Lindblad master equation. The true evolution is not Markovian (but in a very weak sense). The only memory effect relies on the fact that the system remembers the initial time when the (factorizable) initial conditions were imposed. This effect appears in the time dependence of the coefficients that is responsible also for enforcing positivity in an interesting way (see [45,47] for some discussion on the way positivity follows from the exact master equation). As a final comment, we would like to mention the fact that exact master equations are rather rare, but the above equation for QBM is not the only interesting exact master equation known. For example, it is possible to derive an exact master equation that has strong similarities with the one for QBM (i.e., an equation that is local in time and has time-dependent coefficients) for the model of a two–level system coupled to a bosonic bath through the Hamiltonian (54) (this equation was derived first in [50] and rediscovered by other means in [51]).

4 Einselection in Quantum Brownian Motion

4.1 Decoherence of a Superposition of Two Coherent States

We will analyze here the decoherence process in a simple example: the linear quantum Brownian motion model whose exact master equation is given in (68). For this we will first set up an initial state that is delocalized in position (or momentum) space and examine its temporal evolution, paying special attention to the fate of interference effects. Thus, we will consider a state of the form [52,47]

$$\Psi(x,t=0)=\Psi_1(x)+\Psi_2(x)\ , \tag{80}$$

where

$$\Psi_{1,2}(x)=N\exp\left(-\frac{(x\mp L_0)^2}{2\delta^2}\right)\ \exp\left(\pm\mathrm{i}P_0x\right), \tag{81}$$

$$N^2\equiv\frac{\bar{N}^2}{\pi\delta^2}=\frac{1}{2\pi^2\delta^2}\left[1+\exp\left(-\frac{L_0{}^2}{\delta^2}-\delta^2P_0^2\right)\right]^{-1}\ . \tag{82}$$

Note that we assumed (just for simplicity) that the two wave packets are symmetrically located in phase space. The above expression allows us to study two extreme cases: the coherent states are separated in position or in momentum. In both cases, as a consequence of quantum interference, the Wigner function oscillates and becomes negative in some regions of phase space (and therefore cannot be interpreted as a probability distribution). When the coherent states are separated in position (momentum), the fringes are aligned along the p (x) axis.

To evolve this initial state, we should solve the master equation (68). Rather than doing this, one can use the explicit form of the evolution operator (65) and obtain the exact form of the reduced density matrix or the Wigner function at any time. We will adopt this strategy but will use the master equation (68) and the equation for the Wigner function (78) as a guide to interpret our results and to obtain simple estimates for the most important effects that take place as a result of the interaction between the system and the environment. The exact evolution of the above initial state is such that the Wigner function can be written always as the sum of two Gaussian peaks and an interference term,

$$W(x,p,t)=W_1(x,p,t)+W_2(x,p,t)+W_{\mathrm{int}}(x,p,t)\ , \tag{83}$$

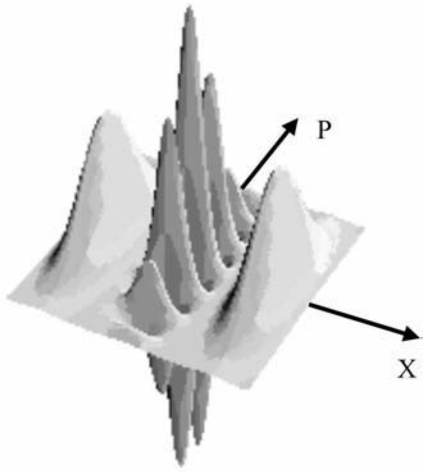


Fig. 4. Wigner function for a quantum state which is a superposition of two Gaussian wave–packets separated in position. The interference fringes are aligned along the p axis.

where

$$\begin{aligned} W_{1,2}(x,p,t) &= \frac{\bar{N}^2}{\pi}\frac{\delta_2}{\delta_1}\exp\left(-\frac{(x\mp x_{\rm c})^2}{{\delta_1}^2}\right)\exp\left(-{\delta_2}^2(p\mp p_{\rm c}-\beta(x\mp x_{\rm c}))^2\right) , \\ W_{\rm int}(x,p,t) &= 2\frac{\bar{N}^2}{\pi}\frac{\delta_2}{\delta_1}{\delta_2}^2(p-\beta x)^2 \\ &\quad\times\cos\left(2\kappa_p p+2(\kappa_x-\beta\kappa_p)x\right) . \end{aligned} \tag{84}$$

All the coefficients appearing in these expressions are somewhat complicated functions of time that are determined by the coefficients that appear in the propagator (65) and the initial state (in the same way, they also depend on temperature and on the spectral density of the environment). The initial state is such that ${\delta_1}^2={\delta_2}^2=\delta^2$, $\kappa_x=P_0=p_{\rm c}$, $\kappa_p=L_0=x_{\rm c}$ and $A_{\rm int}=0$.

From the form of the exact solution, it is clear what the qualitative behavior of the quantum state is. The two Gaussian peaks follow the two classical trajectories (which get distorted by the interaction with the environment) and change their width along their evolution. On top of this, the interference fringes change their wavelength and also rotate somewhat following the rotation of the two wave packets. The effect of decoherence is clearly manifested in the damping of the interference fringes that, in the above formulae, is produced by the exponential term $\exp(-A_{\rm int})$. Thus, we will look carefully at this term, which can be seen to be the "fringe visibility factor" defined as

$$\exp\left(-A_{\rm int}\right)=\frac{1}{2}\frac{W_{\rm int}(x,p)|_{\rm peak}}{\left(W_1(x,p)|_{\rm peak}W_2(x,p)|_{\rm peak}\right)^{1/2}} . \tag{85}$$

A close analysis of the definition of $A_{\rm int}$ shows that it vanishes initially and is always bounded from above, i.e.,

$$A_{\rm int} \leq \frac{{L_0}^2}{\delta^2} + \delta^2 {P_0}^2 = A_{\rm int}|_{\rm max} \ . \tag{86}$$

The value of $A_{\rm int}$ cannot grow to infinity as a consequence of the fact that the two Gaussian initial states have a finite overlap that is proportional to $\exp\left(-A_{\rm int}|_{\rm max}\right)$.

To understand qualitatively and quantitatively the time dependence of the fringe visibility factor, it is interesting to obtain an evolution equation for $A_{\rm int}$. Using its definition, we know that

$$\dot{A}_{\rm int} = \frac{\dot{W}_{\rm int}}{W_{\rm int}}|_{\rm peak} - \frac{1}{2}\left(\frac{\dot{W}_1}{W_1} + \frac{\dot{W}_2}{W_2}\right)|_{\rm peak} \ . \tag{87}$$

This, after using the form of the Wigner function together with the evolution equation, can be transformed into

$$\dot{A}_{\rm int} = 4D(t){\kappa_p}^2 - 4f(t)\kappa_p(\kappa_x - \beta\kappa_p) \ . \tag{88}$$

This equation enables us to obtain a clear picture of the time evolution of the fringe visibility function. Thus, we can see that the first term on the right-hand side is always positive and corresponds to the effect of normal diffusion. The normal diffusion will tend to wash out interference. The initial rate at which $A_{\rm int}$ grows is determined by the diffusion coefficient and by the initial wavelength of the fringes in the momentum direction (remember that initially we have $\kappa_p = L_0/\hbar$. As time goes by, we see that the effect of this term will be less important as the effective wavelength of the fringes grows (making κ_p decrease).

Various simple estimates of the temporal behavior of the fringe visibility factor can be obtained from this equation. The most naive one is to neglect the time dependence of the diffusion coefficient and assume that the fringes always stay more or less frozen, as in the initial state. In such a case, we have $A_{\rm int} \approx 4L_0^2 Dt/\hbar^2$. Thus, if we use the asymptotic expression of the diffusion coefficient, we obtain (at high temperatures) $A_{\rm int} \approx \gamma t 4L_0^2/\lambda_{\rm DB}^2$ where $\lambda_{\rm DB}$ is the thermal de Broglie wavelength. Consequently, we find that decoherence takes place at a rate

$$t_{\rm dec} = \gamma_0^{-1}(\lambda_{DB}/L_0)^2 \ , \tag{89}$$

which is the relaxation rate multiplied by a factor that could be very large in the macroscopic domain (this is the result originally obtained by one of us, see [52] where it is shown that for typical macroscopic parameters, i.e., room temperature, centimeter-scale distances and masses on the order of a gram, the factor $4L_0^2/\lambda_{DB}^2$ can be as large as 10^{40}).

By analyzing the temporal behavior of $A_{\rm int}$ obtained by using the exact solution, we can check that this naive estimate is an excellent approximation in many important situations. However, it may fail in other important cases. Here, we want to stress a message that we believe is very important (see [55]): It may be rather dangerous to draw conclusions that are too general from the theoretical analysis of simple models of decoherence (like the one of linear QBM). The reason is that simple estimates like the one corresponding to the decoherence time scale (89) are just that: simple estimates that apply to specific situations. They do not apply in other circumstances, some of which we will describe here (and in the next section). For example, the above simple estimate of the decoherence time scale fails in the simple case of "ultrafast" decoherence. For, in the high-temperature approximation of the master equation we neglected (among other things) initial transients occurring in the time scale fixed by the cutoff. Nothing (not even decoherence) can happen faster than the cutoff time scale since only after such time scale the diffusion coefficient reaches a sizable value. Thus, studying the initial time behavior of the normal diffusion coefficient one realizes that for very short times, A_{int} always grows quadratically (and not linearly). In fact, we have

$$A_{\mathrm{int}} \simeq \frac{4M\gamma_0 k_{\mathrm{B}} T L_0^2}{\hbar^2} \Lambda t^2 \ .$$

From this expression one sees that in this case A_{int} is smaller than the one obtained under the assumption of a constant diffusion coefficient (at least for times $t \leq \Lambda^{-1}$). In this case, the decoherence time scale may be longer than the one corresponding to the high temperature approximation,

$$t'_{\mathrm{dec}} = \frac{\hbar}{2L_0\sqrt{M\gamma_0 \Lambda k_{\mathrm{B}} T}} \ . \tag{91}$$

On the other hand, the above estimate for A_{int} also fails to take into account the fact that A_{int} does not grow forever because it finally saturates to the value fixed by (86). Saturation is achieved in a time scale that can be estimated to be $t_{\mathrm{sat}} \approx \gamma_0^{-1}(\hbar\Omega/k_{\mathrm{B}}T)$. At approximately this time the saturation of A_{int} takes place (it is clear that this is a very short time, much shorter than any dynamical time scale).

The high-temperature approximation to the behavior of A_{int} will clearly fail at very low temperatures (however, it is quite remarkable how robust an approximation this is; see [47] for a detailed analysis). We will comment in the next section about the effects arising at low temperatures giving more accurate estimates for A_{int} in such a domain.

4.2 Predictability Sieve and Preferred States for QBM

The most important consequence of the decoherence process is the dynamical selection of a set of stable, preferred states. These are, by definition, the least affected by the interaction with the environment in the sense that they are the ones that become less entangled with it. To obtain these states, a systematic ("predictability sieve") criterion has been proposed [3,53]. The basic idea is the following: To find the pointer states, one should consider all possible pure initial states for the system and compute the entropy associated with its reduced density matrix after some time t. The pointer states are the ones that minimize the entropy production for a dynamic range of times.

The predictability sieve can be applied to the simplest models of a quantum measurement, for which the Hamiltonian of the system can be completely neglected. In such a case, the pointer states are directly associated with the eigenstates of the interaction Hamiltonian (actually, to its eigensubspaces that may be degenerate). In other more realistic situations where the self–Hamiltonian of the system is not negligible the pointer states are not going to be picked only by the interaction Hamiltonian but by the interplay between it and the evolution produced by the systems own Hamiltonian. The best example where we can explicitly compute these pointer states is the QBM model we have been studying in this section. To do this, the master equation is, as we will see, a very convenient tool.

To find pointer states, we should minimize the entropy production at some time (varying over times to find a stable answer). However, to make our task simpler, instead of using the von Neumann entropy, we will simply study the evolution of the purity of the system as measured by $\varsigma = \mathrm{Tr}\rho^2$. This quantity is equal to one for a pure state and decreases when the state of the system gets mixed because entanglement is generated by the evolution. The master equation directly enables us to write down an evolution equation for the purity ς. Thus, using the definition of ς and the (49) we obtain[3]:

$$\dot{\varsigma} = 2\gamma\varsigma - 4D\mathrm{Tr}(\rho^2 x^2 - \rho x \rho x) - 2f\mathrm{Tr}(\rho^2(xp + px) - 2\rho x \rho p). \tag{92}$$

To simplify our treatment, we will once again use a perturbative approximation and substitute in the right hand side of this equation the expression for the free Heisenberg operators: $x(t) = x\cos(\Omega t) + p/M\Omega\sin(\Omega t)$ and $p(t) = p\cos(\Omega t) - M\Omega x\sin(\Omega t)$. Moreover, we will average over one period of the harmonic motion, assuming that the coefficients of the master equation do not vary during that time (clearly, this is a crude approximation, and we will comment later about what happens when we relax it). We also assume that the initial state is pure (and use the fact that in that case $\rho^2 = \rho$). Moreover, we neglect the effect of the friction term because, as we see, this term will always try to increase the purity in a way that is not sensitive to the state itself (thus, friction always tries to localize the state competing against diffusion that has the opposite effect). Doing this, we find out that the change in purity over one period is simply given by

$$\varsigma(T) - \varsigma(0) = -2D(\Delta x^2 + \Delta p^2/M^2\Omega^2)\ . \tag{93}$$

where Δx and Δp are respectively the position and momentum dispersion of the initial state. The anomalous diffusion term does not produce any net entropy increase (or purity decrease) because its effect averages out over one oscillation. The term responsible for purity decrease is simply coming from diffusion, and to minimize it, we should vary over all possible initial states. This can easily be implemented by varying over all values of the initial dispersion in position and momentum in such a way that the right–hand side of (93) is minimized. Because $\Delta x \Delta p \geq \hbar/2$ must always be satisfied, it is clear that the minimum is obtained when the state saturates uncertainty relations. From the resulting equation we obtain the pointer states as having $\Delta x^2 = \hbar/2M\Omega$ and $\Delta p^2 = \hbar M\Omega/2$. Therefore, the pointer states are simply given by coherent states with minimum uncertainty. This result is simple and satisfying. In fact, coherent states are the closest we can get to points in phase space. They are preferred states in QBM because they turn out to be the most robust, and the most effectively resist the combined effect of the system and the environment. They are also well localized in position and, therefore, are not significantly perturbed by the environment monitoring their position. Moreover, because of their symmetry, they are also not drastically altered by the evolution induced by the Hamiltonian of the system.

4.3 Energy Eigenstates Can Also Be Selected by the Environment!

So far, we have discussed two regimes in which the predictability sieve can be successfully applied. We first mentioned the case of a measurement (where the Hamiltonian of the system is negligible), and we just studied the case where both the system and the environment induce nontrivial evolution. There is a third regime that is interesting to study and is one in which the evolution of the environment is very slow as compared with the dynamic time scales of the system. If this is the case, it is possible to show [54] that the preferred states are simply the eigenstates of the Hamiltonian of the system. However, it is interesting to note that to find out this result, it is not possible to use a model like the linear QBM we described before. In fact, in such a model we can see that if we consider a very slow environment (with frequencies much smaller than the one belonging to the system) the master equation (49), which is still applicable, has time-dependent coefficients that are oscillatory functions of time with no well-defined sign. Therefore, the predictability sieve criterion does not give a robust set of states in this case.

However, the third regime of einselection can be examined using a simple argument based on an adiabatic solution of the full Schrödinger equation. The main ingredient we need is, as will be shown below, a slow environment that couples to the system through an interaction Hamiltonian that has a nonzero expectation value in the energy eigenstates of the system. To see this, we will solve the full Schrödinger equation treating the environment adiabatically. Suppose that the initial state of the universe given as $|\Psi(0)\rangle = \sum_n c_n|\phi_n\rangle|\epsilon_0\rangle$ where the states $|\phi_n\rangle$ are nondegenerate eigenstates of the Hamiltonian of the system (with distinct energies E_n), and $|\epsilon_0\rangle$ is a state of the environment that, for simplicity, we will consider as a coherent state (the vacuum, for example). We can solve the full Schrödinger equation in the adiabatic approximation and show that this state evolves into $|\Psi(t)\rangle = \sum_n c_n \exp(-\mathrm{i}E_n t/\hbar)|\phi_n\rangle|\epsilon_n(t)\rangle$ where the state $|\epsilon_n(t)\rangle$, that gets correlated with the n-th energy eigenstate of the system, obeys the following Schrödinger equation

$$\mathrm{i}\hbar\frac{d}{\mathrm{d}t}|\epsilon_n\rangle = \langle\phi_n|H_{\mathrm{int}}|\phi_n\rangle|\epsilon_n\rangle. \tag{94}$$

Note that in this equation the operator $\langle\phi_n|H_{\rm int}|\phi_n\rangle$ acts on the Hilbert space of the environment and depends parametrically on the energy eigenstates of the system. We will assume that the interaction is such that the Hamiltonian is of the form $H_{\rm int} = S \otimes \Pi_{\mathcal{E}}$, where the operator S acts on the system Hilbert space, and the environment operator $\Pi_{\mathcal{E}}$ acts on the environment as a translation generator (it could be the momentum operator, for example, but from our discussion it will be clear that the choice of momentum here is not crucial).

The decoherence in energy eigenbasis can easily be established as follows. Because $\Pi_{\mathcal{E}}$ is a momentum operator and the initial state of the environment is a coherent state, the evolution turns out to be simply such that $|\epsilon_n(t)\rangle = |\epsilon_0 + S_{nn}t\rangle$, where $S_{nn} = \langle\phi_n|S|\phi_n\rangle$. Therefore, the overlap between the two states that correlate with different energy eigenstates can be estimated as $\langle\epsilon_n(t)|\epsilon_m(t)\rangle \approx \exp(-t^2(S_{nn}-S_{mm})^2\,\hbar^2)$. Consequently, in this case, we see einselection of energy eigenstates (superpositions of energy eigenstates are degraded while pure energy eigenstates are not affected). For this reason, pointer states are energy eigenstates. This result has a rather natural interpretation. It just tells us that the environment is not able to react before the system has time to evolve and therefore only probes time-averaged quantities of the system. Energy, being the only observable that does not average out to zero is therefore the preferred observable. The conditions for energy eigenstates to become the pointer basis are the following: the environment must behave adiabatically (and be slow as compared with the dynamics of the system), and the interaction with the system must be through an observable with a nonvanishing expectation value in energy eigenstates.

5 Deconstructing Decoherence: Landscape Beyond the Standard Models

Simple models of decoherence, like the one we discussed so far (linear quantum Brownian motion) are important for illustrating the simplicity and high efficiency of the decoherence process (two characteristics that may be interpreted as indicating its generality). However, it is important to keep in mind that no generic conclusions should be drawn from simple estimates. This is especially important in view of the possibility of carrying out experiments to test decoherence in a controlled manner. In such cases, it is essential to study specific models of the decoherence process in the correct context. Estimates of the decoherence time scale, nature of pointer states, and other characteristics of decoherence obtained in models like QBM should be taken as indications rather than as strong predictions.

In this section, we would like to stress the fact that some of the simple features that have became identified as “standard lore” in the decoherence process for the simplest case of linear QBM are not generic by showing explicitly how they fail in two specific examples. We will address basically two issues. First we will consider the status of one of the simplest predictions arising from studying decoherence in linear QBM: the “decoherence rate grows quadratically with distance”. We will show that this is not the case for more realistic models where local interactions between particles and fields (rather than oscillators) are taken into account. Second, we will consider the status of predictions of the decoherence time scale like the ones in (89) at low temperature. In this case, by analyzing the same linear QBM at low temperatures, we will show that the decoherence process may be more complicated, allowing even for nonmonotonic behavior.

5.1 Saturation of the Decoherence Rate at Large Distances

One of the results obtained studying the decoherence process in linear QBM models is that the decoherence rate grows quadratically with the separation between different pieces of the system wave function. This result is natural (delocalized wave packets decohere faster) but would certainly not be physical if it held for arbitrarily large separations. Apart from any arguments involving cutoff (see the discussion following (89)), it is clear that the environment should have a coherence length so that separations that are bigger than this natural length scale should be equivalent and therefore induce the saturation of the decoherence rate.

However, saturation is not present in the linear QBM model, as is clear from the discussion above. One therefore asks what kinds of models predict saturation. We will describe here the simplest of such models. The environment is formed by a quantum scalar field; the system is a quantum particle, and the interaction between them is local. This is the model whose perturbative master equation we derived in Sect. 3.4. It is important to stress once more that the linear QBM model is obtained from the particle–field model by means of the dipole approximation. Thus, saturation in this context arises only if we do not make the dipole approximation (which is certainly not well justified for large separations). The issue of the saturation of the decoherence rate was analyzed first in [56] and also discussed in [55]. In this review we present a simpler discussion than the one of [55] that captures the main ingredients necessary for saturation and enables us to obtain the principal results without complicated calculations (some experimental results related to these issues were reported in [7]).

As we discussed in Sect. 3.4, the reduced density matrix of the particle obeys the perturbative master equation (62). In this equation, the Heisenberg operator of the particle $x(t_1)$ appears. To simplify our argument, we will consider the system that is a free and very massive particle and therefore replace $x(t) = x(0)$ in (62) (corrections to this approximation can be computed also). In the simplest example, we will consider as environment a massless scalar field (and replace the corresponding expressions for the Fourier transform of the two point functions; see (44)). In this case, we can express the master equation in the position representation as

$$\dot{\rho}(x,x') = -\mathit{\Gamma}(x-x')\rho(x,x') + \dots \ , \tag{95}$$

where only the term producing decoherence has been written out, and the function $\Gamma(x,x')$ is defined as

$$\Gamma(x,x') = -8\pi\frac{e^2}{\hbar^2}\int_0^\infty \mathrm{d}kW(k)\sin(kt)\coth(\beta k/2)(1-\frac{\sin(kr)}{kr})\ , \tag{96}$$

where $r = |x-x'|$ (and, as before, $W(k)$ is the Fourier transform of the window function that introduces a natural high–frequency cutoff). It is simplest to analyze the high–frequency limit of the above expression. In that case, the integral can be exactly computed and turns out to be

$$\begin{aligned}\Gamma(r) &= 8\pi^2\frac{e^2}{\beta\hbar^2}(\frac{\sinh(\Lambda r)}{\Lambda r}-1)\exp(-\Lambda t) \qquad \text{if } r\leq t\\ &= 8\pi^2\frac{e^2}{\beta\hbar^2}(1-\frac{t}{r}-\exp(-\Lambda t)+\frac{\sinh(\Lambda t)}{\Lambda r}\exp(-\Lambda r)) \qquad \text{if } r\geq t\end{aligned} \tag{97}$$

From this expression we clearly see the saturation. Thus, the solution of the master equation in the "decoherence dominated" approximation (neglecting all terms except the one producing decoherence) is simply $\rho(x,x') \approx \exp(-\int_0^t \mathrm{d}t_1\ \Gamma(x-x',t_1))\rho(x,x')$. The dependence of $\Gamma(r)$ for long distances is given by the second instance in (97) that approaches a constant as r grows larger than $1/\Lambda$ and t. On the other hand, the quadratic dependence of the decoherence rate is recovered for small distances: by expanding the function $\Gamma(r)$ around $r\approx 0$, we obtain a quadratic behavior.

5.2 Decoherence at Zero Temperature

A simple estimate for the low temperature behavior of the fringe visibility function can be obtained as follows. Use the asymptotic form of the diffusion coefficient for low temperatures given by perturbation theory and integrate the equation for A_{int}, neglecting both its time dependence as well as the temporal evolution of the wavelength of the fringes. In this way, we obtain $A_{\mathrm{int}} \approx \gamma_0 t(4L_0^2/\Delta x^2)\coth(\beta\Omega/2)$. However, this is not always a good approximation. On the one hand, if this behavior were correct, we could estimate the saturation time (the time for which A_{int} would approach its maximum value) to be on the order of $t_{\mathrm{sat}} \approx \gamma_0^{-1}\tanh(\beta\Omega/2)$, which for very low temperatures can be very close to, or even larger than, a dynamical time scale. Note that this does not imply that decoherence occurs in a dynamical time scale: For that, the important fact is the actual value of A_{int} and not how close to the maximum value we are. The decoherence time-scale at low temperatures is on the order of $t_{\mathrm{dec}} \approx \gamma_0^{-1}(\Delta x/2L_0)^2$, which is still much shorter than γ_0^{-1} for macroscopic parameters). The fact that the naive estimate for the saturation time scale becomes larger than typical dynamical times means that A_{int} does not have a monotonic behavior in time. In fact, it turns out that at very low temperatures, the role of the anomalous diffusion term in the master equation starts to be relevant (its value is of the same order of magnitude as the normal diffusion coefficient). The contribution of this term to the evolution of $A_{\rm int}$ is clearly seen in (88) where we see that the second term (associated with anomalous diffusion) does not have a well-defined sign (its sign changes as the interference fringes rotate in phase space). From this observation, one expects that if at low temperatures the fringe visibility factor does not saturate, its time dependence should exhibit some oscillatory behavior (modulating an overall increase dictated by normal diffusion). The periods of slower decoherence coincide with the moments when fringes get oriented along the position axis (this coincides with the instant when the two wavepackets are most separated in momentum). This qualitative prediction concerning the behavior of $A_{\rm int}$ is confirmed by the exact numerical calculations shown in Fig. 5. In this figure, the oscillations are clearly seen.

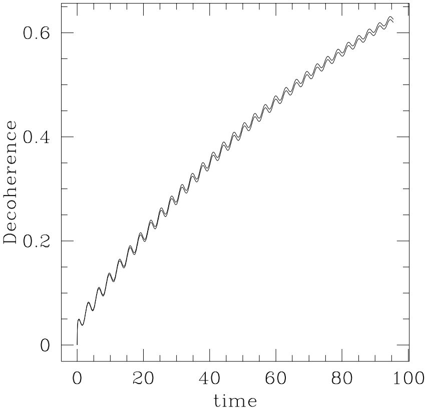


Fig. 5. Decoherence at zero temperature proceeds nonmonotonically. Here, the time dependence of $A_{\rm int}$ for a harmonic oscillator interacting with a zero temperature environment is displayed. Oscillations correspond to the change in orientation of the interference fringes.

A very simple and interesting expression for $A_{\rm int}$ can be obtained for the QBM model. Thus, in [45] it has been shown that the fringe visibility factor can always be written as follows:

$$A_{\rm int} = \frac{1}{2}\left(\frac{2L_0}{\Delta x}\right)^2 \coth^2(\beta\Omega/2)\left(1-\left(\ddot{S}^2/\Omega^4+\dot{S}^2/\Omega^2\right)\right) \ , \tag{98}$$

where Δx^2 is the position dispersion in thermal equilibrium (i.e., $\Delta x^2 = \hbar\coth(\beta\Omega/2)/M\Omega$) and S is the normalized position autocorrelation function defined as

$$\Delta x^2 S(t) = \frac{1}{2}\langle\{x(t),x\}\rangle - \langle x(t)\rangle\langle x\rangle \ . \tag{99}$$

This equation enables us to obtain very simple qualitative estimates of the efficiency of decoherence. More interestingly, it clearly shows that decoherence has the same physical origin as other dissipative effects (and is closely related to the decay of the autocorrelation function through (98)). However, in spite of their common origin, the decay of correlations and the decoherence process have very different time scales. In fact, from the above equation we can estimate how much the correlation functions have to decay in order for the system to decohere. Thus, at the time for which A_{int} approaches unity, the spatial correlations in the system should have decayed by a factor $S(t_{\mathrm{dec}})/S(0) = \sqrt{1 - \Delta x^2/4L_0^2}$, which is indeed very small (note that Δx approaches the thermal de Broglie wavelength at high temperatures and the spread of the ground state at zero temperature).

On the other hand, the above formula (98) can also allow us to estimate correctly A_{int} both at high and low temperatures for the underdamped Brownian motion model. In fact, we just need to obtain a reasonable approximation for the position correlation function. For example, assuming a simple exponential decay would lead us to conclude that

$$A_{\mathrm{int}} = \frac{1}{2}\frac{4L_0^2}{\Delta x^2}\coth^2(\beta\Omega/2)(1 - \exp(-\gamma_0 t)). \tag{100}$$

This is a crude but very reasonable approximation that is, for example, not only very good at high temperatures and very early times but also exhibits the correct saturation behavior for long times. It can be further improved by better approximating the position correlation function. For example, computing $S(t)$ in the highly underdamped regime we obtain

$$\begin{aligned} A_{\mathrm{int}} &= \frac{1}{2}\frac{4L_0^2}{\Delta x^2}\coth^2(\beta\Omega/2) \\ &\quad\times \left(1 - \exp(-\gamma_0 t)\left(1 + \gamma_0^2\sin^2(\Omega t)/2\Omega_0^2 - \gamma_0\sin(2\Omega_0 t)/2\Omega_0\right)\right), \end{aligned} \tag{101}$$

which is a very good approximation for the low–temperature (low–damping) behavior exhibited in Fig. 5.

5.3 Preexisting Correlations Between the System and the Environment

Almost all papers concerning decoherence assume that the initial state has no correlations between the system and the environment (i.e., that the state can be factored). In this section we will analyze what happens if we consider more general initial conditions. In particular, we are interested in analyzing initial conditions that are closer to what we encounter experimentally. Thus, we consider a situation in which the system and the environment are initially in a thermal equilibrium state at some temperature (which could be zero) and at the initial time we perform a measurement on the system to prepare an initial state. This measurement could be imperfect (i.e., may be characterized not by a projection operator, but by a POVM). After this measurement, we consider the evolution of the system coupled to the environment in the usual way. Under these circumstances, the initial state of the universe is generally not a product. Moreover, in the case when the initial state is a product (i.e., when the measurement performed on the system is perfect), the state of the environment depends functionally on the state of the system. This type of initial states can generally be written as

$$\rho_0 = \sum_j A_j \rho_\beta A'_j \ , \tag{102}$$

where A_j and A'_j are Krauss operators (not necessarily projectors) acting on the Hilbert space of the system (see [29] for a good review).

We are not going to present any details of the calculations leading to the (exact) solution of this model. Our presentation follows closely the one in [45] where the influence of initial correlations on decoherence was examined. Here we present a summary of the results obtained in that paper.

First, it is worth stressing the fact that it is still possible to find a relatively simple master equation for the reduced density matrix of the system. However, the existence of initial correlations prevents us from expressing this equation entirely in terms of the reduced density matrix. Thus, the evolution of ρ not only depends on ρ itself but also on initial correlations between the system and the environment. Interestingly enough, for the case of the linear QBM model, an exact master equation that is very similar to (49) can be obtained. It reads as follows:

$$\begin{aligned}\dot{\rho}(q,q',t) = {} & \mathrm{i}\left(\frac{1}{2}\left(\partial_q^2 - \partial_{q'}^2\right) - \frac{1}{2}\Omega^2(t)(q^2 - {q'}^2)\right)\rho(q,q',t) \\ & - \gamma(t)\ (q-q')\ (\partial_q - \partial_{q'})\,\rho(q,q',t) \\ & - D_1(t)\ (q-q')^2\ \rho(q,q',t) \\ & - \mathrm{i}D_2(t)\ (q-q')\ (\partial_q + \partial_{q'})\,\rho(q,q',t) \\ & + \mathrm{i}\tilde{C}_1(t)\ (q-q')\ \rho_{11}(q,q',t) \\ & - \mathrm{i}\tilde{C}_2(t)\ (q-q')\ \rho_{12}(q,q',t)\ .\end{aligned} \tag{103}$$

It is important to stress that this equation is exact and valid for all spectral densities and initial temperatures. The time-dependent coefficients appearing in (103) are functions of time and temperature (and of the spectral density of the environment, of course). Explicit formulae are given in [45]. The interpretation of the first three lines of this equation is identical to the ordinary case where no correlations are present. The initial correlations appear in the time dependence of the coefficients but, for realistic environments, this dependence is very weak (thus, these coefficients are qualitatively the same as before). The last two lines make this equation nonhomogeneous. In fact, these terms are present because of the correlated nature of the initial state. Thus, in that case, the master equation cannot be entirely written in terms of the reduced density matrix. It can be shown that the two density matrices ρ_{11} and ρ_{12} are obtained by propagating two different initial states given by the "density matrices" $\rho_{11} = \{q, \rho\}$ and $\rho_{12} = \mathrm{i}[q, \rho]$. The evolution of ρ_{1i} can also be studied with this formalism because (apart from not being normalized) they belong to the class of initial conditions defined by (102). Therefore, the evolution equation obeyed by these operators is also (103), with new inhomogeneous terms. Thus, a hierarchy of equations, which are coupled because of the initial correlations, can be derived in this way (see [45] for more details).

The time dependence of all the coefficients has been studied in detail in [45] and the conclusion is that, for an ohmic environment at arbitrary temperatures, the coefficients $\tilde{C}_1$ and $\tilde{C}_2$, entering in the inhomogeneous terms of the master equation are exceedingly small and become negligible after a time that is on the order of the cutoff time scale. After this short initial transient, the impact of the initial correlations on the future evolution of the system can be entirely neglected. Of course, in less realistic situations, it is possible to show that these coefficients have an important effect. For example, the formalism we described could be applied to the case of two coupled oscillators in which one considers one of them as the system and the other one as the environment. In this case, when the size of the system and environment are comparable, initial correlations play an important role. The time dependence of the other coefficients of the master equation is also affected by the correlations but they all behave qualitatively in a similar way as in the absence of such initial correlations (see [45] for a detailed study of these coefficients).

It is interesting to analyze the evolution of a delocalized initial state to see how decoherence takes place in this model, which includes the effect of initial correlations. For this, we consider the initial condition (102 with the operators associated to a projection onto a Schrödinger cat state (say, a superposition of two coherent states separated in position). Thus, we take

$$\rho = \frac{\hat{P}\rho_\beta\hat{P}}{\mathrm{Tr}(\rho_\beta\hat{P})} \ ,$$

where $\hat{P}$ is a projector onto a pure state of the system $\hat{P} = |\Psi\rangle\langle\Psi|$ and the state $|\Psi\rangle$ is itself a Schrödinger's cat state (i.e., a superposition of two Gaussian packets),

$$|\Psi\rangle = |\Psi_+\rangle + |\Psi_-\rangle \ , \tag{105}$$

where $|\Psi_\pm\rangle$ are such that

$$\langle x|\Psi_\pm\rangle \ = N \ \exp\left[-\frac{(q \mp L_0)^2}{2\delta^2} \pm \mathrm{i}P_0 q\right] \ . \tag{106}$$

The decoherence process for this initial state has been analyzed in the previous section in the absence of initial correlations. The fate of this state is not very different from the behavior we described before but there are some subtle differences. Thus, initial correlations distort the Gaussian peaks in the initial Wigner function as well as the intermediate interference fringes. An exact solution of the problem is possible (see [45]) and it turns out that it is no longer true that the Wigner function can be written as the sum of two Gaussian peaks plus interference fringes. In fact, it turns out that each Gaussian peak is distorted in such a way that it can be written as the sum of two nearby Gaussians with a term between them. The same is true for the interference fringes, which get distorted and split into several (actually ten) terms. However, for realistic (ohmic) environments, this effect is very small (as discussed in [45]), and the decoherence process goes qualitatively in the same way as described in the previous section (in fact, in Fig. 5, the two curves for the decoherence factor are almost indistinguishable from each other: one corresponds to an initially uncorrelated state while the other to the case described in this section).

6 Decoherence and Chaos

Here we investigate environment induced superselection in the context of quantum chaos (i.e., quantum dynamics of systems that are classically chaotic). We first argue [60] that the evolution of a chaotic macroscopic (but, ultimately, quantum) system is not just difficult to predict (requiring accuracy exponentially increasing with time) but quickly ceases to be deterministic in principle as a result of the Heisenberg indeterminacy (which limits the resolution available in the initial conditions). This happens after a time $t_\hbar$, which is only logarithmic in the Planck constant. A definitely macroscopic (if somewhat outrageous) example [61] is afforded by various components of the solar system that are chaotic, with the Lyapunov time scales ranging from a bit more then a month (Hyperion, a prolate moon of Saturn[57]) to millions of years (planetary system as a whole [58,59]). On the time scale $t_\hbar$ the initial minimum uncertainty wave packets corresponding to celestial bodies would be smeared over distances of the order of the radii of their orbits into “Schrödinger cat like” states, and the concept of a trajectory would cease to apply. In reality, such paradoxical states are eliminated by decoherence that helps restore quantum-classical correspondence. We shall also see that the price for the recovery of classicality is the loss of predictability. In the classical limit (associated with effective decoherence, and not just with the smallness of $\hbar$) the rate of increase of the von Neumann entropy of the decohering system is independent of the strength of the coupling to the environment and equal to the sum of the positive Lyapunov exponents.

6.1 Quantum Predictability Horizon: How the Correspondence Is Lost

As a result of chaotic evolution, a patch in the phase space that corresponds to some regular (and classically “reasonable”) initial condition becomes drastically deformed. Classical chaotic dynamics is characterized by the exponential divergence of trajectories. Moreover, conservation of the volume in the phase space in the course of Hamiltonian evolution (which is initially a good approximation for sufficiently regular initial conditions, even in cases that are ultimately quantum) implies that the exponential divergence in some of the directions must be balanced by the exponential squeezing – convergence of trajectories – in other directions. It is that squeezing that forces a chaotic system to explore the quantum regime. As the wave packet becomes narrow in the direction corresponding to momentum,

$$\Delta p(t) \;=\; \Delta p_0 \;\exp(-\lambda t) \;, \tag{107}$$

(where Δp_0 is its initial extent in momentum, and λ is the relevant Lyapunov exponent) the position becomes delocalized: The wave packet becomes coherent over the distance $\ell(t)$ that can be inferred from Heisenberg's principle,

$$\ell(t) \geq (\hbar/\Delta p_0)\exp(\lambda t) \ . \tag{108}$$

Coherent spreading of the wavepacket over large domains of space is disturbing in its own right. Moreover, it may lead to a breakdown of the correspondence principle at an even more serious level. Predictions of the classical and quantum dynamics concerning some of the expectation values no longer coincide after a time $t_\hbar$ when the wave packet coherence length $\ell(t)$ reaches the scale on which the potential is nonlinear.

Such a scale χ can usually be defined by comparing the classical force (given by the gradient of the potential $\partial_x V$) with the leading order nonlinear contribution $\sim \partial_x^3 V$,

$$\chi \ \simeq \ \sqrt{\frac{\partial_x V}{\partial_x^3 V}} \ . \tag{109}$$

For instance, for the gravitational potential $\chi \simeq R/\sqrt{2}$, where R is a size of the system (i.e., a size of the orbit of the planet). The reason for the breakdown of the correspondence is that when the coherence length of the wave packet reaches the scale of nonlinearity,

$$\ell(t) \ \simeq \chi \ , \tag{110}$$

the effect of the potential energy on the motion can no longer be represented by the classical expression for the force [60], $F(x) = \partial_x V(x)$, because it is not even clear where the gradient is to be evaluated for a delocalized wave packet. As a consequence, after a time given by

$$t_\hbar \ = \ \lambda^{-1} \ln \frac{\Delta p_0 \chi}{\hbar} \ , \tag{111}$$

the expectation value of some of the observables of the system may even begin to exhibit noticeable deviations from the classical evolution [64].

This is also close to the time beyond which the combination of classical chaos and Heisenberg's indeterminacy makes it impossible *in principle* to employ the concept of a trajectory. Over the time $\sim t_\hbar$ a chaotic system will spread from a regular Planck-sized volume in the phase space into a (possibly quite complicated) wave packet with the dimensions of its envelope comparable to the range of the system. This time scale defines the quantum predictability horizon – a time beyond which the combination of classical chaos and quantum indeterminacy makes predictions not just exponentially difficult, but impossible in principle. The shift of the origin of the loss of predictability from classical deterministic chaos to quantum indeterminacy amplified by exponential instabilities is just one of the symptoms of the inability of classical evolution to track the underlying quantum dynamics.

This breakdown of correspondence can be investigated more rigorously by following the evolution of the Wigner function (defined in (77)) for the possibly macroscopic, yet ultimately quantum system. Dynamics of the Wigner function is generated by the *Moyal bracket* (that is simply the Wigner transform of the right–hand side of von Neumann equation for the density matrix). This Moyal bracket can be expressed through the familiar classical Poisson bracket:

$$\dot{W} \ = \{H, W\}_{\mathrm{MB}} \ = \ -\mathrm{i}\sin(\mathrm{i}\hbar\{H, W\}_{\mathrm{PB}})/\hbar \ . \tag{112}$$

Above, H is the Hamiltonian of the system, and W is the Wigner transform of the density matrix.

When the potential V in H is analytic, the Moyal bracket can be expanded in powers of the Planck constant. Consequently, the evolution of the Wigner function is given by

$$\dot{W} \ = \ \{H, W\}_{\mathrm{PB}} \ + \sum_{n\geq 1} \frac{\hbar(-)^n}{2^{2n}(2n+1)!} \partial_x^{2n+1} V(x) \partial_p^{2n+1} W(x,p). \tag{113}$$

Correction terms above will be negligible when $W(x,p)$ is a reasonably smooth function of p, that is, when the higher derivatives of W with respect to momentum are small. However, the Poisson bracket alone predicts that, in the chaotic system, they will increase exponentially quickly as a result of the "squeezing" of W in momentum, (107). Hence, after $t_\hbar$, quantum "corrections" will become comparable to the first classical term on the right–hand side of (113). At that point, the Poisson bracket will no longer suffice as an approximate generator of evolution. The phase space distribution will be coherently extended over macroscopic distances, and interference between the fragments of W will play a crucial role.

The time scale on which the quantum-classical correspondence is lost in a chaotic system can also be estimated (or rather, bounded from above) by the formula [62,63]

$$t_r \ = \ \lambda^{-1} \ln(I/\hbar) \ , \tag{114}$$

where I is the action.

6.2 Exponential Instability vs. Decoherence

In a quantum chaotic system weakly coupled to the environment, the process of decoherence briefly sketched above will compete with the tendency for coherent delocalization, which occurs on the characteristic time scale given by the Lyapunov exponent λ. Exponential instability would spread the wave packet to the "paradoxical" size, but monitoring by the environment will attempt to limit its coherent extent by smoothing out interference fringes. The two processes shall reach *status quo* when their rates are comparable,

$$\tau_D(\delta x) \ \lambda \ \simeq \ 1. \tag{115}$$

Because the decoherence rate depends on δx, this equation can be solved for the critical, steady state coherence length, which yields $\ell_{\mathrm{c}} \sim \Lambda_{\mathrm{dB}}(T) \times \sqrt{\lambda/\gamma}$.

A more careful analysis can be based on the combination of the Moyal bracket and the master equation approach to decoherence we have just sketched. In many cases, (including the situation of large bodies immersed in the typical environment of photons, rarefied gases, etc.) an effective approximate equation can be derived and translated into the phase space by performing a Wigner transform of the master equation. Then:

$$\begin{aligned}\dot{W} &= \{H,W\}_{\mathrm{PB}} + \sum_{n\geq 1} \frac{\hbar^{2n}(-1)^n}{2^{2n}(2n+1)!}\partial_x^{2n+1}V(x)\partial_p^{2n+1}W(x,p) \\ &\quad + 2\gamma\partial_p pW + D\partial_p^2 W \ . \end{aligned} \qquad (116)$$

As before, we are interested in the regime where we can neglet the term that causes relaxation, which, in the macroscopic limit, can be made very small without decreasing the effect of decoherence caused by the last, diffusive term. As we saw in the previous section, the role of this decoherence term is to destroy the quantum coherence of the fragments of the wave function between spatially separated regions. Thus, in effect, this *decoherence term* can esure that the Poisson bracket is always reasonably accurate. Diffusion prevents the wave packet from becoming too finely structured in momentum, which would have caused the failure of the correspondence principle. In the case of the thermal environment, the diffusion coefficient $D = \eta k_{\mathrm{B}}T$, where η is the viscosity. The competition between the squeezing resulting from the chaotic instability and spreading resulting from diffusion leads to a standoff when the Wigner function becomes coherently spread over

$$\ell_{\mathrm{c}} = \hbar\sqrt{\frac{\lambda}{2D}} = \Lambda_{\mathrm{dB}}(T)\ \sqrt{\lambda/2\gamma}\ . \qquad (117)$$

This translates into the critical (spatial) momentum scale of

$$\sigma_{\mathrm{c}} = \sqrt{\frac{2D}{\lambda}}\ , \qquad (118)$$

which nearly coincides with the quick estimate given by equation (115).

Returning to an outrageous example of the solar system, for a planet of the size of Jupiter a chaotic instability on the four–million–year time scale and the consequent delocalization would be easily halted even by a very rarefied medium (0.1 atoms/cm^3, comparable to the density of interplanetary gas in the vicinity of massive outer planets) at a temperature of 100 K (comparable to the surface temperature of major planets): The resulting ℓ_{c} is on the order of 10^{-29} cm! Thus, decoherence is exceedingly effective in preventing the packet from spreading; $\ell_{\mathrm{c}} \ll \chi$, by an enormous margin. Hence, the paradox we have described in the first part of the paper has no chance of materializing.

The example of quantum chaos in the solar system is a dramatic illustration of the effectiveness of decoherence, but its consequences are, obviously, not restricted to celestial bodies: Schrödinger cats, Wigners friends, and, generally, all of the systems that are in principle quantum but sufficiently macroscopic will be forced to behave in accordance with classical mechanics as a result of the environment–induced superselection [1,2]. This will be the case whenever

$$\ell_{\mathrm{c}} \ll \chi \ , \tag{119}$$

because ℓ_{c} is a measure of the resolution of "measurements" carried out by the environment.

This incredible efficiency of the environment in monitoring (and, therefore, localizing) states of quantum objects is actually not all that surprising. We know (through direct experience) that photons are capable of maintaining an excellent record of the location of Jupiter (or any other macroscopic body). This must be the case, because we obtain our visual information about the universe by intercepting a minute fraction of the reflected (or emitted) radiation with our eyes.

Our discussion extends and complements developments that go back more than a decade [65]. We have established a simple criterion for the recovery of the correspondence, (119), which is generously met in the macroscopic examples discussed above. And, above all, we have demonstrated that the *very same* process of decoherence that delivers "pointer basis" in the measuring apparatus can guard against violation of the quantum-classical correspondence in dynamics.

6.3 The Arrow of Time: A Price of Classicality?

Decoherence is caused by the continuous measurement-like interactions between the system and the environment. Measurements involve the transfer of information, and decoherence is no exception: The state of the environment acquires information about the system. For an observer who has measured the state of the system at some initial instant the information he will still have at some later time will be influenced (and, in general, diminished) by the subsequent interaction between the system and the environment. When the observer and the environment monitor the same set of observables, information losses will be minimized. This is in fact the idea behind the *predictability sieve*[3,4] – an information-based tool which allows one to look for the einselected, effectively classical states under quite general circumstances. When, however, the state implied by the information acquired by the observer either differs right away from the preferred basis selected by the environment, or – as will be the case here – evolves dynamically into such a "discordant" state, the environment will proceed to measure it in the preferred basis, and, from the observer's point of view, information loss will ensue.

This information loss can be analyzed in several ways. The simplest is to compute the (von Neumann) entropy increase in the system. This will be our objective in this section. However, it is enlightening to complement this "external" view by looking at the consequences of decoherence from the point of view of the observer, who is repeatedly monitoring the system and updating his records [66]. The loss of information can be quantified by the increase of the von Neumann entropy,

$$\mathcal{H} \;=\; -\mathrm{Tr}\,\rho \ln \rho \;, \tag{120}$$

where ρ is the reduced density matrix of the system. We shall now focus on the rate of increase of the von Neumann entropy in a dynamically evolving system subject to decoherence. As we have seen before, decoherence restricts the spatial extent of the quantum-coherent patches to the critical coherence length ℓ_c, (117). A coherent wavepacket that overlaps a region larger than ℓ_c will decohere rapidly, on a time-scale τ_D shorter than the one associated with the classical predictability loss rate given by the Lyapunov exponent λ. Such a wave packet will deteriorate into a mixture of states, each of which is coherent over a scale of dimension ℓ_c by $\sigma_\mathrm{c} = \hbar/\ell_\mathrm{c}$. Consequently, the density matrix can be approximated by an incoherent sum of reasonably localized and approximately pure states. When N such states contribute more or less equally to the density matrix, the resulting entropy is $\mathcal{H} \;\simeq\; \ln N$.

The coherence length ℓ_c determines the resolution with which the environment is monitoring the position of the state of a chaotic quantum system. That is, by making an appropriate measurement on the environment, one could in principle localize the system to within ℓ_c. As time goes on, the initial phase space patch characterizing the observer's information about the state of the system will be smeared over an exponentially increasing range of the coordinate, (108). When the evolution is reversible, such stretching does not matter, at least in principle: It is matched by the squeezing of the probability density in the complementary directions (corresponding to negative Lyapunov exponents). Moreover, in the quantum case folding will result in the interference fringes – tell-tale signature of the long range quantum coherence, best visible in the structure of the Wigner functions.

Narrow wave packets, and, especially, small-scale interference fringes are exceedingly susceptible to monitoring by the environment. Thus, the situation changes dramatically as a result of decoherence. In a chaotic quantum system, the number of independent eigenstates of the density matrix will increase as

$$N \;\simeq\; \ell(t)/\ell_\mathrm{c} \;\simeq\; \frac{\hbar}{\Delta p_0 \ell_\mathrm{c}}\; \exp(\lambda t)\;. \tag{121}$$

Consequently, the von Neumann entropy will grow at the rate:

$$\dot{\mathcal{H}} \simeq \;\frac{\mathrm{d}}{\mathrm{d}t} \ln(\ell(t)/\ell_\mathrm{c}) \;\simeq\; \lambda\;. \tag{122}$$

This equation emerged as a "corollary" of our discussion, but perhaps it is even its key result: Decoherence will help restore the quantum-classical correspondence. But we have now seen that this will happen at a price. Loss of information is an inevitable consequence of the eradication of the "Schrödinger cat" states that were otherwise induced by the chaotic dynamics. They disappear because the environment is "keeping an eye" on the phase space, monitoring the location of the system with an accuracy set by ℓ_c.

Throughout this section we have "saved" on notation, using "λ" to denote (somewhat vaguely) the rate of divergence of the trajectories of the hypothetical chaotic system. It is now useful to become a bit more precise. A Hamiltonian system with $\mathcal{D}$ degrees of freedom will have in general many ($\mathcal{D}$) pairs of Lyapunov exponents with the same absolute value but with opposite signs. These global Lyapunov exponents are obtained by averaging local Lyapunov exponents, which are the eigenvalues of the Jacobian of the local transformation, and which describe the rates at which a small patch centered on a trajectory passing through a certain location in the phase space is being deformed.

The evolution of the Wigner function in the phase space is governed by the local dynamics. However, over the long haul, and in the macroscopic case, the patch that supports the probability density of the system will be exponentially stretched. This stretching and folding will produce a phase-space structure that differs from the classical probability distribution because of the presence of the interference fringes, with the fine structure whose typical scale is on the order of $\hbar/\ell^{(i)}(t)$. In an isolated system, this fine structure will saturate only when the envelope of the Wigner function fills in the available phase space volume. Monitoring by the environment destroys these small–scale interference fringes and keeps W from becoming narrower than σ_{c} in momentum. As a result – and in accord with (122) above – the entropy production will asymptotically approach the rate given by the sum of the positive Lyapunov exponents,

$$\dot{\mathcal{H}} = \sum_{i=1}^{\mathcal{D}} \lambda_{+}^{(i)} \ . \tag{123}$$

This result [60] is at the same time familiar and quite surprising. It is familiar because it coincides with the Kolmogorov-Sinai formula for the entropy production rate for a *classical* chaotic system. Here we have seen underpinnings of its more fundamental quantum counterpart. All the same, it is surprising because it is independent of the strength of the coupling between the system and the environment, even though the process of decoherence (caused by the coupling to the environment) is the ultimate source of entropy increase. Over the last few years, the argument we presented above has been investigated and confirmed, using numerical simulations (see [67,68,69,70,71]). Figure 6 presents clear evidence showing that in the chaotic regime the entropy production rate approaches the value set by the Lyapunov exponent (data correspond to studies of a quantum particle moving in a harmonically driven double well potential [71]).

This independence is indeed remarkable, and leads one to suspect that the cause of the arrow of time may be traced to the same phenomena that are responsible for the emergence of classicality in chaotic dynamics, and elsewhere (i.e., in quantum measurements). In a sense, this is of course not a complete surprise: Von Neumann knew that the measurements are irreversible [18]. And Zeh [32] emphasized the close kinship between the irreversibility of the "collapse" in quantum measurements and in the second law, cautioning against circularity of using one to solve the other. However, what is surprising is that both the classical-looking result ultimately has quantum roots, and that these roots are

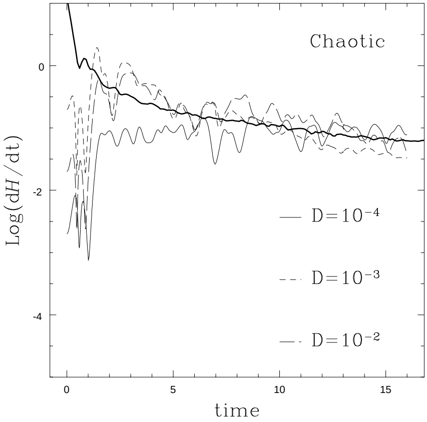


Fig. 6. Entropy production resulting from decoherence for a classically chaotic system becomes, after an initial transient, independent of the value of the diffusion constant and set by the Lyapunov exponent. See [71].

so well hidden from view that the entropy production rate depends solely on the classical Lyapunov exponents.

Environment may not enter explicitly into the entropy production rate, (123), but it will help determine when this asymptotic formula becomes valid. The Lyapunov exponents will "kick in" as the dimensions of the patch begin to exceed the critical sizes in the corresponding directions, $\ell^{(i)}(t)/\ell_{\mathrm{c}}^{(i)} > 1$. The instant when that happens will be set by the strength of the interaction with the environment, which determines ℓ_{c}. This "border territory" may be ultimately the best place to test the transition from quantum to classical. One may, for example, imagine a situation where the above inequality is comfortably satisfied in some directions in the phase space, but not in the others. In that case, the rate of the entropy production will be lowered to include only these Lyapunov exponents for which decoherence is effective.

6.4 Decoherence, Einselection, and the Entropy Production

The significance of the efficiency of decoherence goes beyond the example of the solar system or the task of reconciling quantum and classical predictions for classically chaotic systems. Every degree of freedom coupled to the environment will suffer loss of quantum coherence. Objects that are more macroscopic are generally more susceptible. In particular, the "hardware" responsible for our perceptions of the external universe and for keeping records of the information acquired in the course of our observations is obviously very susceptible to decoherence. Neurons are strongly coupled to the environment and are definitely macroscopic enough to behave in an effectively classical fashion. That is, they have a decoherence time scale many orders of magnitude smaller than the relatively sluggish time scale on which they can exchange and process information.

As a result, in spite of the undeniably quantum nature of the fundamental physics involved, perception and memory have to rely on the information stored in the decohered (and, therefore, effectively classical) degrees of freedom.

An excellent illustration of the constraint imposed on information processing by decoherence comes from the recent discussions of the possibility of implementing real quantum computers. Decoherence is viewed as perhaps the most serious threat to the ability of a quantum information processing system to carry out a superposition of computations [30,11]. Yet, precisely such an ability to "compute" in an arbitrary superposition would be necessary for an observer to be able to "perceive" an arbitrary quantum state. Moreover, in the external universe only those observables that are resistant to decoherence and which correspond to "pointer states" are worth recording. Records are valuable because they allow for predictions, and resistance to decoherence is a precondition to predictability [14,3].

It is too early to claim that all the issues arising in the context of the transition from quantum to classical have been settled with the help of decoherence. Decoherence and einselection are, however, rapidly becoming a part of a standard lore [72,73]. Where expected, they deliver classical states, and – as we have seen above – guard against violations of the correspondence principle. The answers that emerge may not be to everyone's liking, and do not really discriminate between the Copenhagen Interpretation and the Many Worlds approach. Rather, they fit within either mold, effectively providing the missing elements – delineating the quantum-classical border postulated by Bohr (decoherence time fast or slow compared to the dynamical time scales on the two sides of the "border"), and supplying the scheme for defining distinct branches required by Everett (overlap of the branches is eliminated by decoherence).

7 How to Fight Against Decoherence: Quantum Error Correcting Codes

It is clear that decoherence is a process that has a crucial role in the quantum–to–classical transition. But in many cases, physicists are interested in understanding the specific causes of decoherence just because we want to get rid of it. Thus, decoherence is responsible for washing out the quantum interference effects we would very much like to see as a signal in some experiments. This is the type of situation one is clearly facing in quantum computation (and in the physics of quantum information in general). A quantum computer is a gigantic interferometer whose wave function explores an exponential number of classical computations simultaneously. Coherence between branches of the computer wave function should be maintained because the existence of quantum interference between these branches is the basic reason why these computers can outperform their classical counterparts. Thus, decoherence in this context is a major problem.

An obvious way of try to prevent decoherence from damaging quantum states is to reduce the strength of the coupling between the system and its environment. However, it is never possible to reduce this coupling to zero and eliminate decoherence in this way. Remarkably, in recent years new techniques that enable the active protection of the information stored in quantum states from the degrading effect of the interaction with the environment have emerged. They come under the name of "Quantum Error Correcting Codes" (QECC) and were invented by people working on quantum computation [75,74]. They are based on remarkably simple and beautiful ideas and could be found to be useful in other areas of physics. For this reason, we believe it could be interesting to include this final main theme to give a simple–minded presentation of the methods that could enable us in principle to "fight against decoherence" preserving quantum states.

7.1 How to Protect a Classical Bit

To introduce the basic idea of Quantum Error Correcting Codes it is better to start with a short discussion of the simplest ways in which one can protect classical information. Suppose that we have a single qubit b that lives in a noisy environment. Because of the effect of the noise we will assume that the bit has a probability p to flip after some time. Therefore, if we look at the bit after this time, the probability of the bit being unaltered by the noise is $1-p$ and therefore, the information is degraded. Can we protect this classical bit? The answer is "yes" and the way to do it is by using an error correcting code. The simplest such procedure is based on the brute force use of redundancy as follows. We can "encode" this one bit of information using more carriers, mapping the state of the bit into many identical copies (i.i. $b \to (b, b, \ldots, b)$). If we do this, we can recover the initial information after the noise occurred by voting on and adopting as our result the one that gets the majority of votes. In this way we also discover which carriers were altered by the noise (i.e., the minority) and recuperate the information. Of course, this works if the error probability is small enough. To be precise, let us assume that we encode the information in three carrier bits (this is the simplest repetition code). The probability that no flip occurs is $P(\text{no flip}) = (1-p)^3$, and the others are simply $P(\text{one flip}) = 3p(1-p)^2$, $P(\text{two flips}) = 3p^2(1-p)$, and $P(\text{three flips}) = p^3$. Thus, the above error–correcting strategy (encoding one into three bits and voting at the end) increases the probability of keeping the information intact from $1-p$ to $1-3p^2+2p^3 = 1-O(p^2)$, which is close to unity, provided p is small enough. This example illustrates the simplest classical error correction code. Of course, much more sophisticated codes exist, and we are probably not doing justice to the beautiful theory of classical error–correcting codes (see [76]) by using this naive code as an example. However, we think it is enough for the purpose of our discussion.

7.2 How to Protect a Quantum Bit

The basic question then becomes if it is possible to generalize this simple procedure to quantum mechanics. One may be tempted to guess that this task is impossible because a quantum version of the naive repetition code described above could never work as a consequence of the nonclonability of quantum states. Also, the fact that measurements drastically affect the state of quantum systems [12,13] is somehow suggestive of the difficulties of implementing an error–correcting quantum strategy naively translating the classical error–correcting ideas. However, these expectations were proven to be incorrect when in 1995 Peter Shor created the first quantum error–correcting code [75]. His work, once again triggered a lot of activity and over the last four years the theory of Quantum Error Correcting Codes was fully developed. So far, there have been some experimental demonstrations showing the workings of these codes (only in NMR experiments) but in our view, the interesting ideas of QECC still are waiting for phycisists to give a definite answer to whether or not they will be useful for other purposes than the ones that originally motivated them. For this reason, we find it interesting to bring these issues to this review.

Let us now describe how it is possible to create QECC. For this, we consider a quantum bit prepared in an arbitrary quantum state $\Psi = \alpha|0\rangle + \beta|1\rangle$. To be precise, we will first describe how noise affects the state of the qubit whose state we want to protect. We will first consider the simplest case of a noise that just produces "dephasing". We assume that the noise introduces a random phase with a probability p or leaves the state intact with a probability $1-p$. Although this is not the most general kind of operation that a noisy environment can produce on a quantum system, we will later show that this is not a restrictive assumption and that the treatment we present here can be generalized to include all of the effects that the noise can produce. So, for the moment we will just consider this "dephasing" noise. The dephasing can be simply described by the action of a σ_z operator on the state of the system. In this chapter, we will adopt the following notation. The Pauli matrices $\sigma_{x,y,z}$ are simply denoted as X, Y, Z. Thus, if the initial state of the system is $\Psi_0 = \alpha|0\rangle + \beta|1\rangle$ the final state (after the noise has occurred) is described by a density matrix as,

$$\rho_{\text{out}} = (1-p)\rho_{\text{in}} + pZ\rho_{\text{in}}Z \ , \tag{124}$$

where $\rho_{\text{in}} = |\Psi_0\rangle\langle\Psi_0|$. It is easy to see that the interaction with the noise degrades the quantum state, causing the loss of quantum coherence. As a measure of this degradation, we can compute the "fidelity" of the process that is simply given by the overlap between the ideal state and the actual state. Using the above form for the density matrix, we find out that fidelity is reduced to $F = \text{Tr}(\rho_{\text{out}}\rho_{\text{in}}) = 1 - 4p|\alpha\beta|^2$. Thus, fidelity is reduced by an amount that is linear in the error probability p. Another measure of the degradation is given by the loss of purity of the final state that can be measured, for example, by $\text{Tr}(\rho^2_{\text{out}}) = 1 - 8p(1-p)|\alpha\beta|^2$. In what follows, we will present a method that enables us to protect the quantum state in such a way that the fidelity (or the loss of purity) does not decay linearly with the error probability but it does so quadratically.

So, let us present a way to protect the state of our qubit from the effect of a dephasing environment. As in the classical case, we will use many carriers to protect one qubit of information (in our example, we use three qubits to protect one). But the use of redundancy has to be more subtle in the the quantum case.

The key idea is to encode the logical states into entangled states of the three qubits in such a way that when an error occurs, the logical states are mapped into other orthogonal subspaces (one subspace for each error we want to correct). If this is the case, we can learn about the error by measuring an observable that just tells us in what two–dimensional subspace the state is in. In this way, we learn what the error was without getting any information about the state itself. Once we know the error, we can correct it and start the process all over again. This idea is illustrated clearly (we hope!) by the three-qubit example. In this case, we can use the following encodinf for the logical states:

$$\begin{aligned}|0\rangle_L &= \frac{1}{2}(|000\rangle + |110\rangle + |101\rangle + |011\rangle),\\ |1\rangle_L &= \frac{1}{2}(|111\rangle + |001\rangle + |010\rangle + |100\rangle),\end{aligned} \qquad (125)$$

(the subscript L is used to denote the logical states). The "encoding" process is simply the mapping of the physical states of the three independent carriers onto the above entangled logical states. This task is the first one that one has to do to protect the information and is represented by a unitary operator (the encoding operator E). One takes the qubit whose quantum state is to be protected and applies an operation to it together with the other two carriers we use. This operation maps the initial state into the encoded state, i.e., $E(\alpha|0\rangle + \beta|1\rangle)|00\rangle = \alpha|0\rangle_L + \beta|1\rangle_L$. Later in this section we will describe ways in which the encoding operation can be implemented.

The reason why (125) is a good encoding can be seen as follows. It is a simple exercise to show that when we apply an error operator to any of the two logical states (i.e., when we act with a Z operator on any one of the qubits) we obtain mutually orthogonal states. Thus, one can show that $|0\rangle_L \perp Z_i|0\rangle_L \perp |1\rangle_L \perp Z_i|1\rangle_L$ for $i = 1, 2, 3$, i.e., that the two logical states and their "erroneous descendants" are a set of eight mutually orthogonal states that constitute a basis of the complete Hilbert space of the three qubits. Therefore, the total Hilbert space can be decomposed in the direct sum of four two-dimensional subspaces. The "logical subspace" H_L, which is generated by the two vectors $\{|0\rangle_L, |1\rangle_L\}$, has three "erroneous descendents" which are simply $Z_i H_L$, and the total Hilbert space is the direct sum of H_L and $Z_i H_L$ $(i = 1, 2, 3)$. As a consequence, there is an observable that we could measure to determine in which one of the four subspaces the state is in. In so doing, we discover the error and can correct it trivially.

To complete our description, we just have to exhibit what is this observable whose measurement reveals the error. To do this, it is interesting to look at the symmetries of the logical states (125). It is clear that these states are eigenstates of the operators $M_1 = X_1X_2$ and $M_2 = X_2X_3$ with eigenvalue $+1$ (thus, $|0\rangle_L$ is an homogeneous superposition of all states with an even number of ones and $|1\rangle_L$ contains all states with an odd number of ones; therefore these states are invariant when we flip any two states, which is precisely what the X_iX_j operators do). Moreover, it is easy to show that M_1 and M_2 are two commuting hermitian operators whose eigenvalues are ± 1 (this follows from the fact that these operators square to the identity, i.e., $M_i^2 = 1$). Moreover, it is simple to show that all the "erroneous descendents" of the logical subspace are also eigenspaces of M_i. For example, the subspace $Z_1 H_L$ is formed by linear superpositions of the vectors $\{z_1|0\rangle_L, |1\rangle_L\}$ that are eigenstates of M_1 and M_2 with eigenvalues equal to -1. This follows from the fact that as the error operator Z_1 anticommutes with M_1 and M_2, it transforms eigenstates of these operators into eigenstates with a different eigenvalue (i.e., if $M_i|\phi\rangle = |\phi\rangle$, then $M_i Z_1|\phi\rangle = -Z_1|\phi\rangle$). Therefore, if our goal is to find out in which of the four two-dimensional subspaces the state is in, we just have to measure the two operators M_1 and M_2. The result of this measurement is always represented by a set of two numbers that are ± 1 (the two eigenvalues of M_i) and each of the four possible alternatives (that are known as the error syndromes) identify uniquely one of the four subspaces (H_L corresponds to the syndrome $(+1, +1)$, $Z_1 H_L$ to $(-1, -1)$, $Z_2 H_L$ to $(-1, +1)$ and $Z_3 H_L$ to $(+1, -1)$).

It is also interesting to think about what kind of physical procedure we should follow to perform this kind of measurement. As discussed, we need to measure the operators M_i that are constructed as tensor products of Pauli matrices. However, it is very important to realize that we must do this **without** measuring individually the factors appearing in these products! Thus, in our case, we need to measure only $M_1 = X_1 X_2$ and $M_2 = X_1 X_3$, but we cannot do this by measuring the three operators X_i individually. If we were to do this, we would be measuring a complete set of commuting observables and causing the system to collapse into a particular state. Instead, quantum error correction needs measuring, not a complete set of observables but only enough observables to gain information about the error without destroying the coherence in the state of the system (thus, we want our measurement to project the state into a two-dimensional subspace and not to collapse it into one ray).

It is not hard to find a systematic way to devise a strategy that will enable us to measure any operator that is the tensor product of Pauli matrices without measuring the individual factors. To do this, it is clear that because the observables we measure are collective, we should induce an interaction between the qubits in such a way that after the interaction, the result of the measurement is "written" on only one particle. For example, suppose that we have two particles and we want to measure the operator $M = X_1 X_2$. Suppose also that we find a unitary operator D satisfying the condition $Z_2 D = DM$. This condition implies that the operator D will transform an eigenstate of M with eigenvalue m (that can only be ± 1) into an eigenstate of Z with eigenvalue m. Therefore, if we want to measure M, we can first apply the unitary operation D and then measure Z_2 (in other words, D is the operator that changes basis from M to Z_2 eigenstates). Thus, now we just need to construct this operator. This can be done by using a simple quantum circuit. In fact, the quantum circuit for the operator D is shown in Fig. 7. We just have to apply a Hadammard rotation to each qubit and then do a `c-not` using the first qubit as the control and the second one as the target. To show that this is the correct circuit for D, we just have to show that the relation $Z_2 D = DM$ is satisfied. For this purpose, we apply $M = X_1 X_2$ to the left of the circuit and start moving the X_i operators to the right. As these operators satisfy that $RX = ZR$, they transform into Z operators when they pass through the Hadammard rotations. Then, the Z operator in the control goes through the end of the circuit but the one acting on the target generates an extra Z in the control qubit that cancels the first one. Therefore, this implies that the circuit satisfies the required identity. Using this simple idea, if is possible to design simple quantum circuits that can be used to measure any collective observables built as tensor products of Pauli matrices. Moreover, this can be generalized to any number of qubits. For example, the circuit to measure $M_1 = X_1 X_2$ and $M_2 = X_1 X_3$ is given in Fig. 6 and consists of three Hadammard rotations (one in each qubit) followed by two `c-not` gates with the first qubit acting as the control. It is easy to see that if we measure the second and third qubits after the circuit, we learn about the syndrome and therefore find out what the error was.

To recover from the error, we just have to apply a simple operation to the remaining qubit that we do not measure (the first one in our example). This qubit contains the quantum state up to some unitary transformation that we can undo. To find out how to recover from the error, the idea is simply to see what the circuit does to the errors themselves. In fact, it is easy to show that the operator D associated with the decoding circuit appearing in Fig. 7 satisfies $Z_1 D = D X_1 X_2 X_3$, and that $Z_2 D = D X_2$, $Z_3 D = D X_3$. Therefore, this means that if the encoded state is affected by a Z_1 type error, the resulting state after decoding will have the last two qubits set to one (we already knew that this was the syndrome corresponding to this error), and the first qubit will be affected by an X rotation that we should undo. On the contrary, the other two errors (Z_2 and Z_3) do not require any corrective action.

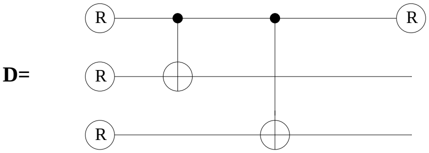


Fig. 7. Decoding circuit for the three qubit quantum error–correcting code

So, to summarize, the error–correcting procedure is the following: (1) We encode the qubit in three carriers applying the encoding circuit shown in Fig. 7. (2) After the errors act on the system, we decode the state, detect the syndromes, and apply corrective operations. (3) We refresh the syndrome qubits (resetting them to the zero state) and encoding again. It is clear that measurement of the syndrome is not really necessary because it can always be replaced by a corrective operation performed by means of a quantum circuit (in our case a `c-c-not` that is controlled by the second and third qubits). The essential part of this method is the refreshing of the syndrome qubits that is the part responsible for taking away the “entropy” generated by errors.

Two final comments are worth making before giving a more formal presentation. First, we should remark that our discussion so far assumed that errors were applied by some agent that acted on a single (unknown) qubit. However, we can extend this method to consider a situation in which there is a probability p for any one qubit to be affected by a Z–type error. In this case, the state of the three qubits before the decoding and corrective circuit is applied is given by the following density matrix:

$$\begin{aligned}\rho_{\mathrm{out}} &= (1-p)^3\rho_{\mathrm{in}} + p(1-p)^2\sum_i Z_i\rho_{\mathrm{in}}Z_i \\ &\quad + p^2(1-p)\sum_{i\neq j} Z_iZ_j\rho_{\mathrm{in}}Z_iZ_j + p^3 Z_1Z_2Z_3\rho_{\mathrm{in}}Z_1Z_2Z_3\ . \end{aligned} \tag{126}$$

After we apply the decoding and corrective procedure to this density matrix, it is clear that the first two terms will now be simply proportional to ρ_{in}. Thus, in this way we have completely eliminated the term that is linear in the error probability p. The final state differs from the ideal one only through terms that are quadratic in the error probability. Therefore, the fidelity of the whole process will be given by $F = 1 - O(p^2)$. On this linear–to–quadratic change in the dependence of F on p relies the whole power of quantum error correction (which clearly only has a good chance of working at this level, without concatenation, if p is small enough).

Finally, we could worry about not having considered more general classes of errors. However, it should be clear by now that the general idea described so far could be generalized to include more general operations. It is important to realize that to take into account all possible effects the environment could cause on a qubit, we should protect not only against phase errors (associated with Z operators) but also against bit flips (associated with X operators) and on a combination of both (associated with Y operators). It is clear that if we are able to protect against three types of independent errors, we could also fight efficiently against arbitrary unitary (or non unitary) errors that can always be written in terms of operators that are linear combinations of these three elements and of the identity. So, the question is how to invent codes that protect against arbitrary errors affecting any one of the carrier qubits. A code like this was first presented by Peter Shor [75] and can be constructed using our previous three qubit QECC as a building block. In fact, Shor encodes one qubit using nine carriers organized in three blocks of three qubits each. The logical states are a product of three factors like the ones shown in (125). This code has the following eight symmetry operations (the previous one had the two symmetries, M_1 and M_2): First, we easily find six symmetry operators that generalize the previous M_1 and M_2 in the three blocks of three qubits. Second, we find two other independent symmetries corresponding to the fact that the three blocks are repeated: $M_7 = Z_1Z_2Z_3Z_4Z_5Z_6$ and $M_8 = Z_1Z_2Z_3Z_7Z_8Z_9$. It is easy to show that each of the 27 different errors that can affect the nine carriers corresponds to a different syndrome (and therefore maps the logical states into orthogonal subspaces). The decoding should be done by measuring the above eight operators that reveal the syndrome and allow us to know the error that took place enabling us to correct it. A decoding circuit for this code can be easily constructed following the same discussion presented above for the three qubits. It is interesting to note that the code presented by Shor is by no means the most efficient way to correct errors. In fact, we notice that we are using an enormous Hilbert space of dimension $2^9 = 512$, but we would only need a space with enough room to accommodate for all the subspaces where we would map independent errors (in this case we require for this purpose only $2(1 + 3 \times 9) = 56$). Smaller codes have been developed, and the smallest one that corrects general one-qubit errors requires five qubits [77], because $n = 5$ saturates the identity $2^n = 2(3n + 1)$. This is the so-called "perfect" QECC and has the following symmetry operators $M_1 = Z_2Z_3Z_4Z_5$, $M_2 = Y_1Z_3X_4Y_5$, $M_3 = Z_1X_2Z_3X_5$, and $M_4 = Z_1Y_3Y_4Z_5$. To show that these symmetry operators constitute a good QECC requires showing that all independent errors produce a different anticommutation pattern with the M_i operator (this is left as an exercise). The construction of an encoding–decoding circuit for this code can also be done by generalizing the ideas we have described before.

7.3 Stabilizer Quantum Error-Correcting Codes

A more formal description of the principles underlying the theory of quantum error-correcting codes can be given (following the presentation of [78] we restrict ourselves to discuss a rather wide class of codes known as stabilizer codes; (for more general codes and for a more thorough discussion of QECC we refer the reader to [81,79,80]). We can consider codes that protect k qubits by encoding them into n carriers. Here, the code space $\mathcal{H}_k$ (or logical space) is a 2^k dimensional subspace of the total Hilbert space of the n carriers. $\mathcal{H}_n$ is a tensor product of n two–dimensional factors and has a natural basis whose elements are product states of the individual carriers. This is the "physical basis" that can be formed with the common eigenstates of the operators $\{Z_1, \ldots, Z_n\}$ (for convenience, we label states of this basis not by the eigenvalues of the corresponding operators, which are ±1, but by the eigenvalues of the projectors onto the −1 subspace, which are 0 or 1: thus, the label $z_j = 0$ ($z_j = 1$) corresponds to a +1 (−1) eigenvalue of the operator Z_j). Furthermore, we order the n carriers in such a way that the last k qubits are the ones whose state we encode, and the first $n-k$ are the ancillary carriers. Therefore, states of the physical basis are of the form $|s, z\rangle_P = |s\rangle_P \otimes |z\rangle_P$ (where the strings $s = (s_1, \ldots, s_{n-k})$, $z = (z_1, \ldots, z_k)$ store the corresponding eigenvalues and the subscript P is used to identify the states of the physical basis).

An error-correcting code is a mapping from the physical product states $|0\rangle_P \otimes |\Psi\rangle_P$ onto the code space $\mathcal{H}_k$, which is formed by entangled states of n carriers. A rather general class of codes can be described in terms of their stabilizer group (see [80]). The stabilizer of the code is an Abelian group formed by all operators that are tensor products of Pauli matrices and have $\mathcal{H}_k$ as an eigenspace with an eigenvalue equal to +1. Every element of the stabilizer, which is a finite group with 2^{n-k} elements, can be obtained by appropriately multiplying $n-k$ generators, which will be denoted as $M_1, \ldots, M_{n-k}$. The elements of the stabilizer are completely degenerate in the code space $\mathcal{H}_k$ (since all states in $\mathcal{H}_k$ are eigenstates with eigenvalue +1 of all M_j). To define a basis in the code space, we choose k extra operators $L_1, \ldots, L_k$, which being tensor products of Pauli matrices commute with all elements of the stabilizer. These operators $L_{j'}$, $j' = 1, \ldots, k$ are the "logical pointers" because they define the directions in $\mathcal{H}_k$ associated with the logical states $|0\rangle_L, \ldots, |2^k - 1\rangle_L$ (logical pointers belong to the group of operators that commute with the stabilizer, known as the normalizer).

The $n-k$ generators of the stabilizer together with the k logical pointers are a Complete Set of Commuting Operators (CSCO) whose common eigenstates form a complete basis of the Hilbert space $\mathcal{H}_n$. Elements of this "logical basis", labeled by n quantum numbers, are denoted as $|m, l\rangle_L$, where the bit strings $m = (m_1, \ldots, m_{n-k})$, and $l = (l_1, \ldots, l_k)$ identify the corresponding eigenvalues, and the subscript L refers to logical states. The CSCO formed by the generators of the stabilizer and the logical pointers defines a prescription for decomposing the original Hilbert space of the n carriers into a tensor product of a 2^k-dimensional logical space $\mathcal{L}$ and a 2^{n-k}-dimensional syndrome space $\mathcal{Y}$. In fact, elements of the logical basis (which are entangled states of the n-carriers) are tensor products of states belonging to $\mathcal{L}$ and $\mathcal{Y}$: $|m, l\rangle_L = |m\rangle_L \otimes |l\rangle_L$. Encoded states, which belong to $\mathcal{H}_k$, are also product states of the form $|\Psi\rangle = |0\rangle_L \otimes \sum_l c_l |l\rangle_L$.

The code protects quantum states against any error E_a whose action on states of the logical basis is to change the logical syndrome and, eventually, rotate the logical state in $\mathcal{L}$ in a syndrome–dependent way,

$$E_a \; |m\rangle_L \otimes |l\rangle_L = e^{\mathrm{i}\phi_{ma}} \; |m + c_a\rangle_L \otimes U_a |l\rangle_L \; . \tag{127}$$

Here, U_a is a unitary operator acting on the collective logical space $\mathcal{L}$, and ϕ_{ma} is a phase that may depend on the syndrome and the error. The error E_a changes the syndrome from m to $m + c_a$ where c_a is the bit string storing the commutation pattern between the error and the generators of the stabilizer (the jth bit of this string is one if the error anticommutes with M_j and is zero otherwise). The reason for this is that when acting on a logical state, the error E_a changes the eigenvalue of the operator M_j only if $\{M_j, E_a\} = 0$. The label a used to identify errors is arbitrary and, for the case of nondegenerate codes (which are the only ones we will consider here) it is always possible to label errors E_a using simply the commutation pattern c_a (i.e., we can choose $a = c_a$).

To correct against the action of any of the errors E_a (or against any linear superposition of them) one can first detect the error by measuring the collective syndrome (i.e., measuring the observables M_j, $j = 1, \ldots, n-k$) and later recover from the error by applying the corresponding operator $U_a^\dagger$. This detection–recovery process can be conveniently described as a quantum operation defined by the following mapping from the erroneous density matrix ρ_{in} into the corrected one ρ_{out},

$$\rho_{\text{out}} = \sum_{m=0}^{N} R_m \rho_{\text{in}} R_m^\dagger \ , \tag{128}$$

where the sum runs over all syndromes ($N = 2^{n-k} - 1$), and the recovery operator for each syndrome is

$$R_m = |0\rangle_L \ {}_L\langle m| \otimes U_m^\dagger \ . \tag{129}$$

By construction, these operators satisfy the identity $\sum_{m=0}^{N} R_m^\dagger R_m = I$.

Because our description of error detection–recovery process is entirely formulated on the logical basis, it does not involve so far any reference to encoding or decoding operations that can be simply defined as a change of basis. The encoding operator C is a unitary operator mapping the physical basis, formed by product states of the n carriers, onto the logical basis, formed by entangled states. Accordingly, C transforms the operators Z_i (whose eigenvalues define states on the physical basis) into the operators M_j, $L_{j'}$ (that label states on the logical basis). Thus, the encoding operator C is such that $Z_j = C^\dagger M_j C$, $j = 1, \ldots, n-k$, and $Z_{n-k+j'} = C^\dagger L_{j'} C$, $j' = 1, \ldots, k$. Taking this into account, the action of the operator R_m can be described, in the physical basis, as the following sequence of operations: i) decode the state, ii) measure the syndrome in the physical basis by measuring Z_j in the first $(n-k)$ carriers, iii) if the result of the measurement is the string s, apply the syndrome–dependent recovery operator $U_s^\dagger$ resetting the syndrome back to zero, and iv) encode the resulting state.

Finding a stabilizer code correcting a given set of errors is a rather hard task that involves designing generators having appropriate commutation patterns with the errors. Once the generators are found and the logical pointers are chosen, an encoding or decoding operator can be constructed (strategies for designing encoding or decoding circuits from the stabilizer are known; see [82,83]). The recovery operators depend on the encoding or decoding strategy and can be explicitly found from the encoding circuit by running errors through it.

As we mentioned above, the simplest code protecting $k = 1$ qubit using $n = 3$ carriers correcting against phase errors in any of the carriers can be understood as a particular example of this general stabilizer code class. In such a case, the basic errors to correct are $E_1 = Z_1$, $E_2 = Z_2$ and $E_3 = Z_3$. The stabilizer of the code can be chosen to be generated by $M_1 = X_1 X_2$ and $M_2 = X_1 X_3$. The commutation pattern associated with each error is $c_1 = 11$ (because the error Z_1 anticommutes with both M_1 and M_2), $c_2 = 10$, $c_3 = 01$ (note that we could relabel the errors ordering them according to their commutation pattern). The decoding circuits exhibited in Fig. 7 has the properties

$$C^\dagger Z_1 C = X_1 X_2 X_3, \quad C^\dagger Z_2 C = X_2, \ \text{and} \ \ C^\dagger Z_3 C = X_3. \tag{130}$$

These properties entirely determine the action of the errors Z_i in the logical basis. For example, the last identity implies that $E_3 |m\rangle_L |l\rangle_L = |m + c_3\rangle_L |l + 1\rangle_L$. Thus, the error E_3 not only changes the syndrome but also modifies the logical state by flipping it. This means that the recovery operator for this error is $U_3 = X$.

Analogously, we can find how the other errors act on the logical basis, showing that $U_1 = U_2 = I$.

8 Discussion

We have seen "decoherence in action" in a variety of settings. Our aim was not to review all of the studies of decoherence done in recent years. Thus, we left aside from our review the discussion of some very interesting physical problems where the role of environment-induced decoherence is relevant. For example, in cosmology, the way in which decoherence can account for the quantum to classical transition of density fluctuations (and of spacetime itself) has been – and still is – a matter of debate (see [84] for an incomplete list of relevant papers). Fortunately, there are also other areas where decoherence can be analyzed and tested in the laboratory. Among them, the use of systems of trapped and cold atoms (or ions) may offer the possibility of engeneering the environment (effectively choosing the pointer states) as proposed in [85]. Trapped atoms inside cavities were discussed [86] and the relation between decoherence and other cavity QED effects (such as Casimir effect) was analyzed [87]. On the mesoscopic scale, the nature of decoherence may receive increasing attention specially in the context of BEC both as a key ingredient in the phenomenological description [88] and as a threat to the longevity of BEC Schrödinger cats [89]. Moreover, the nature of decoherence is being studied experimentally in the context of condensed matter systems (see, for example, [90]).

The aim of this section is to describe briefly what is (and point out what is not) accomplished by decoherence, and to show how it facilitates understanding the transition from quantum to classical. Environment-induced superselection is clearly the key interpretational benefit arising from decoherence. The quantum principle of superposition does not apply to open quantum systems. States in the Hilbert space are no longer "equal". Under a broad variety of realistic physical assumptions, one is now forced to conclude that for macroscopic objects only a small subset of states can ever contribute to the "familiar classical reality". Only the einselected pointer states will persist for long enough to retain useful (stable) correlations with, say, the memories of the observers, or, more generally, with other stable states. By contrast, their superpositions will degrade into mixtures that are diagonal in the pointer basis.

The precondition for "perception" (as in "perception of classical reality") is the ability of the state to persist, or to evolve in a more or less predictable manner during a time interval over which the observer is monitoring it. This time interval can occasionally be quite short, but it should not be as unreasonably short as the typical decoherence time for the macroscopic systems. Thus, the only states that have a chance of being perceived as "real" are the preferred (pointer) states. Indeed, given the limited accuracy of the observer's efforts, it may be more precise to say that broad superpositions of pointer states are definitely ruled out.

It is important to emphasize that the environment-induced superselection leads to a probability distribution that is diagonal in the preferred basis, and not to a single pointer state. Thus, the uniqueness of perceptions of the observer has its roots in the stability of the correlations between the states of the macroscopic objects in the outside world and the records in the observer's memory (which, incidentally, must also use preferred states of, say, neurons to store records of the observations).

The information possessed by the observer is not an abstract, esoteric entity. Rather, "information is physical" [30] and "there is no information without representation" [3]. In practice, this means that the state of the observer is in part determined by what he knows about the rest of the universe. Thus, the physical existence of long-lasting records underlies the essence of the process of perception. Observers will be aware of their own records, and of the external universe in a state consistent with these records. This viewpoint known as the *existential interpretation* [3,17] accounts for the apparent collapse, but is consistent with either the Many Worlds or Copenhagen Interpretation.

The nature of the preferred states is dynamically negotiated in the course of the interaction between the system and the environment, but, as we have already seen, the self-Hamiltonian of the system plays an important role. Truly realistic models are difficult to treat, but lessons of the predictability sieve applied to simple models allow one to infer with some confidence that, in general, pointer states will be localized in position. After all, most interactions depend on distance. Thus, localization is an inescapable consequence [2,14]. Nevertheless, as we have already seen in perhaps the most relevant exactly solvable case of a decohering harmonic oscillator, preferred states tend to be localized in both position and momentum and can be regarded as quantum counterparts of classical points.

Investigation of the coexistence of decoherence with chaos is an example of a bit more complicated case. There, we have seen that localization is effectively enforced (even if such systems cannot be treated analytically, and extensive numerical studies are required).

An exciting "corollary" of decoherence in the setting of quantum chaos is the quantum derivation of the classically anticipated entropy production rate, given by the sum of positive Lyapunov exponents. This suggests a quantum origin of the second law of thermodynamics. Indeed, it seems that the resolution of the two outstanding puzzles of physics – the arrow of time and the apparent classicality – may originate from the same essentially quantum source, from decoherence and einselection.

The study of decoherence and einselection over the past two decades has yielded a new paradigm of emergent, effective classicality. It leads to a new understanding of the quantum origins of the classical. To be sure, not all of the interpretational questions have been settled, and much further work is required. Nevertheless, as a result of this paradigm shift, the quantum-to-classical transition has become a subject of experimental investigations, while previously it was mostly a domain of philosophy.

Acknowledgment

This research was supported in part by NSA. We gratefully acknowledge Isaac Newton Institute in Cambridge, England, where much of this manuscript was prepared. JPP was also supported by grants from ANPCyT (PICT 01014) and Ubacyt (TW23).

References

1. Zurek, W. H., *Phys. Rev.* **D 24**, 1516-1524 (1981).
2. Zurek, W. H., *Phys. Rev.* **D 26**, 1862-1880 (1982).
3. Zurek, W. H., *Progr. Theor. Phys.* **89**, 281-302 (1993).
4. Zurek, W. H., Habib, S., and Paz, J. P., *Phys. Rev. Lett* **70**, 1187-1190 (1993); Anglin, J. R., and Zurek, W. H., *Phys Rev.* **D53**, 7327-7335 (1996).
5. Gallis, M. R., *Phys. Rev.* **A53**, 655-660 (1996); Tegmark, M., and Shapiro, H. S., *Phys. Rev.* **E50**, 2538-2547 (1994).
6. Brune, M., Hagley, E., Dreyer, J., Maître, X., Maali, A., Wunderlich, C., Raimond, J-M., and Haroche, S., *Phys. Rev. Lett.* **77**, 4887-4890 (1996).
7. Cheng, C. C., and Raymer, M. G., *Phys. Rev. Lett*, **82**, 4802 (1999)
8. Myatt, C. J., et al., *Nature*, **403**, 269 (2000).
9. Ammann, H., Gray, R., Shvarchuk, I., and Christensen, N., *Phys. Rev. Lett.* **80**, 4111 (1998).
10. Klappauf, B. G., Oskay, W. H., Steck, D. A., and Raizen, M. G., *Phys. Rev. Lett.* **81**, 1203 (1998); Erratum in *Phys Rev. Lett.* **82** 241 (1999).
11. Bennett, C. H., *Physics Today* **48**, No. 10 (1995); Bennett, C. H., and DiVincenzo, D. P., *Nature 404*, 247 (2000).
12. Wootters, W. K., and Zurek, W. H., *Nature* **299**, 802 (1982).
13. Dieks, D., *Phys. Lett.* **A 92**, 271 (1982).
14. Zurek, W. H., *Physics Today* **44**, 36 (1991).
15. Tegmark, M., *Phys. Rev.* **E 61**, 4194 (2000).
16. Zurek, W. H., *Physica Scripta* **T76**, 186 (1998), also available at quant-ph/9802054.
17. Zurek, W. H., *Phil. Trans. R. Soc. Lond.* **A356**, 1793 (1998), also available at quant-ph/9805065.
18. von Neumann, J., "Measurement and reversibility" and "The measuring process", chapters V and VI if *Mathematische Grundlagen der Quantenmechanik*, (Springer, Berlin, 1932); English translation by R. T. Beyer *Mathematical Foundations of Quantum Mechanics*, (Princeton Univ. Press, Princeton, 1955).
19. Monroe, C., Meekhof, D. M., King, B. E., and Wineland, D. J., *Science*, **272**, 1131-1136 (1996).
20. Everett III, H., *Rev. Mod. Phys.* **29**, 454 (1957).
21. Zurek, W. H., pp. 175-212 in *Physical Origins of Time Asymmetry*, Halliwell, J. J., Pérez-Mercader, J., and Zurek, W. H., eds. (Cambridge University Press, Cambridge, 1994).
22. Zurek, W. H., "Information transfer in quantum measurements", pp. 87-116 in *Quantum Optics, Experimental Gravity, and the Measurement Theory*, P. Meystre and M. O. Scully, eds. (Plenum, New York, 1983).
23. Bohm, D., *Quantum Theory*, (Prentice-Hall, Engelwood Cliffs, 1951).
24. Rauch, H., *Physica Scripta* **T76**, 24 (1998).
25. Pfau, T., et al., *Phys. Rev. Lett.* **73**, 1223 (1994).

26. Chapman, M. S., et al., *Phys. Rev. Lett.* **75**, 3783 (1995).
27. Zurek, W. H., manuscript in preparation (2000).
28. Lloyd, S, *Phys. Rev.* **A 55**, 1613 (1996)
29. Schumacher, B., *Phys. Rev.* **A 54**, 2614 (1996).
30. Landauer, R., *Phil. Trans. R. Soc.* **353** 367 (1995); also, in *Proc. of the Drexel-4 Symposium on Quantum Nonintegrability: Quantum – Classical Correspondence*, D. H. Feng and B.-L. Hu, eds. (World Scientific, Singapore, 1998); Unruh, W. G., *Phys. Rev* **A51**, 992 (1995) Chuang, I. L., Laflamme, R., Shor, P., and Zurek, W. H., *Science*, **270**, 1633-1635 (1995).
31. Zeh, H. D., *Found. Phys.* **3**, 109 (1973).
32. Zeh, H. D., *The Physical Basis of the Direction of Time*, (Springer, Berlin, 1989).
33. Albrecht, A., *Phys. Rev.* **D 46**, 5504 (1992).
34. Albrecht, A., *Phys. Rev.* **D 48**, 3768 (1993).
35. Walls D.F. and Milburn G.J., *Quantum Optics*, (Springer Verlag, Berlin, 1994).
36. Chaturvedy, S. and Shibata, F., *Z. Phys.* **B35**, 297 (1979), see also Desposito, M. and Hernandez, S. H., *Physica* **227A**, 248 (1996).
37. Hu, B. L., Paz, J. P., and Zhang, Y., *Phys. Rev.* **D 45**, 2843 (1992).
38. Leggett, A. J., Chakravarty, S., Dorsey, A. T., Fisher, M. P. A., Garg, A., and Zwerger, W., Rev. Mod. Phys. **59**, 1 (1987).
39. Paz, J. P. pp. 213-220 in *Physical Origin of Time Asymmetry*, Halliwell, J. J., Pérez-Mercader, J., and Zurek, W. H., eds. (Cambridge University Press, 1992).
40. Caldeira, A. O., and Leggett, A. J., *Physica* **121A**, 587-616 (1983); *Phys. Rev.* **A 31**, 1059 (1985).
41. Unruh, W. G., and Zurek, W. H., *Phys. Rev.* **D 40**, 1071-1094 (1989).
42. F.Haake and R.Reibold, Phys.Rev, **32**, 2462, (1985).
43. Hu, B. L., Paz, J. P., and Zhang, Y., *Phys. Rev.* **D 47**, 1576 (1993).
44. Grabert, H., Shramm, P., and Ingold, G. L., *Phys. Rep.* **168**, 115 (1988).
45. Davila Romero, L. and Paz, J. P., *Phys. Rev.* **A 53**, 4070 (1997).
46. Feynman R. P., and Vernon F. L., *Ann. Phys.* **24**, 118 (1963).
47. Paz, J. P., Habib, S., and Zurek, W. H., *Phys. Rev.* **D 47**, 488 (1993).
48. Wigner, E. P., *Phys. Rev.* **40**, 749 (1932). For a review, see Hillery, M., O'Connell, R. F., Scully, M. O., and Wigner, E. P., *Phys. Rep.* **106**, 121 (1984).
49. Lindblad, G., *Comm. Math. Phys.* **40**, 119-130 (1976).
50. Garraway, B. M., *Phys. Rev.* **A 55**, 4636 (1997), *ibid* **A 55**, 2290 (1997).
51. Anastopoulos, C. and Hu, B.L., e–print quant-ph/9901078.
52. Zurek, W. H., pp. 145-149 in *Frontiers of Non-equilibrium Statistical Mechanics*, G. T. Moore and M. O. Scully, eds. (Plenum, New York, 1986).
53. Zurek, W. H., Habib, S., and Paz, J. P., *Phys. Rev. Lett.*, **70**, 1187, (1993).
54. Paz, J. P., and Zurek, W. H., *Phys. Rev. Lett.* **82**, 5181 (1999).
55. Anglin, J. R., Paz, J. P., and Zurek, W. H., *Phys. Rev.* **A 53**, 4041 (1997).
56. Gallis, M. R., and Fleming, G. N., *Phys. Rev.* **A 42**, 38 (1990); **A 43**, 5778 (1991); Gallis, M. R., *Phys. Rev.* **A 48**, 1023 (1993).
57. Wisdom, J., Peale, S. J., and Maignard, F. *Icarus* **58**, 137 (1984); see also Wisdom, J., *Icarus* **63**, 272 (1985).
58. Laskar, J., *Nature* **338**, 237 (1989).
59. Sussman, G. J., and Wisdom, J., *Science* **257**, 56-62 (1992).
60. Zurek, W. H., and Paz, J. P., *Phys. Rev. Lett.* **72**, 2508-2511 (1994); *ibid.* **75**, 351 (1995).
61. Zurek, W. H. and Paz, J. P., *Physica* **D83**, 300 (1995).

62. see selected papers in Casati, G., and Chrikov, B., *Quantum Chaos* (Cambridge University Press, Cambridge, 1995).
63. Berman, G. P., and Zaslavsky, G. M., *Physica* (Amsterdam) **A91**, 450 (1978).
64. Habib, S., Shizume, K., and Zurek, W. H., *Phys. Rev. Lett*, **80**, 4361 (1998).
65. Ott, E., Antonsen, T. M., and Hanson, J, *Phys. Rev. Lett.* **35**, 2187 (1984); Dittrich, T., and Graham, R., *Phys. Rev.* **A 42**, 4647 (1990), and references therein.
66. This point of view is related to the one expressed by C. Caves and co–workers who emphasize on "hipersensitivity to perturbations" as the defining aspect of quantum chaos. See Caves, C., and Schack, R., *Hypersensitivity to perturbation: An information-theoretical characterization of classical and quantum chaos*, in *Quantum Communication, Computing, and Measurement*, edited by Hirota, O., Holevo, A. S., and Caves, C. M., (Plenum Press, New York, 1997), pp. 317-330. This criterion was introduced by A. Peres (see Peres, A. *Quantum Theory Concepts and Methods*, Kluger (1995)).
67. Shiokawa, K., and Hu, B. L., *Phys. Rev* **E 52**, 2497 (1995).
68. Miller, P. A., and Sarkar, S. *Phys. Rev.* **E 58**, 4217 (1998); **E 60**, 1542 (1999).
69. Pattanayak, A. K., *Phys. Rev. Lett.* **83**, 4526 (2000).
70. Pastawski, H., Usaj, G., and Levstein, P. *"Quantum chaos: an answer to the Boltzmann–Loschmidt controversy?"*, preprint Famaf (2000); for interesting related experimental work using NMR techniques see also Pastawski, H., Usaj, G., and Levstein, P., *Chem. Phys. Lett.* **261** 329 (1996).
71. Monteoliva, D., and Paz, J. P., (2000) *to appear.*
72. Gell-Mann, M., and Hartle, J. B., in *Complexity, Entropy, and the Physics of Information*, Zurek, W. H., ed. (Addison-Wesley, Reading, 1990).
73. Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O., and Zeh, H. D., *Decoherence and the Appearance of a Classical World in Quantum Theory*, (Springer, Berlin, 1996).
74. Steane A., 1996, *Phys. Rev. Lett.* **77**, 793. Steane A., 1996, *Proc. Roy. Soc. Lond.* **A452**, 2551.
75. Shor P., 1995, *Phys. Rev.* **A 52**, 2493.
76. Mc Williams and Sloane, *"Theory of Error Correcting Codes"* (Elsevier, Amsterdam, 1977).
77. Laflamme R., Miquel C., Paz J.P. and Zurek W.H., *Phys. Rev. Lett.* **77**, 198 (1996).
78. Paz, J. P., and Zurek, W. H., *Proc. Roy. Soc. London* **A 454**, 355 (1998).
79. Calderbank, A. R., Rains, E. M., Shor, P. W., and Sloane, N. J. A., *Phys. Rev. Lett.*, **78**, 405 (1997).
80. Gottesman, D. 1998, Caltech PhD Thesis, quant-ph, see also "Stabilizer codes and quantum error correction", Preprint quant-ph/9705052; *Phys. Rev.* **A 54** 1862 (1996).
81. Knill, E. and Laflamme, R., Preprint quant-ph/9608012; *Phys. Rev.* **A 55**, 900 (1997).
82. Cleve, R., and Gottesman, D., *Phys. Rev.* **56** 76 (1997).
83. H. Pringe, MsC Thesis (unpublished), Buenos Aires University (1997)
84. Halliwell, J. J., *Phys. Rev.* **D 39** 2912 (1989); Kiefer, C., *Class. Quantum Grav.* **4** 1369 (1987); Paz, J. P., and Sinha, S., *Phys. Rev.* **D 45** 2823 (1992); *ibid* **D 44** 1038 (1991); for more recent discussion see Lombardo, F., Mazzitelli, F. D., and Monteoliva, D. *Phys. Rev.* **D** (2000) *to appear.*
85. Poyatos, J. F., Cirac, J. I. and Zoller, P., *Phys. Rev. Lett.* **77** 4728 (1997).
86. Davidovich, L., Brune, M., Raimond, J. M., and Haroche, S., *Phys. Rev.* **A 53** 1295 (1996).

87. Dalvit, D., and Maia Neto, P., *Phys. Rev. Lett.* **87** 798 (2000); *see also* quant-ph/0004057.
88. Anglin, J., *Phys. Rev. Lett.* **79** 6 (1997).
89. Dalvit, D., Dziamarmaga, J., Zurek, W. H., *Phys. Rev.* **A** (2000) *to appear.*
90. Mohanty, P., Jariwada, E. M. Q., and Webb, R. A., *Phys. Rev. Lett.* **77** 3366 (1995); Mohanty, P., and Webb, R. A., *Phys. Rev.* **B55**, R13 452 (1997).

Quantum Information Science Using Photons

Dik Bouwmeester[1], John C. Howell[2], and Antia Lamas- Linares[3]

[1] University of California Santa Barbara, Department of Physics, Santa Barbara CA 93106, USA.
[2] University of Rochester, Department of Physics and Astronomy, Rochester NY, USA.
[3] University of Oxford, Centre for Quantum Computation, Oxford OX1 3PU, UK.

Abstract. Quantum optics provides an ideal medium for implementing some of the most interesting quantum information schemes. This chapter gives an overview of photons as carriers of quantum information and their use in protocols such as quantum teleportation, quantum cryptography and quantum cloning.

1 Introduction

1.1 A Humble Point of View

We are living on a tiny planet rotating around the sun, which is just a small star floating in a sea of many others. Hundreds of millions of stars make up our galaxy, itself a part of an even larger structure. The larger the astronomic structures get the more humble seems our existence.

Looking the other way towards smaller and smaller structures, we enter the wonderful world of biological cells, DNA, smaller molecules, atoms, nuclei, and quarks. Interestingly, our physical size is practically in the middle between the largest astronomical structures and the tiniest structures of which we are aware. It is all a magnificent miracle and we are obviously far, far away from understanding our own position in it all.

Here we are, open eyes, stunned by beauty, shocked by human behavior. Drifting away from our down-to-earth problems we try to understand the large and the small. Surprisingly, our minds seem to be able to cope quite well with the structure of the large scale universe although we are rather disturbed by the notion of infinity which seems to be lurking in the distance. However, it is in trying to understand the small–scale structures that our poor brains really start to struggle.

The fundamental laws of quantum physics that govern the small–scale world (and indirectly the large–scale world) were discovered in the 20th century but so far we have failed to find a logical (generally accepted) underlying explanation for them. The following anecdote about Einstein and Bohr illustrates different points of view concerning our capability of understanding quantum physics. Einstein liked to use phrases such as: "The Lord is subtle but not malicious." This reflected his expectation that his mind should be capable of deeply understanding the fundamental laws of quantum physics. On one of such occasions, Bohr

W. Dieter Heiss (Ed.): LNP 587, pp. 149–197, 2002.
© Springer-Verlag Berlin Heidelberg 2002

Fig. 1. Einstein liked to use phrases such as: "The Lord is subtle but not malicious", or "God does not play dice". On one of such occasions Bohr replied to Einstein: "Please stop telling God what to do!"

replied to Einstein: "Please stop telling God what to do!" Clearly, Bohr placed mankind in a more humble position.

The recent development of Quantum Information Science is somewhat in the spirit of Bohr's remark. We accept the laws of quantum physics, mysterious and beautiful as they are, and try to explore them to design surprising and potentially useful devices. By doing so, it turns out that we do gain more and more insight into the workings of quantum physics. To some extent our understanding of classical physics has followed a similar route. It is through familiarization with the laws rather than through obtaining deeper understanding that we accept the laws of physics as they are.

1.2 Quantum Mystery

The wave–particle duality for microscopic particles is one of the most fundamental and astonishing features of quantum mechanics. The famous double–slit experiments performed at the level of individual particles reveal clear interference patterns after combining the experimental results from many single particle runs. The single-particle interference implies the superposition principle in quantum mechanics which, in our classical language, means that an object experiences different "routes" at the same time. Quantum mechanics provides a description of the "wave function" of the particles and predicts with excellent accuracy the interferences. Despite its simplicity the quantum wave function seems rather mysterious. To extract experimental predictions from it one has to take the square of its amplitudes (in the appropriate basis corresponding to the specific type of measurement) which should be interpreted as probabilities for certain measurement outcomes. This description places probabilities and the concept of information at a most fundamental level. Classically, one uses probabilistic descriptions if the available information about the system under consideration is incomplete (although, in principle, complete information could be obtained). Quantum mechanics tells us that *complete* quantum information can lead to *fundamentally indeterministic* measurement outcomes.

The discussion about the completeness of quantum mechanics and the interpretation of the quantum wave function sharpens when one considers *entangled* states. Entanglement refers to a quantum state of two or more subsystems. For example, consider two spatially separated spin-1/2 particles, one in spatial mode a and one in mode b, described by the two-particle superposition state (along a chosen z axis)

$$\frac{1}{\sqrt{2}}\left(|+\frac{1}{2}\rangle_a|-\frac{1}{2}\rangle_b - |-\frac{1}{2}\rangle_a|\frac{1}{2}\rangle_b\right) . \tag{1}$$

This state expresses that a spin measurement on one of the two particles will have a random outcome (+1/2 or −1/2 along the chosen measurement axis), but once the outcome is known its partner must have the opposite spin. In other words an entangled state describes relations between subsystems rather than individual properties of the subsystems. Most important is the phase factor between the two terms (here a minus sign) which implies that if the spin is measured along a different axis, fixed relations between the particles are still obtained. It is this basis-independent feature of the perfect correlations between measurement outcomes on the two spatially separated particles that makes an entangled state nonclassical and gives it nonlocal features. We will discuss these properties in detail in Sect. 2 on the Einstein–Podolski–Rosen paradox and Bell's inequalities, and show how entangled photons can be produced, Sect. 3, and how they can be used for quantum cryptography, Sect. 6, dense coding, Sect. 7, and quantum teleportation, Sects. 8 and 9.

If one takes the point of view that the quantum-mechanical wave function describes nothing else than the information known about the multi-particle system then it is perhaps not too surprising that the measurement outcome of one particle provides, at the position of the particle subjected to the measurement, the information about the state of the other particle that could be light-years away. Any measurement on this second particle will still have a random outcome for a local observer since it is physically impossible for this observer to have instantaneously access to the information obtained by measurements on the first particle.

The notion of information is at the heart of many discussions in this book and therefore we will spend some time on some classical and quantum mechanical considerations about information starting with a review of the rise and fall of Maxwell's Demon.

1.3 Maxwell's Demon

Naively one might think that information is a very abstract concept which doesn't have much to do with physical concepts like energy and entropy. However, as soon as one realizes that information is the storage of data for a certain period of time it is clear that a physical storage mechanism is required. In order to elucidate that the notion of information is intimately linked to physical systems and to the concept of thermodynamic entropy we review the history of Maxwell's Demon (we refer to the book on "Maxwell's Demon", edited by Leff and Rex, for reprints of the most relevant papers on the topic [1]).

The story of the Demon started with the publication of Maxwell's book in 1871. Near the end of the book in a section on "Limitations of the Second Law of Thermodynamics" Maxwell wrote: "One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. *But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is impossible to us.*" Maxwell continues by considering the situation that a vessel filled with a gas at uniform temperature is divided in two partitions A and B, by a division in which there is a small hole. The supernatural being (later referred to as Maxwell's Demon), who can see the motion of the individual molecules, opens and closes the hole, so as to allow only the slower ones to pass from A to B and only the swifter molecules to pass from B to A. He will thus without expenditure of work, raise the temperature of A and lower that of B, in contradiction to the second law of thermodynamics. Figure 2a is an artist impression of Maxwell's demon who proudly overrules the laws of thermodynamics.

It seems that Maxwell anticipated the breakdown of the second law of thermodynamics at the microscopic level. Many attempts have been made to circumvent such a drastic conclusion by considering, for example, the influence of the measurements performed by the Demon, or the influence of thermodynamical fluctuations on the microscopic Demon. But none of the attempts succeeded and it took over one century before a resolutions was found based on information theory.

The resolution lies in the fact that Maxwell's Demon obtains information about the motion of the gas molecules in making his selection of when to open and close the hole in the partition. In order for the demon to return to its initial state, its memory has to be erased. It is this information erasure, illustrated by RESET in Fig. 2b, that transfers heat to the system and that saves the second law of thermodynamics. To understand this, we review some important historical developments in information theory.

In 1961 Landauer argued that computing machines involve devices which perform logic functions that do not have a single-valued inverse (e.g. the erasure of memory, implicit in the AND and OR logic gates, which takes a computer memory from a certain arbitrary state to a unique standard reference state). This logic irreversibility is associated with physical irreversibility and requires a minimal heat generation, per machine cycle, typically of the order of kT for each irreversible function. A simple way to understand the minimum requirement of dissipation is to note that a binary device must have at least one degree of freedom associated with the information. Classically a degree of freedom is associated with kT of thermal energy. Any switching signals passing between devices must therefore have this much energy to override the thermal noise. Dissipation is essential in order to avoid oscillation back and forth between the logic 0 and 1 states after the switching signal has been applied. In fact any old-fashioned light switch will "click" in order to dissipate the applied switching energy, thus preventing an oscillation between the light on and off.

Although it later turned out that reversible computation is possible (by running a computation backwards after read-out of the computational result, thus avoiding erasure of memory), Landauer's consideration concerning heat dissipation due to erasure of memory is most relevant for resolving the Maxwell's Demon problem. In 1982 Bennett simplified the discussion to its bare essentials by viewing the demon's memory as a two-state system (say, states L and R) that is set in a standard state, N, prior to a measurement and by considering only a single molecule in a vessel. Initially the molecule is freely moving through the vessel. At some moment the demon, unaware of the position of the molecule, inserts a partition trapping the molecule on the left- or right-hand side of the vessel. Next the demon performs a measurement to learn about the position of the molecule,

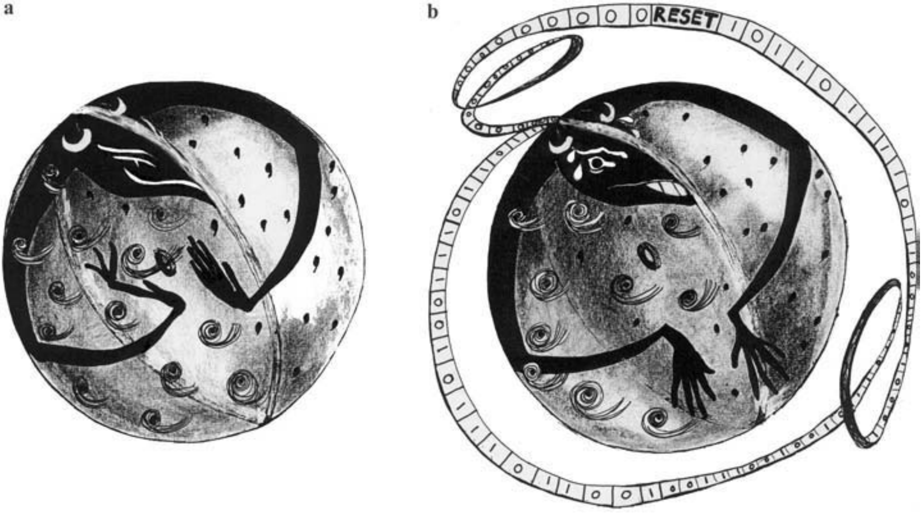


Fig. 2. Artist impression of; (a) Maxwell's demon who proudly seems to violate the second law of thermodynamics by allowing fast (slow) moving gas molecules to go from the right (left) side of the glass bulb to the left (right) side through a hole in the partition wall. (b) Maxwell's demon starts sweating because he is producing heat in the process of erasing his memory about the motion of the molecules. The erasure of information will guaranteed that the second law of thermodynamics is not violated after all.

leaving his mind in state R or L depending on whether the molecule has been detected on the right- or left-hand side. Knowing on which side the particle is, the demon can now extract work from the molecule, by inserting a piston on the side not containing the molecule and allowing the molecule to expand against the piston to fill the whole vessel again. The amount of work extracted is given by

$$W = \int_{V_1}^{V_2} P(V)\mathrm{d}V = \int_{V_1}^{V_2} \frac{kT}{V}\mathrm{d}V = kT \ln \frac{V_2}{V_1} = kT \ln 2\,. \tag{2}$$

Recall from thermodynamics that the amount of work W performed by a system is given by $W = T\Delta S$ where ΔS is the difference between the entropy of the final and initial state. Since the final and initial state of the single molecule in the vessel are identical, the entropy of the demon must have changed. For equilibrium the entropy S, usually defined at the macroscopic level, must be maximized which is associated with maximum disorder in the system. The important link between the entropy S and the microscopic notion of disorder follows from the fact that entropy is an additive quantity, i.e. $S_{1+2} = S_1 + S_2$, where S_1, S_2 and S_{12} are the entropies of systems 1, 2 and the union of systems 1 and 2, whereas the probability of finding the system simultaneously in a particular configuration is multiplicative, i.e. $g_{1+2} = g_1 g_2$. Therefore the relation between S and g has the form $S = k \ln g$, where k is the Boltzmann constant. Since the result of the measurement by the demon transfers his state of mind from N to either R or L, with equal probability, the entropy increases by $k \ln 2$. To return to the initial state of the demon involves erasure of the stored information which corresponds to a two-fold compression of the volume of the demon's phase space. To achieve this compression the same amount of work obtained by letting the molecule expend over the whole vessel is required and will have to be converted into heat, i.e. it will be dissipated.

The important conclusion to be drawn from the above discussion is that *information is physical* and discarding information results in heat transfer to the environment.

1.4 Shannon Entropy

Shannon introduced the concept of entropy into classical information theory to study the optimum compression of data and the maximum rate of reliable communication over a noisy communication channel.

The Shannon Entropy of a system is the information of the system expressed in units of classical bits. To see the analogy between entropy S, as described in the previous section, and the Shannon entropy H we introduce an 'information function' $I(E)$ whose value is determined by the event E. We assume that $I(E)$ is a function only of the probability p for the event E to take place. The interpretation of I as an information function implies that the information gained when two independent events occur with individual probabilities p and q is simply the sum of the information gained from each event alone, i.e.

$$I(pq) = I(p) + I(q)\,. \tag{3}$$

Whereas the notion of information is additive the probabilities are multiplicative. If we further require that we express information in ‘bits’ the functional dependence of I on p becomes

$$I(p) = C \log p \,, \tag{4}$$

for some constant C and for the log taken to be base two. Choosing $C = -1$ the average information gain when one of a mutually exclusive set of events with probabilities $p_1, ..., p_n$ occurs is defined as the Shannon entropy

$$H(p_1, ..., p_n) \equiv - \sum_x p_x \log p_x \,, \tag{5}$$

with the additional convention that $0 \log 0 \sim 0$.

The entropy H was introduced by Shannon in order to quantify the minimum resources needed to store information. As an example, let us consider that we have an information source that transmits words composed out of letters of an alphabet of eight elements. If all letters have equal probability, $p_i = 1/8$, $i = 1, ..., 8$, to appear per transmitted letter, a compact way of representing the letters in binary notation would use the obvious 3-bit code.

$$\begin{aligned} 1^{\mathrm{st}} &\rightarrow 000, \\ 2^{\mathrm{nd}} &\rightarrow 001, \\ &\dots \\ 8^{\mathrm{th}} &\rightarrow 111. \end{aligned} \tag{6}$$

This optimum 3-bit code is confirmed by the Shannon entropy

$$H = - \sum_i p_i \log p_i = 8 \frac{1}{8} \log 8 = 3 \,. \tag{7}$$

If, like in our own alphabet, the letters are not equally likely to appear (e.g. the letters “a” and “e” appear much more frequent than “q” and “z”) a 3-bit code is certainly not the optimum way of encoding the eight letters. For example if a certain letter is much more likely to appear than others it would be profitable to encode this letter in a bit string as short as possible and use longer strings for less frequently occurring letters. One must be careful if one does not use the same code length for different letters because in that case it is not a priori obvious which bits of a long string of bits should be taken together to encode a letter. One possible way is to use the convention that the bit value 0 will indicate the end of a code, in addition to specifying the maximum length of a code. The most frequently occurring letter can then be encoded by 0 itself, the one-but-most-frequently occurring letter by 10, and the most unlikely letter by the maximum length code 1...1. It is far from obvious what the most compact way of encoding is and the Shannon entropy provides exactly the answer to that question. If, for example, the probabilities for the eight letters are given by: $p_6 = 1/2$, $p_1 = p_3 = p_8 = 1/8$, $p_2 = p_4 = p_5 = p_7 = 1/32$ the Shannon information entropy, H, is 2.25 bits.

1.5 Von Neumann Entropy

Various familiar notions from thermodynamics can be rephrased in information theory. For example, entropy measures the disorder in a system whereas the Shannon entropy measures the uncertainty associated with a classical probability distribution. The notion of Shannon entropy can be applied to Quantum Mechanics where classical probability distributions are replaced by density operators. Von Neumann defined the entropy of a quantum state, given by the density matrix ρ, by the formula

$$S(\rho) \equiv -k_B \mathrm{tr}(\rho \ln \rho)\,, \tag{8}$$

which for qubits is modified to give,

$$S(\rho) \equiv -\mathrm{tr}(\rho \log \rho)\,, \tag{9}$$

with the log taken to base two. If λ_x are the eigenvalues of ρ then the von Neumann entropy can be expressed as

$$S(\rho) = -\sum_x \lambda_x \log \lambda_x\,, \tag{10}$$

which strongly resembles the definition of the Shannon entropy H.

As an simple application of the von Neumann entropy we describe how to quantify entanglement in a pure state of two subsystems A and B. The general state of A and B (one of dimension N and the other of dimension $M \leq N$) can be written as (the so-called Schmidt decomposition)

$$|\Psi_{AB}\rangle = \sum_{i=1}^{N} c_i |u_i\rangle |v_i\rangle\,, \tag{11}$$

where $\{|u_i\rangle\}$ is a basis for A and $\{|v_i\rangle\}$ is a basis for B. The von Neumann entropy of system B, i.e. the amount of uncertainty about the state of system B before a measurement is made on system A is $S(\rho_B)$, where ρ_B is the reduced density matrix of system B

$$\rho_B = \mathrm{Tr}_A \rho_{AB} = \sum_p |c_p|^2 |v_p\rangle\langle v_p|\,. \tag{12}$$

After the measurement on system A the state of B will be uniquely defined, i.e. if we obtain $\{|u_i\rangle\}$ for A then the state of B is $\{|u_i\rangle\}$, and $S(\rho_B)$ will be 0. So the information gained is $S(\rho_B) = S(\rho_A)$. Thus A and B are maximally entangled if the reduced density matrices are maximally mixed. For more than two entangled subsystems and for mixed states in general, a Schmidt decomposition does no longer exist and the quantification of the amount of entanglement becomes very difficult.

In the previous sections we briefly touched upon concepts in information theory. In general, classical and quantum information theory deals with transmission, optimal compression, and optimal protection against noise of information.

In addition, quantum information theory includes the quantization and classification of entanglement which is closely related to entanglement purification and entanglement distillation. We refer to the excellent lectures notes by Preskill [2], and to [3], for much more information and further references.

2 Einstein–Podolsky–Rosen Paradox and Bell's Inequalities

In 1935 Einstein, Podolsky and Rosen (EPR) put forward a seminal paper questioning the completeness of quantum mechanics [4]. They analyzed the situation of two spatially separated particles that are entangled in position and momentum. In 1951, Bohm framed the EPR problem in terms of a Gedanken experiment based on the dissociation of a spin-zero two-atom molecule where each atom had a spin-1/2 [5]. After dissociation the wave function of the system has the form $\frac{1}{\sqrt{2}}\left(|+\frac{1}{2},-\frac{1}{2}\rangle - |-\frac{1}{2},+\frac{1}{2}\rangle\right)$. Concerning measurements on the two-atom system Bohm wrote: "Suppose now that one measures the spin angular momentum of any one of the particles...Because of the existence of correlations, one can immediately conclude that the angular-momentum vector of the other particle is equal and opposite".

The debate about the completeness of quantum mechanics was considered to be merely philosophical until 1964 when John Bell showed that quantum mechanics and hidden-variable theories were mathematically incompatible [6]. He derived an inequality based on Bohm-type quantum systems which showed that any local realistic theory and quantum mechanics predicted two different probabilistic outcomes. His work was further elaborated on by Clauser, Horne, Shimony and Holt (CHSH) [7]. The first experimental tests of nonlocality used Bohm-type polarization correlations of two-photon cascade (for a review of early two-photon cascade experiments see [8]). With advancements in technology, improved sources of polarization entangled photons were obtained using parametric down-conversion in nonlinear crystals (see Sect. 3). Recently, two entangled two-level ions were also used to violate a Bell Inequality [9].

The derivation of Bell's inequalities start with the following consideration. If a local realistic theory can account for the correlations that Alice and Bob measure, then it must be true that Alice's measurement outcome A ($+\frac{1}{2}$ or $-\frac{1}{2}$) must be independent of Bob's analyzer orientation b and measurement outcome B ($+\frac{1}{2}$ or $-\frac{1}{2}$) and vice versa. Alice's and Bob's probabilistic measurement outcomes must then decouple as

$$P(A,B|a,b,\lambda) = P(A|a,\lambda)P(B|b,\lambda) \tag{13}$$

where λ accounts for all possible local hidden variables. Assuming local hidden variables one can define a measurement outcome given by

$$E^{\mathrm{HV}}(a,b) \equiv \int \mathrm{d}\lambda f(\lambda)\overline{A}(a,\lambda)\overline{B}(b,\lambda). \tag{14}$$

where

$$\overline{A}(a,\lambda) \equiv P(+\frac{1}{2}|a,\lambda) - P(-\frac{1}{2}|a,\lambda), \tag{15}$$

$$\overline{B}(b,\lambda) \equiv P(+\frac{1}{2}|b,\lambda) - P(-\frac{1}{2}|b,\lambda)\,. \tag{16}$$

Here $P(\pm\frac{1}{2}|a,\lambda)$ is the probability of obtaining measurement result $\pm\frac{1}{2}$ for orientation a of the analyzer. Because the signs of the probabilities in $\overline{A}(a,\lambda)$ and $\overline{B}(b,\lambda)$ are different, it must be true that $|\overline{A}(a,\lambda)| \leq 1$ and $|\overline{B}(b,\lambda)| \leq 1$. The derivation proceeds as

$$E^{\mathrm{HV}}(a,b) - E^{\mathrm{HV}}(a,b') = \int \mathrm{d}\lambda f(\lambda)\overline{A}(a,\lambda)(\overline{B}(b,\lambda) - \overline{B}(b',\lambda)). \tag{17}$$

Since $|\overline{A}_1(a,\lambda)| \leq 1$, then

$$|E^{\mathrm{HV}}(a,b) - E^{\mathrm{HV}}(a,b')| \leq \int \mathrm{d}\lambda f(\lambda)|(\overline{B}(b,\lambda) - \overline{B}(b',\lambda))|; \tag{18}$$

similarly,

$$|E^{\mathrm{HV}}(a',b) + E^{\mathrm{HV}}(a',b')| \leq \int \mathrm{d}\lambda f(\lambda)|(\overline{B}(b,\lambda) + \overline{B}(b',\lambda))|\,. \tag{19}$$

Since $|\overline{B}(b,\lambda)| \leq 1$, then

$$|\overline{B}(b,\lambda) - \overline{B}(b',\lambda) + \overline{B}(b,\lambda) + \overline{B}(b',\lambda)| \leq 2 \tag{20}$$

which implies

$$S \equiv |E(a,b) - E(a,b') + E(a',b) + E(a',b')| \leq 2 \tag{21}$$

Hence, the maximum possible value of S that can be achieved, assuming locally explicable outcomes is 2. On the other hand quantum mechanics predicts

$$\begin{aligned} E^{\mathrm{QM}}(a,b) = {} & P(+\frac{1}{2},+\frac{1}{2}|a,b) - P(+\frac{1}{2},-\frac{1}{2}|a,b) \\ & - P(-\frac{1}{2},+\frac{1}{2}|a,b) + P(-\frac{1}{2},-\frac{1}{2}|a,b) \end{aligned} \tag{22}$$

Using the Bell inequality in Eqn. (21) a theoretical maximum violation of $2\sqrt{2}$ is obtained. This prediction is obtained using analyzer rotations of $a = 0^o$, $a' = 45^o$, $b = 22.5^o$, and $b' = 67.5^o$. Any value of $S > 2$, violates a Bell inequality and thus favors a quantum mechanical explanation and contradicts predictions based on the assumption of local realism. Many experiments have been used to verify Bell's predictions. Experiments of interest are found in [8,10,11,12,13,9,14]. In Sect. 12 we will discuss a generalization to spin-1 Bell inequalities.

3 Producing Entangled Particles

3.1 Introduction

The initial methods for creating entangled particles were inspired by Bohm's original idea of producing a pair of spin (polarization) entangled particles. In the early experimental test of Bell's inequality an atomic cascade was used to produce pairs of polarization-entangled photons [8]. This source did not have a very high entanglement purity due to the recoil of the atoms, and the light emitting from them was hard to collect efficiently. In the early 90's the nonlinear optical process of parametric down-conversion (PDC) was pinpointed as having promise for the generation of these states. Nowadays the most widely used source of entangled photons is still based on this method, although with important developments since the initial implementations. There are several very interesting experiments dealing with the creation of entanglement between particles other than photons but a discussion of these experiments is beyond the scope of these notes. A good overview of experiments with trapped ions, Rydberg atoms, and nuclear spins can be found in [3].

3.2 Parametric Down-Conversion

Parametric down-conversion (PDC) is a process by which a pump photon in a nonlinear crystal has a small probability of splitting into two photons of lower frequency. Of course, energy and momentum have to be conserved when this happens, giving us the *phase matching conditions*,

$$\omega_{\mathrm{p}} = \omega_1 + \omega_2, \tag{23}$$

$$\boldsymbol{k}_{\mathrm{p}} = \boldsymbol{k}_1 + \boldsymbol{k}_2. \tag{24}$$

The subindices p, 1 and 2 correspond to *pump*, and the two down-converted photons 1 in mode a and 2 in mode b (in the literature photons 1 and 2 are often referred to signal and idler photons). The magnitude of the vector $\boldsymbol{k}$ inside the crystal is given by $|\boldsymbol{k}| = \frac{\omega}{c} n(\omega)$. In general due to dispersion, that is, light of different frequencies have different refractive indices, the two conditions above cannot be satisfied simultaneously for light with the same polarization. We can however explore the fact that nonlinear crystals can also be birefringent, that is each photon produced will see a different refractive index depending on its polarization and its propagation direction with respect to the optical crystal axis. Depending on which photon sees what refractive index we distinguish two different classes of down-conversion. In type-I the two down-converted photons emerge with the same polarization –say the extraordinary– while the pump was originally in the orthogonal– in this case the ordinary. Type-II occurs when the two down-converted photons have orthogonal polarizations. The rather complex emission spectra arises from these phase matching conditions as all the possibilities of fulfilling them are realized.

One of the most interesting features of PDC is that the down-converted photons are always produced in pairs and are extremely tightly correlated in their time of emission [15] as well as in their energy. Taken individually each of them has a very large bandwidth (typically larger than 30nm) and its arrival time is essentially determined by the pump coherence length. Already there is here a form of entanglement but it is not in a convenient form so some refinements will have to be added.

3.3 Franson's Proposal

An ingenious way of entangling photons came from a proposal by Franson [16]. The main requirement was to have a mechanism by which pairs of photons are created simultaneously, these photons were each directed into an unbalanced Mach–Zender (MZ) interferometer and the output analyzed in coincidence.

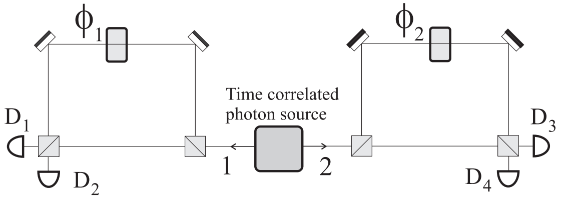


Fig. 3. Schematic drawing of Franson's proposal [16] for generating entangled pairs of photons.

By looking at Fig. 3 it is simple to see that there are only four distinct possibilities at the output of the MZ:

$$|\psi\rangle = |s_1, s_2\rangle + |s_1, l_2\rangle + |l_1, s_2\rangle + |l_1, l_2\rangle \,, \tag{25}$$

where s and l stand for the short and long arms in the MZ interferometers. By looking in coincidence, that is observing one photon on each side that should have taken both the long or both the short paths, and by introducing phase shifters in the setup we obtain the following entangled state

$$|\psi\rangle = |s_1, s_2\rangle + e^{\phi_i + \phi_s} |l_1, l_2\rangle \,. \tag{26}$$

It is important that the difference in path length between the arms of the interferometer is longer than the coherence length of a single photon –in order to avoid single-photon fringes– but shorter than the coherence length of the pump so that the times of arrival are still adequately correlated and the states $|ss\rangle$ and $|ll\rangle$ are indistinguishable. If these conditions are verified, the coincidence rate, say between D_1 and D_3, is given by

$$|1 + e^{\mathrm{i}\phi}|^2 = 2 + 2\cos\left(\frac{\omega_\mathrm{p}}{2c}(\Delta L_i + \Delta L_s)\right) \tag{27}$$

An experimental realization of the Franson interferometer can be found in [17].

3.4 Polarization Entanglement

A very successful and widely used method [11] for generating entangled photons is based on type-II PDC and produces polarization entanglement. In principle it might seem irrelevant in what degree of freedom the quantum correlations occur, but of course in practice some variables are more manageable than others. A brief look at the literature shows that the great majority of discussions of the EPR paradox do not use position and momentum (as originally envisaged by EPR) but non orthogonal polarization bases. Not only is this easier to understand from a conceptual point of view, it is experimentally very convenient. There exists a wealth of very accurate apparatuses for measuring and manipulating the polarization of light that would be difficult to equal for any other equivalent variable.

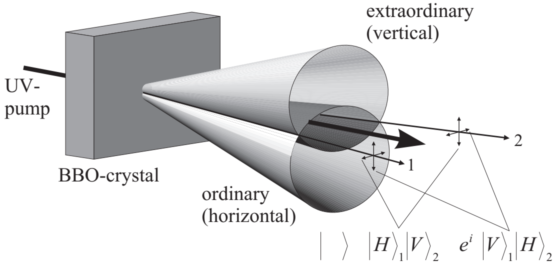

extraordinary
(vertical)
UV-
pump
BBO-crystal
ordinary
(horizontal)
2
1
$| \quad \rangle \quad |H\rangle_1|V\rangle_2 \quad e^{i} \quad |V\rangle_1|H\rangle_2$

Fig. 4. Type-II parametric down-conversion degenerate emission cones [3]. In the direction where the ordinary and extraordinary cones intersect an entangled state in polarization can be produced.

Consider the non collinear type-II PDC emission spectrum. The phase matching conditions force photons with conjugate frequencies to be emitted along two non concentric cones, one with the ordinarily polarized photons while the other carries the extraordinarily polarized ones. In addition, the angle between the cones is determined by the orientation of the optical axis with respect to the incident beam and thus we can control how close the cones are by simply tilting the crystal. If we limit our discussion to the degenerate case –$\omega_1 = \omega_2$– we can see that (Fig. 4), in the directions where the cones intersect each other, there is a superposition of two possibilities, namely $|H_1V_2\rangle$ and $|V_1H_2\rangle$. In order to obtain a quantum superposition of these two terms there must be no way other than by their polarization by which the terms could be distinguished. Unfortunately, one small detail is keeping a tag on the two possibilities and making them in principle distinguishable: the birefringence of the crystal. When propagating through a birefringent media, the ordinary and extraordinary beams see different refractive indices and will travel at different speeds, so only pairs of photons created at the end-face of the crystal will arrive simultaneously at the detectors, while those produced at the in-face will have a delay between H and V photons of $\delta T = L\left(\frac{1}{u_{\mathrm{o}}} - \frac{1}{u_{\mathrm{e}}}\right)$. If this time delay is larger than the coherence time of the photons the two processes are distinguishable and will not lead to polarization entangled photons. The solution is to engineer an additional time delay that will restore the indistinguishability. This can be done by putting a crystal similar to the one used to produce the photons but of half the thickness in each of the paths and rotate it by 90^0. This will reverse the effect of the original delay by half such that pairs coming from symmetric points with respect to the center of the crystal generate a proper superposition of indistinguishable HV or VH emissions in the same time slot (see Fig. 5). This is clearly demonstrated by introducing these so called compensators which enable the production of high quality (97% purity) entangled states of the form

$$\psi = \frac{1}{\sqrt{2}}(|H_1V_2\rangle + e^{\mathrm{i}\phi}|V_1H_2\rangle)|, \tag{28}$$

As an additional advantage, tilting one of the compensators (the one in the path of photon 1) provides a fine tune of the relative phase between H_1 and V_1 which is an easy way of modifying the phase ϕ in (28).

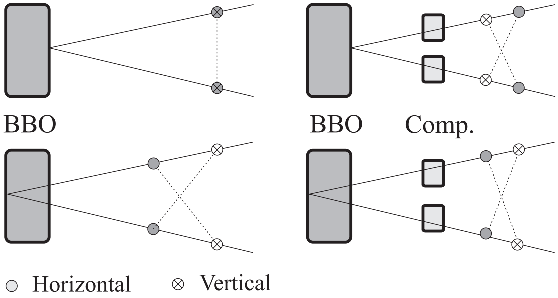


Fig. 5. This figure shows how using compensating elements eliminates the time labelling produced by the crystal birefringence. At the left the photons created some depth inside the crystal come out at different times, the one with extraordinary polarization always ahead of the ordinary. In the figure we see that only those photons created just before getting out of the crystal are properly entangled. When using compensators (right) the time delays are reversed by half so that pairs coming from symmetric points with respect to the center of the crystal generate a superposition of indistinguishable HV or VH emissions in the same time slot.

Recently, yet another source of polarization entanglement was proposed by Kwiat et al. [18] The idea is to take type-I down-conversion crystals with the optical axes rotated by 90 degrees with respect the each other and use a pump with linear polarization oriented at 45 degrees. Each crystal can create pairs of photons, and as long as you cannot tell from which crystal the photons came, you will have a superposition of the two possible two-photon states,

$$|\psi\rangle = |HH\rangle + e^{i\phi}|VV\rangle \,. \tag{29}$$

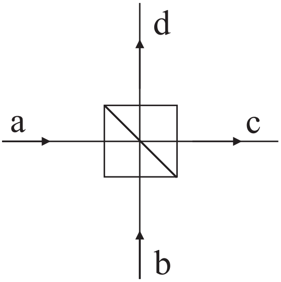


Fig. 6. Beam splitter with input modes *a* and *b* and output modes *c* and *d*.

4 The Beam Splitter Action on a Two-Photon State

4.1 Beamsplitter Transformation

Beam splitter (BS) play and important role in many quantum-optical implementations of quantum information protocols. The action of a beam splitter BS with incoming light modes *a*, *b* and out coming modes *c*, *d*, see Fig. 6, is characterized by the following unitary transformation

$$c = ta + \mathrm{i}rb \tag{30}$$

$$d = ira + tb\,, \tag{31}$$

where t and r are real-valued transmission and reflection coefficients ($r^2+t^2 = 1$). Let us take as input the state $|\Psi\rangle_{\text{in}}$ of one horizontal (H) photon in mode a and one horizontal (H) photon in mode b

$$|\Psi_{\text{in}}\rangle = a^\dagger_H b^\dagger_H|\text{vac}\rangle = |1_{a_H}, 1_{b_H}\rangle\,, \tag{32}$$

where $|\text{vac}\rangle$ represents the vacuum radiation field and $a^\dagger$ ($b^\dagger$), a (b) are the bosonic creation and annihilation operators for each mode:

$$a^\dagger|N\rangle = \sqrt{N+1}|N+1\rangle, \tag{33}$$

$$a|N\rangle \ = \sqrt{N}|N-1\rangle\,. \tag{34}$$

Applying the inverse of the transformations (30, 31) to this input state, one obtains the output state:

$$|\Psi_{\text{out}}\rangle = (T-R)|1_{c_H}, 1_{d_H}\rangle - i\sqrt{2RT}|2_{c_H}, 0_d\rangle - i\sqrt{2RT}|0_c, 2_{d_H}\rangle\,, \tag{35}$$

where $R = r^2$ and $T = t^2$ are the reflectivity and the transmissivity of the BS.

For a perfect 50:50 BS, $T = R = 0.5$, the first term in (35) is zero by virtue of destructive interference of the two-photon probability amplitudes. The remaining two terms imply that the two photons will exit the BS at the same output port. This effect is the basis of the Hong–Ou–Mandel [15] dip in coincidence recording between two single-photon detectors in modes c and d as function of delay of arrival time at the detectors of photons initially from modes a and b. It is very important that the two input photons are indistinguishable in polarization and in time of arrival at the detectors. If, for example, the photon in input mode a was V polarized instead of H the final state becomes

$$|\Psi_{\text{out}}\rangle = \frac{1}{2}(|1_{c_V}, 1_{d_H}\rangle - i|1_{c_V}, 1_{c_H}\rangle - i|1_{d_V}, 1_{d_H}\rangle - |1_{c_H}, 1_{d_V}\rangle) , \tag{36}$$

and no terms will cancel.

Although the action of a BS on the input of two photons is often referred to as a two-photon interference effect it should be pointed out that the effect is insensitive to the phase of either of the incoming photons nor to their relative phase.

4.2 Bell-State Analyzer

We now explain how 50:50 beam splitters and polarizing beam splitters can be used to distinguish between various polarization-entangled photon states. In the literature one often finds the state of a pair of polarization-entangled photons given by one of the four maximally-entangled Bell states

$$|\Psi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 \pm |V\rangle_1|H\rangle_2), \tag{37}$$

$$|\Phi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|H\rangle_1|H\rangle_2 \pm |V\rangle_1|V\rangle_2) . \tag{38}$$

These four states form a complete basis in which any two-photon polarization state can be expressed. One might be surprised by the fact that the $|\Psi^-\rangle_{12}$ is anti symmetric under particle exchange (that is, exchanging labels 1 and 2 results in -$|\Psi^-\rangle_{12}$) whereas the remain three states are symmetric. How could it be that two photons have anti symmetric, that is fermionic, statistics? The reason for the confusion is that the states given in (37) and (38) only refer to the polarization properties of the photons but not to their full wave function. The complete picture only emerges if one simultaneously considers the remaining degrees of freedom of the particles. What one finds is that the polarization entanglement is ultimately linked to entanglement in some other degrees of freedom; if two bosons are in an anti symmetric polarization entangled state they have to be in an anti symmetric state in some other degrees of freedom in order to obey the overall bosonic symmetry. Let us consider the spatial modes as an additional degree of freedom. If we restrict our attention to two photons, one in spatial mode a and one in spatial mode b, then the symmetric and anti-symmetric (entangled) spatial states are given by

$$|\Psi_S\rangle_{12} = \frac{1}{\sqrt{2}}(|a\rangle_1|b\rangle_2 + |b\rangle_1|a\rangle_2), \tag{39}$$

$$|\Psi_A\rangle_{12} = \frac{1}{\sqrt{2}}(|a\rangle_1|b\rangle_2 - |b\rangle_1|a\rangle_2 . \tag{40}$$

The anti symmetric entangled polarization state $|\Psi^-\rangle_{12}$ has to be linked to $|\Psi_A\rangle_{12}$ and the remaining three symmetric states $|\Psi^+\rangle_{12}$ and $|\Phi^\pm\rangle_{12}$ are linked to $|\Psi_S\rangle_{12}$. It is through this link that beam splitters acting on the spatial part of the two-photon wave function can be used to distinguish polarization entangled states.

A very convenient notation for the four Bell states that incorporates the link between polarization and spatial entanglement is given by

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(a_H^\dagger b_V^\dagger + a_V^\dagger b_H^\dagger)|\mathrm{vac}\rangle, \tag{41}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(a_H^\dagger b_V^\dagger - a_V^\dagger b_H^\dagger)|\mathrm{vac}\rangle, \tag{42}$$

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(a_H^\dagger b_H^\dagger + a_V^\dagger b_V^\dagger)|\mathrm{vac}\rangle, \tag{43}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(a_H^\dagger b_H^\dagger - a_V^\dagger b_V^\dagger)|\mathrm{vac}\rangle. \tag{44}$$

It is now straightforward to verify that using the transformations described by (30) and (31) for a 50:50 BS results in

$$|\Psi^+\rangle \rightarrow \frac{-i}{\sqrt{2}}(c_H^\dagger c_V^\dagger + d_H^\dagger d_V^\dagger)|\mathrm{vac}\rangle, \tag{45}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(c_H^\dagger d_V^\dagger - c_V^\dagger d_H^\dagger)|\mathrm{vac}\rangle, \tag{46}$$

$$|\Phi^+\rangle = \frac{-i}{2\sqrt{2}}({c_H^\dagger}^2 + {c_V^\dagger}^2 + {d_H^\dagger}^2 + {d_V^\dagger}^2)|\mathrm{vac}\rangle, \tag{47}$$

$$|\Phi^-\rangle = \frac{-i}{2\sqrt{2}}({c_H^\dagger}^2 - {c_V^\dagger}^2 + {d_H^\dagger}^2 - {d_V^\dagger}^2)|\mathrm{vac}\rangle. \tag{48}$$

After passing the beam splitter, the $|\Psi^-\rangle$ state is the only one that results in one photon in each output modes (c and d). In fact $|\Psi^-\rangle$ is the eigenstate of the beam splitter. For the three remaining states $|\Psi^+\rangle$ is distinguished from the other two by the photons in each term have orthogonal polarizations. This state can be uniquely identified by inserting additional polarizing beam splitters in modes c and d and by observing coincidence recordings at a pair of single-photon detectors behind one of the polarizing beam splitters. The states $|\Phi^+\rangle$ and $|\Phi^-\rangle$ remain degenerate for this detection scheme but could jointly be identified if detectors are used that can distinguish two-photon impact from a one-photon impact.

5 No-Cloning Theorem

The no-cloning theorem in quantum physics is most important for understanding the virtues of quantum cryptography (to be discussed in Sect. 6) and quantum teleportation (to be discussed in Sect. 8). The discussion about cloning of quantum states was initiated by Herbert in his interesting, but soon to be discarded, proposal for superluminal communication. The idea was to combine entanglement and quantum-state cloning to achieve non local communication between our friends Alice and Bob. The scheme goes as follows. Alice and Bob share a polarization entangled pair of photons in the state

$$|\Psi^-\rangle_{12} = \frac{1}{\sqrt{2}}(|H\rangle_A|V\rangle_B - |V\rangle_A|H\rangle_B)\,. \tag{49}$$

Alice performs a polarization measurement in either the $|H\rangle$, $|V\rangle$ basis or in the 45° rotated basis ($|+45^\circ\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$, $|-45^\circ\rangle = (|H\rangle - |V\rangle)/\sqrt{2}$). As soon as Alice obtains a measurement result, instantaneously the state of the particle in Bob's hands is reduced to the state orthogonal to the one measured by Alice. If Bob could uniquely determine the polarization of his particle he would be able to tell in which basis Alice performed her measurement, thereby establishing super luminal communication. However, this seems impossible since Bob has to choose at random a projection basis, which might well be different from the one that Alice used, for the detection of his photon. But now imagine that Bob has a machine that can make copies of the incoming one photon state, e.g. via a stimulated emission process. This would provide him with an ensemble of identical particles from which it is possible to determine the polarization.

Wootters and Zurek [19] and Dieks [20] showed that the cloning procedure as invoked in the proposal by Herbert is in conflict with the basic laws of quantum mechanics. To see this consider a device, initially in state $|\psi_0\rangle$, which produced perfect copies of V and H photons, i.e. which performs the following unitary transformation:

$$|V\rangle|\psi_0\rangle \to |V\rangle|V\rangle|\psi_V\rangle, \tag{50}$$
$$|H\rangle|\psi_0\rangle \to |H\rangle|H\rangle|\psi_H\rangle. \tag{51}$$

Next consider the action of this device on the $|45^\circ\rangle$ polarization state,

$$|45^\circ\rangle = \frac{1}{\sqrt{2}}(|V\rangle + |H\rangle)|\psi_0\rangle \to \frac{1}{\sqrt{2}}(|V\rangle|V\rangle|\psi_V\rangle + |H\rangle|H\rangle|\psi_H\rangle). \tag{52}$$

Clearly, even in the case that $|\psi_V\rangle = |\psi_H\rangle = |\psi_{45^\circ}\rangle$ this output is not of the desired form $|45^\circ\rangle|45^\circ\rangle|\psi_{45^\circ}\rangle$. It is the linearity of quantum mechanics that simply forbids the cloning of an arbitrary quantum state. It was also pointed out by Mandel [21] and by Milonni and Hardies [22] that perfect cloning in stimulated emission, such as proposed by Herbert is prevented by the unavoidable presence of spontaneous emission.

Although it clear that perfect cloning is not possible it is also clear that some polarization information can be extracted from a measurement of a single photon. If this would not be that case one could not even extract any information from an ensemble of identical photons. With the information obtained from measurements on a single photon one could produce new photons that are "poor" copies of the initial state. This implies that imperfect cloning is possible and the natural question that arises is: How well can one clone an arbitrary quantum state? This question was answered by Buzek and Hillery in 1996 [22]. They proposed a device that would produce two approximate copies for an arbitrary input qubit. The quality of the copies can be quantified by the fidelity F with respect to the input state, $F = \langle\psi|\rho_a|\psi\rangle$, where $|\psi\rangle$ is the state of the original qubit and ρ_a is the reduced density matrix of one of the two copies. The universal, that is state independent, value for the fidelity turned out to be $F = 5/6$. A generalization to the case of N to M quantum cloning has been by Gisin and Massar [24].

The value of 5/6 for the fidelity of 1 to 2 cloning can be understood by taking into account a physical system that would perform the cloning. Since cloning is similar to stimulated emission let us consider a polarization independent gain medium in which the stimulated emission takes place. This situation has been discussed by Simon et al.[25]. All that needs to be considered are the properties of the creation operator for a photon in the desired (polarization-degenerate) field mode; $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$. If the spatial mode a is occupied by a given photon with a specific polarization state (say $|1_H, 0_V\rangle$) then the action of the creation operator is $a^\dagger_H|1_H, 0_V\rangle = \sqrt{2}|2_H, 0_V\rangle$. This produces a perfect copy of the input qubit. In addition we have to take spontaneous emission into the orthogonal polarization mode into account.

This is described by the action of the creation operator on the vacuum state which produces a single photon in the mode, ${}^\dagger_V|1_H, 0_V\rangle$. The probability for stimulated emission is twice as high $(\sqrt{2})^2$ as the probability for spontaneous emission. The ratio of having two perfect copies to having the initial state together with its orthogonal state after passing the gain medium is therefore 5 to 1. This yields the fidelity of 5/6.

6 Quantum Cryptography

Cryptology, which in Greek means “hidden word”, is the science of secure communication. The use of cryptology dates back to the ancient Egyptians, Hebrews, Assyrians and Greeks. Numerous methods for masking messages have been developed over the last 2500 years. The ability to hide the meaning of a plain message through cryptographic means is of obvious import in everything from military applications to internet banking.

Employment of a secret key, along with a publicly-distributed cipher-text, is the masking method described herein. Access to the key, say a long string of randomly generated logic zeros and ones, is limited to the selected users, Alice and Bob. As the originator of the message, Alice has a binary plaintext message. She then employs her secret key to encrypt the plaintext binary string using her secret key making a ciphertext. The ciphertext is simply obtained by addition modulo 2 of each bit in the key (with the same length as the message) to each bit of the message. The ciphertext, which is now as random as the key, is distributed over a public channel making it accessible to Bob. Bob uses his key to decrypt the message, again by addition modulo 2. Thus, the security of the message depends entirely on the security of the key. Quantum mechanics will enable Alice and Bob to obtain a random key of which the security can be checked before using it!

Concurrent with many of the advances in cryptography have been the advances in cryptanalysis, the science of deciphering a message without a key or "code-breaking". Currently, the industry standard for key decryption is based on the computational difficulty of factoring a large number into two prime numbers. While it is easy to multiply two prime numbers, it is very difficult, at least classically, to do the problem in reverse. However, Peter Shor showed that the large state spaces of quantum computers can solve this problem exponentially faster than classical computers [26]. The experimental realization of a quantum computer which could perform such calculations is arguably in the distant future. However, Shor's factoring scheme does point out the susceptibility of classical cryptographic schemes.

In the last two decades, it has been shown that basic laws of quantum mechanics can be used for cryptographic purposes. Surprisingly, quantum cryptosystems have found a firm foothold in the field of cryptography. The original quantum cryptographical protocol was proposed by Bennett and Brassard in 1984 [27]. It was based on the fact that it is impossible to clone (make an exact copy of) an unknown quantum state (see the previous section). Alice sends a quantum particle (e.g. a photon) to Bob in one of 4 polarization orientations of 2 nonorthogonal bases. Bob randomly orients his analyzer in one of the two bases and measures the photon. After the transmission, Alice and Bob publicly announce their bases. Alice and Bob "throw out" the data collected in different bases. The remaining data was collected when Alice and Bob had chosen the same basis for sending and receiving the photon respectively. Hence, assuming a noiseless quantum channel, Alice and Bob should have identical strings of bits. If Eve has "tapped" the quantum channel and analyzed the data in the wrong basis, she will introduce bit flip errors. The exposure of the eavesdropper means that the key must be discarded as the security of the key has been jeopardized.

Another significant quantum cryptosystem was proposed by Ekert in 1991 [28]. The idea was to use the rotationally invariant polarization entangled pair of photons in the state $\frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$. One photon is sent to Alice and the other to Bob. Alice and Bob perform Stern–Gerlach-type measurements on their photons. Alice analyzes the polarization of her photon in one of three randomly chosen analyzer orientations: 0°, 22.5° and 45°. Bob analyzes his photon in one of three randomly chosen analyzer orientations: 22.5°, 45° and 67.5°. After transmission of the key, Alice and Bob then publicly announce their analyzer orientations. The key is generated by keeping the coincidence detections that were measured with the same analyzer orientations (i.e., when Alice and Bob both have analyzer orientations 22.5° or 45°). The horizontally (vertically) polarized photon is counted as a 0(1), for example. Ideally, Alice and Bob would have a perfectly random anti correlated string of binary digits which could be used as a secret key. The security of the transmission is tested by performing a Bell's inequality test on the data taken when Alice and Bob have different analyzer orientations. If an eavesdropper Eve has tapped into the quantum channel, her lack of knowledge about the analyzer orientations introduces errors (destroys quantum correlations) in the channel. With the presence of an eavesdropper, Bell's inequalities will not be violated, which exposes her to the proper users.

7 Quantum Dense Coding

7.1 Theoretical Scheme

Quantum Dense Coding is an example of how entanglement can be used to minimize the number of carriers of information sent from Alice to Bob in order to transmit a certain message. If we consider single-photon transmission one can usually send one bit of information by utilizing the polarization degree of freedom. Note that although a photon is a spin-1 particle one can only explore the transversal polarization degree of freedom, which is two-dimensional and represents a spin-1/2. The longitudinal degree of freedom, in the direction of propagation, cannot be explored since the photon is moving with the maximum velocity allowed by the theory of relativity. The bit value 0 corresponds, for example, to the horizontal polarization ($|0\rangle$) of the photon and the bit value 1 corresponds to the vertical polarization ($|1\rangle$). Quantum dense coding, theoretically proposed by Bennett and Wiesner [29], enhances the information transfer from one bit to two bits. The catch is that entanglement between Alice and Bob must have been established prior to the communication, which requires the transfer of another photon. Although the entanglement itself does not contain any information about the message that Alice wishes to send to Bob, it does enable compression of data.

The obvious classical way to send two bits of information is to send two particles that each carry one bit of information. Identifying 00, 01, 10, and 11, with different information implies that we can encode two bits of information by manipulating *both* particles.

Quantum mechanics also allows one to encode the information in superpositions of the classical combinations. Such superpositions of states of two (or more) particles are the entangled states as described in previous sections. A convenient basis in which to represent such states for two particles, labelled 1 and 2, is formed by the maximally entangled Bell states

$$|\Psi^+\rangle_{12} = (|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2)/\sqrt{2}, \tag{53}$$

$$|\Psi^-\rangle_{12} = (|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2)/\sqrt{2}, \tag{54}$$

$$|\Phi^+\rangle_{12} = (|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2)/\sqrt{2}, \tag{55}$$

$$|\Phi^-\rangle_{12} = (|0\rangle_1|0\rangle_2 - |1\rangle_1|1\rangle_2)/\sqrt{2}\,. \tag{56}$$

Identifying each Bell state with different information we can again encode two bits of information, yet, now by manipulating only *one* of the two particles. This is achieved in the following quantum communication scheme. Initially, Alice and Bob each obtain one particle of an entangled pair, say, in the state $|\Psi^+\rangle_{12}$ given in (53). Bob then performs one out of four possible unitary transformations on his particle (particle 2) alone. The four such transformations are

1. Identity operation (not changing the original two-particle state $|\Psi^+\rangle_{12}$).
2. State exchange ($|0\rangle_2 \rightarrow |1\rangle_2$ and $|1\rangle_2 \rightarrow |0\rangle_2$, changing the two-particle state to $|\Phi^+\rangle_{12}$).
3 State-dependent phase shift (differing by π for $|0\rangle_2$ and $|1\rangle_2$ and transforming to $|\Psi^-\rangle_{12}$).
4. State exchange and phase shift together (giving the state $|\Phi^-\rangle_{12}$).

Since the four manipulations result in the four orthogonal Bell states, four distinguishable messages, i.e. 2 bits of information, can be sent via Bob's two-state particle to Alice, who finally reads the encoded information by determining the Bell state of the two-particle system. This scheme enhances the information capacity of the transmission channel to two bits compared to the classical maximum of one bit.

7.2 Experimental Dense Coding with Qubits

The setup of a quantum-optical demonstration of quantum dense coding, as described in [30], is illustrated in Fig. 7. The polarization entangled photon state is created by non collinear type-II PDC (see Sect. 3.4). Bob performs one out of four possible unitary transformations on one of the two particles before it is sent to Alice. The transformations are performed by using a $\lambda/2$ wave plate (rotated by 0° or 45° with respect to the optical axis of the down-conversion crystal) to change H_1V_2 and V_1H_2 terms into H_1H_2 and V_1V_2 and by using a $\lambda/4$ wave plate (rotated by 0° or by 90° with respect to the optical axis) in order to change the relative phase between the two terms. Alice performs the partial Bell-state measurement as described in Sect. 4.2. Using the 50:50 beam splitter (BS) followed by two polarizing beam splitters (PBS) and photon detectors that can distinguish between one and two-photon impact, Alice can

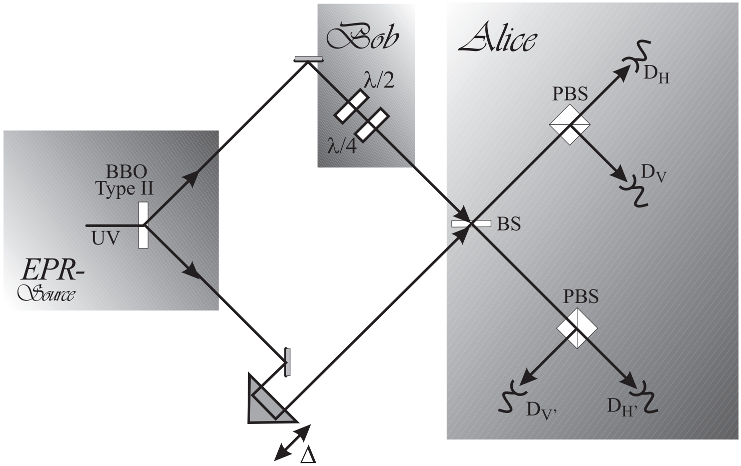


Fig. 7. Experimental setup for quantum dense coding [30].

uniquely identify $|\Psi^+\rangle_{12}$ and $|\Psi^-\rangle_{12}$ and can detect the $|\Phi^+\rangle_{12}$ and $|\Phi^-\rangle_{12}$ in a degenerate way. We note that in the actual experiment a cascade of single-photon detectors have been used (each unable to distinguish two-photon events from single-photon events) in order to have a probabilistic two-photon detector. Taking experimental limitations into account a trit of information (one out of three signals) can be transferred from Bob to Alice per photon sent from Bob to Alice.

8 Quantum Teleportation

8.1 Theoretical Scheme

In Sect. 5 we showed that perfect cloning of a quantum state is impossible, which implies that one cannot obtain full knowledge about a general quantum state by any measurement procedure. Consider now the situation that Alice has a system in an unknown quantum state, say the qubit $|\Psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1$, where $|0\rangle$ and $|1\rangle$ represent two orthogonal states with complex amplitudes α and β satisfying $|\alpha|^2 + |\beta|^2 = 1$. She wishes to transfer this quantum state to Bob but suppose she cannot deliver the particle directly to him. How can she provide Bob with the quantum state?

A solution of the dilemma is the quantum teleportation scheme as proposed by Bennett, Brassard, Crépeau, Jozsa, Peres, and Wootters [31]. The scheme is illustrated in the Fig. 8. The scheme uses an additional pair of entangled particles 2 and 3 (EPR pair), where particle 2 is given to Alice and particle 3 is given to Bob. Let us consider the case in which the entangled pair of particles 2 and 3 shared by Alice and Bob is in the state

$$|\Psi^-\rangle_{23} = \frac{1}{\sqrt{2}}(|0\rangle_2|1\rangle_3 - |1\rangle_2|0\rangle_3)\,. \tag{57}$$

Although initially particles 1 and 2 are not entangled, their joint polarization state can always be expressed as a superposition of the four maximally entangled Bell states, given by (53)–(56), since these states form a complete orthogonal

Table 1. Overview of possible manipulations and detection events of the quantum dense coding experiment with correlated photons.

Bob's setting			
$\lambda/2$	$\lambda/4$	State sent	Alice's registration events
0°	0°	$|\Psi^+\rangle$	coinc. between D_H and D_V or $D_{H'}$ and $D_{V'}$
0°	90°	$|\Psi^-\rangle$	coinc. between D_H and $D_{V'}$ or $D_{H'}$ and D_V
45°	0°	$|\Phi^+\rangle$	2 photons in either D_H, D_V, $D_{H'}$ or $D_{V'}$
45°	90°	$|\Phi^-\rangle$	2 photons in either D_H, D_V, $D_{H'}$ or $D_{V'}$

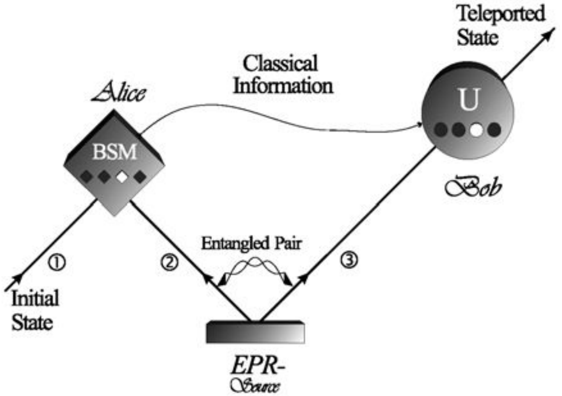


Fig. 8. Principle of quantum teleportation: Alice has a quantum system, particle 1, in an initial state which she wants to teleport to Bob. Alice and Bob also share an ancillary entangled pair of particles 2 and 3 emitted by an Einstein–Podolsky–Rosen (EPR) source. Alice then performs a joint Bell state measurement (BSM) on the initial particle and one of the ancillaries, projecting them also onto an entangled state. After she has sent the result of her measurement as classical information to Bob, he can perform a unitary transformation (U) on the other ancillary particle resulting in it being in the state of the original particle. In the case of quantum teleportation of a qubit, Alice makes a projection measurement onto four orthogonal entangled states (the Bell states) that form a complete basis. Sending the outcome of her measurement, i.e. two bits of classical information, to Bob will enable Bob to reconstruct the initial qubit.

basis. The total state of the 3 particles can be written as:

$$\begin{aligned}|\Psi\rangle_{123} = |\Psi\rangle_1 \otimes |\Psi\rangle_{23} = \frac{1}{2} \; [\; & |\Psi^-\rangle_{12} \, (-\alpha|0\rangle_3 - \beta|1\rangle_3) \\ & + |\Psi^+\rangle_{12} \, (-\alpha|0\rangle_3 + \beta|1\rangle_3) \\ & + |\Phi^-\rangle_{12} \, (\alpha|1\rangle_3 + \beta|0\rangle_3) \\ & + |\Phi^+\rangle_{12} \, (\alpha|1\rangle_3 - \beta|0\rangle_3) \,] \, . \end{aligned} \tag{58}$$

Alice now performs a Bell state measurement (BSM) on particles 1 and 2, that is, she projects her two particles onto one of the four Bell states. As a result of the measurement Bob's particle will be found in a state that is directly related to the initial state. For example, if the result of Alice's Bell state measurement is $|\Phi^-\rangle_{12}$ then particle 3 in the hands of Bob is in the state $\alpha|1\rangle_3 + \beta|0\rangle_3$. All that Alice has to do is to inform Bob via a classical communication channel on her measurement result and Bob can perform the appropriate unitary transformation (U) on particle 3 in order to obtain the initial state of particle 1. This completes the teleportation protocol.

Note that, during the teleportation procedure, the values of α and β remain unknown and that the quantum state initially given to Alice has to be destroyed in order for Bob to obtain the state.

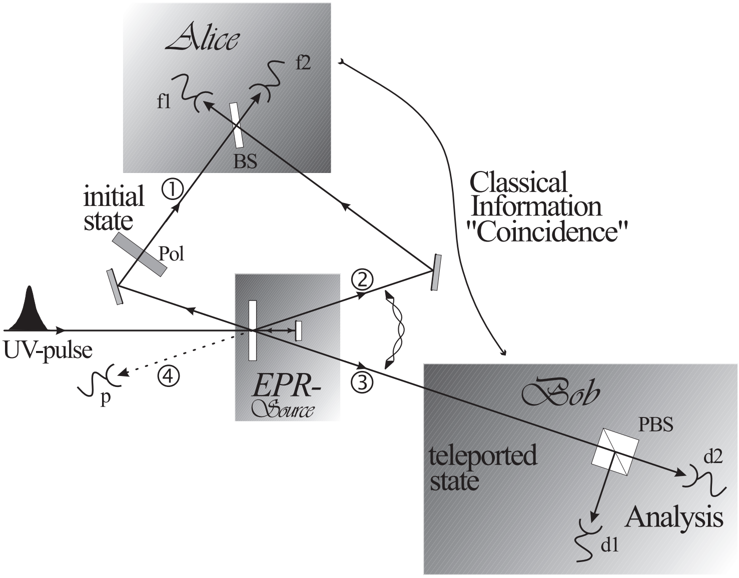


Fig. 9. Schematic drawing of the experimental setup for quantum teleportation of a qubit. A pulse of ultraviolet (UV) light passing through a nonlinear crystal creates the ancillary pair of entangled photons 2 and 3. After retroflection during its second passage through the crystal, the ultraviolet pulse can create another pair of photons, one of which will be prepared in the initial state of photon 1 to be teleported, the other one serving as a trigger indicating that a photon to be teleported is underway. Alice then looks for coincidences after a beam splitter (BS) where the initial photon and one of the ancillaries are superposed. Bob, after receiving the classical information that Alice obtained a coincidence count in detectors f1 and f2 identifying the $|\Psi^-\rangle_{12}$ Bell-state, knows that his photon 3 is in the initial state of photon 1 which he then can check using polarization analysis with the polarizing beam splitter (PBS) and the detectors d1 and d2. The detector P provides the information that photon 1 is underway.

8.2 Experimental Quantum Teleportation of Qubits

The setup for an experimental demonstration of quantum teleportation of qubits [32], encoded in the polarization state of single photons is shown in Fig. 9. It makes use the source of polarization entangled photons described in Sect. 3.4 and the Bell-state analyzer described in Sect. 4.2.

The experimental realization of the quantum teleportation of a qubit presented was restricted to use the $|\Psi^-\rangle_{12}$ Bell-state projection only.[1] The unitary transformation that Bob has to perform when Alice measures photon 1 and 2 in $|\Psi^-\rangle_{12}$ is simply the identity transformation, i.e. Bob should detect a photon in the same state as photon 1.

To avoid photons 1 and 2, which are created independently, being distinguished by their arrival times at the detectors, which would eliminate the possi-

[1] It is possible to extend the Bell-state analyser into an analyser that can uniquely identify both the $|\Psi^-\rangle_{12}$ state and the $|\Psi^+\rangle_{12}$ (see Sect. 4.2).

bility of performing the Bell-state measurement, the following technique is used. Photon 2, together with its entangled partner photon 3, is produced by pulsed parametric down-conversion. The pump pulse, generated by a frequency-doubled mode-locked titanium–sapphire laser, is 200 fs long. The pulse is reflected back through the crystal (see Fig. 9) to create a second pair of photons, photons 1 and 4. Photon 4 is used as a trigger to indicate the presence of photon 1. Photons 1 and 2 are now located within 200 fs long pulses, which can be tuned by a variable delay such that maximal spatial overlap of the photons at the detectors is obtained. However, this does not yet guarantee indistinguishability upon detection since the entangled down-converted photons typically have a coherence length corresponding to about a 50 fs long wavepacket, which is shorter than the pulses from the pump laser. Therefore, coincidence detection of photons 1 and 2 with their partners 3 and 4 with a time resolution better than 50 fs could identify which photons were created together. To achieve indistinguishability upon detection, the photon wave packets should be stretched to a length substantially longer than that of the pump pulse. In the experiment this was done by placing 4 nm narrow interference filters in front of the detectors. These filter out photon wave packets with a time duration of the order of 500 fs, which yields a maximum indistinguishability of photons 1 and 2 of about 85% [33]. To verify that quantum teleportation works for any arbitrary polarization state it is sufficient to demonstrate that the scheme works for three orthogonal states on the polarization (Poincaré) sphere. We refer to [32] for experimental data for such a demonstration using linear polarized light H and 45°, and circular polarized light.

8.3 Teleportation of Entanglement

Instead of using the fourth photon in the experiment described above as a mere trigger to indicate that photon 1 is underway, one can explore the fact that photon 1 and 4 can also be produced in an entangled state, say in the $|\Psi^-\rangle_{14}$ state. The state of photon 1 is therefore completely undetermined and all the information is stored in joint properties of photons 1 and 4. As remarked by Bennett et al. [31], if photon 1 is now subjected to quantum teleportation, photon 3 obtains the properties of photon 1 and therefore becomes entangled with photon 4. Interestingly, photon 4 and photon 3 originate from different sources and never interacted directly with one another, yet they form an entangled pair after the quantum teleportation procedure. An experimental verification of this process of transferring entanglement can be found in [34], and is known as entanglement swapping. A discussion on possible applications can be found in [3,35,36].

8.4 A Two-Particle Scheme for Quantum Teleportation

A limitation of the experimental teleportation scheme discussed above is that Alice could not perform a full Bell-state measurement, which reduced the efficiency of the quantum state teleportation. A full Bell-state measurement would imply a controlled interaction between two photons, which is extremely difficult

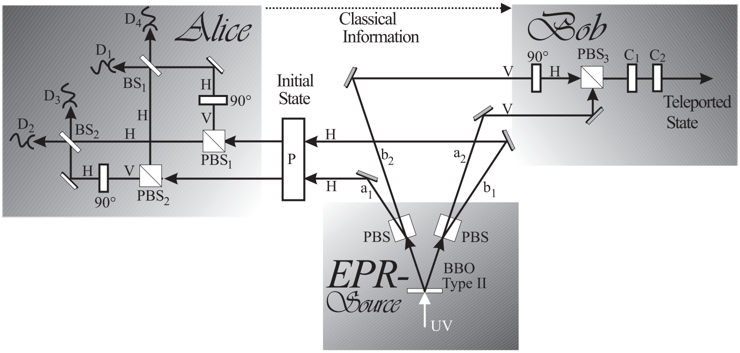


Fig. 10. Experimental scheme for the two-particle protocol for quantum teleportation. The setup consists of a type-II (BBO) down-conversion source for polarization entangled photons, polarizing beam splitters (PBS), 50:50 beam splitters (PB), single-photon detectors (D), 90° polarization rotation plates, the Preparer (P) of the initial quantum state, and polarization transformers (C).

to implement with a reasonable efficiency. S. Popescu proposed an optical scheme that avoids this problem but does place restrictions on the quantum states that can be transferred [37]. An experimental demonstration can be found in [38].

The scheme involves only two particles of which both the polarization and momentum degrees of freedom are explored. The first step is to produce two photons entangled in their direction of propagation, i.e. entangled in momentum, but each with a well-defined polarization. The box representing the EPR source in Fig. 10 shows how this can be achieved [38]. Using type-II parametric down-conversion, one first creates the polarization entangled state

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|H\rangle_1|V\rangle_2 + |V\rangle_1|H\rangle_2\right) , \tag{59}$$

where 1 and 2 label the two output directions of the correlated photons. Following this, both photons are passed through polarizing beam splitters which deflect/transmit horizontal/vertical photons. This transfers the polarization entanglement into momentum entanglement resulting in the state

$$\frac{1}{\sqrt{2}}\left(|a_1\rangle|a_2\rangle + |b_1\rangle|b_2\rangle\right)|H\rangle_1|V\rangle_2 \, . \tag{60}$$

Labels 1 and 2 now indicate the double channels that lead to Alice and Bob respectively. Photons with label 1 are necessarily H polarized and photons with label 2 necessarily V polarized.

On the way to Alice photon 1 is intercepted by the Preparer P who changes the polarization from H to an arbitrary quantum superposition

$$|\Psi\rangle_1 = \alpha|H\rangle_1 + \beta|V\rangle_1 \, . \tag{61}$$

The Preparer affects the polarization in both paths a_1 and b_1 in the same way. The state $|\Psi\rangle_1$ is the quantum state that Alice wants to transmit to Bob. The total state $|\Phi\rangle$ of the two photons after the preparation is

$$|\Phi\rangle = \frac{1}{\sqrt{2}} \left(|a_1\rangle|a_2\rangle + |b_1\rangle|b_2\rangle\right) |\Psi\rangle_1 |V\rangle_2 \,, \tag{62}$$

which is the formal analog of the state $|\Psi\rangle_{123}$ in (58).

The next step in the protocol is that Alice performs a joint (Bell-state) measurement on the initial state $|\Psi\rangle_1$ and on her part of the momentum entangled state. Assuming that there is a way to project photon 1 onto the four Bell states for its polarization and momentum, we obtain the equivalent of (58):

$$\begin{aligned} |\Phi\rangle = \frac{1}{2} \; [\; & (|a_1\rangle|V\rangle_1 + |b_1\rangle|H\rangle_1)(\beta|a_2\rangle + \alpha|b_2\rangle)|H\rangle_2 \\ & + (|a_1\rangle|V\rangle_1 - |b_1\rangle|H\rangle_1)(\alpha|a_2\rangle + \beta|b_2\rangle)|H\rangle_2 \\ & + (|a_1\rangle|H\rangle_1 + |b_1\rangle|V\rangle_1)(\alpha|a_2\rangle - \beta|b_2\rangle)|H\rangle_2 \\ & + (|a_1\rangle|H\rangle_1 - |b_1\rangle|V\rangle_1)(\beta|a_2\rangle - \alpha|b_2\rangle)|H\rangle_2 \;] \, . \end{aligned} \tag{63}$$

The first part of each term corresponds to a Bell state for photon 1 and the second part to the corresponding state of photon 2. For the projection of particle 1 onto the polarization/momentum Bell states we have to entangle the polarization and directional properties of photon 1. This can be done by using polarizing beam splitters in paths a_1 and b_1, and by combining the V component coming from a_1 ($|a_1\rangle|V\rangle_1$) with the H component coming from b_1 ($|b_1\rangle|H\rangle_1$), and vice versa. The combination, sensitive to the relative phase, is obtained by rotating the photons to the same polarization and letting them interfere on a normal beamsplitter. A photon detection by D_1, D_2, D_3, or D_4 now corresponds directly to a projection onto one of the four Bell states.

The final step of the protocol is that Alice informs Bob which detector registered a photon. With this information Bob can reproduce the initial polarization state as follows. He first transforms the momentum superposition of photon 2 (see (63)) into the same superposition in polarization by simply using a 90° rotation plate in paths b_2 (or a_2) and a polarizing beamsplitter to combine the paths. After this, he just switches two optical elements on or off, depending on the information obtained from Alice, to interchange H and V and to provide a relative phase shift of π between H and V. This transforms the polarization state of photon 2 into the polarization state prepared on photon 1, and thus completes the transmission.

An advantage of the scheme is that it uses a full Bell-state measurement. A drawback of the scheme is that it does not allow Alice to teleport the state of an outside particle. Therefore it requires the Preparer's help: the initial polarization state given to Alice has to be prepared on a particle which is momentum entangled with the one given to Bob. Also the state $|\Psi\rangle$ has to be pure, implying that it cannot be part of an entangled state. We refer to [38] for details about the experimental realization of the scheme described above.

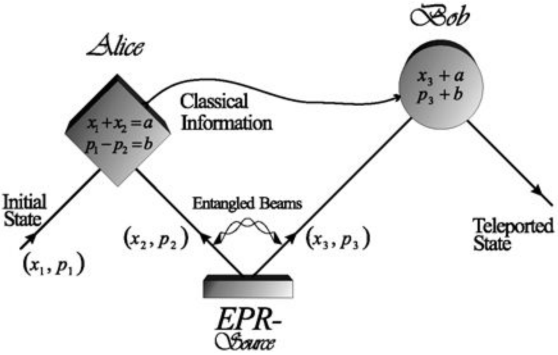


Fig. 11. Schematic drawing of quantum teleportation of continuous variables.

9 Teleportation of Continuous Quantum Variables

9.1 Theoretical Scheme

Quantum teleportation is not restricted to quantum states with a discrete set of basis states. L. Vaidman proposed a teleportation scheme for the position and momentum (which have a continuous set of basis states) of a quantum particle [39]. The scheme was further elaborated on by Braunstein and Kimble [40], and experimentally realised at Caltech [41].

Consider the situation that Alice has a quantum particle with a certain position x_1 and momentum p_1 (see Fig. 11), and she wishes to send this quantum information to Bob who is at a distant location. Due to the Heisenberg uncertainty relation between x and p, Alice cannot measure both x_1 and p_1 with arbitrary precision. The way out of this dilemma is conceptually the same as the protocol described in Sect. 8. An auxiliary pair of particles, entangled in their position and momentum, produced by the EPR source in Fig. 11, has to be distributed between Alice and Bob. Let us consider the case in which the entanglement of particles 2 and 3 is described by the conditions:

$$x_2 + x_3 = 0 \text{ and } p_2 - p_3 = 0\,. \tag{64}$$

The properties of the individual particles, x_2, x_3, p_2, and p_3 are completely undetermined by (64). Instead, their joint properties are defined. Note that, although the operators $\hat{x}$ and $\hat{p}$ do not commute for each particle, the operators for $(x_2 + x_3)$ and $(p_2 - p_3)$ do commute. Therefore, for the entangled state the joint properties, (x_2+x_3) and (p_2-p_3), can both be measured with an arbitrary accuracy.

Next Alice performs the equivalent of a Bell-state measurement on particles 1 and 2. The measurement by Alice yields

$$x_1 + x_2 = a \text{ and } p_1 - p_2 = b\,, \tag{65}$$

where a and b are two real numbers which both have a continuous range of possible values. This indicates that the measurement of the sum of positions and the difference in momenta of the two particles requires the projection onto an ∞-dim Hilbert space.

As a result of the initial entanglement (64) and of Alice's measurement (65), the information obtained about the quantum state in the hands of Bob is

$$x_3 = x_1 - a \text{ and } p_3 = p_1 - b \,. \tag{66}$$

To complete the quantum teleportation protocol, all Alice has to do is to send Bob via a classical channel the results of her measurements, i.e. the measured values a and b, and then Bob just displaces the position and momentum of his particle by a and b, respectively. The final result is that Bob has particle 3 in the initial quantum state of particle 1.

9.2 Quantum Optical Implementation

The experimental implementation of quantum teleportation of continuous quantum variables has been performed at Caltech, California [41]. This implementation does not use the position x and momentum p of particles but uses light beams that can be characterized by parameters obeying the same commutation relations as $\hat{x}$ and $\hat{p}$. The analogy is based on the fact that a single (transversal) mode of the quantized radiation field can be characterized by a quantum harmonic oscillator [42,43].

The classical harmonic oscillator of mass m, frequency ω, displacement x, and momentum p is described by the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m}{2}\omega^2 x^2 \,. \tag{67}$$

To obtain the quantum-mechanical Hamiltonian, x and p should be interpreted as operators ($x \rightarrow \hat{x}$, and $p \rightarrow \hat{p} = \mathrm{i}\hbar\partial/\partial x$) which obey the commutation relation $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$. If we define

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}^\dagger + \hat{a}\right) \,, \tag{68}$$

$$\hat{p} = \mathrm{i}\sqrt{\frac{\hbar m\omega}{2}}\left(\hat{a}^\dagger - \hat{a}\right) \,, \tag{69}$$

then the Hamiltonian for the quantized harmonic oscillator takes the natural form

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \,, \tag{70}$$

where $\hat{a}$ and $\hat{a}^\dagger$ are interpreted as the annihilation and creation operators.

A single transversal mode (frequency ω) of the quantized radiation field can be expressed in terms of the operators $\hat{a}$ and $\hat{a}^\dagger$. In its most basic form, i.e. including all prefactors into a single constant E_0 and considering one polarization direction, the electric field vector operator at a fixed position is given by

$$\hat{\underline{E}}(t) = E_0\left(\hat{a}e^{-\mathrm{i}\omega t} + \hat{a}^\dagger e^{+\mathrm{i}\omega t}\right) \,, \tag{71}$$

where $\hat{a}^\dagger$ and $\hat{a}$ are now interpreted as the photon-creation and photon-annihilation operators. In analogy to the harmonic oscillator, we can define operators $\hat{X}$ and $\hat{P}$ via

$$\hat{X} = \left(\hat{a}^\dagger + \hat{a}\right) , \tag{72}$$

$$\hat{P} = \mathrm{i}\left(\hat{a}^\dagger - \hat{a}\right) . \tag{73}$$

The electric field operator can now be expressed in terms of $\hat{X}$ and $\hat{P}$ as

$$\hat{\hat{E}}(t) = E_0 \left(\hat{X}\cos(\omega t) + \hat{P}\sin(\omega t)\right) . \tag{74}$$

The eigenvalues of $\hat{X}$ and $\hat{P}$, referred to as the quadrature field amplitudes, can be interpreted as the amplitudes of the in- and out-of-phase components of the electric field (with respect to a local oscillator). From the commutation relation $[\hat{X}, \hat{P}] = 2\mathrm{i}$ it follows that $\varDelta X \varDelta P = 1$ ($\langle \varDelta A\rangle^2 = \langle A^2\rangle - \langle A\rangle^2$), which means that the in- and out-of-phase amplitudes cannot be simultaneously measured with arbitrary accuracy, in close analogy to the position x and momentum p of a quantum particle. Hence we have now established the mapping of x and p for a particle to X and P for a single-mode light field.

X and P fulfill the uncertainty relation $\varDelta X \varDelta P \geq 1$. It is possible using nonlinear crystals (inside an optical parametric oscillator OPO) to produce squeezed light fields [44] for which for example $(\varDelta Y)^2 < 1$ and necessarily $(\varDelta X)^2 > 1$. In order to construct entangled light fields consider the case that two light fields, field $\mathcal{A}$ maximally squeezed in X and field $\mathcal{B}$ maximally squeezed Y, enter the two input ports of a 50:50 beam splitter (see Fig. 12). After the beam splitter the fields labelled with 2 and 3 are characterized by the relations

$$X_2 + X_3 = 0 \text{ ,and } P_2 - P_3 = 0 , \tag{75}$$

which specify precisely the desired entangled state [48,46,47,41]. (For polarization entangled light fields see [45,46].)

We now turn to the problem of performing a Bell-state-like measurement. Mixing the initial beam, characterized by (X_1, P_1) which are the continuous quantum variables to be teleported, with one beam coming from the EPR source, represented by (X_2, P_2), onto a 50:50 beamsplitter, yields in the two output ports beams characterized by

$$(X_\mathcal{C}, P_\mathcal{C}) = (X_1 - X_2, P_1 - P_2) , \text{ and } (X_\mathcal{D}, P_\mathcal{D}) = (X_1 + X_2, P_1 + P_2) . \tag{76}$$

Using the balanced homodyne detection method (see e.g. Ref. [48]), Alice can now measure the X component of beam $\mathcal{D}$ and the P component of $\mathcal{C}$, providing her with the values $a = X_1 + X_2$, and $b = P_1 - P_2$, respectively, as required for the quantum teleportation protocol. The balanced homodyne detection method is based on mixing of the signal field with a local oscillator on a 50:50 beam splitter and the recording of the difference in the photocurrent (proportional to the field intensity) between two detectors in the output arms of the beam

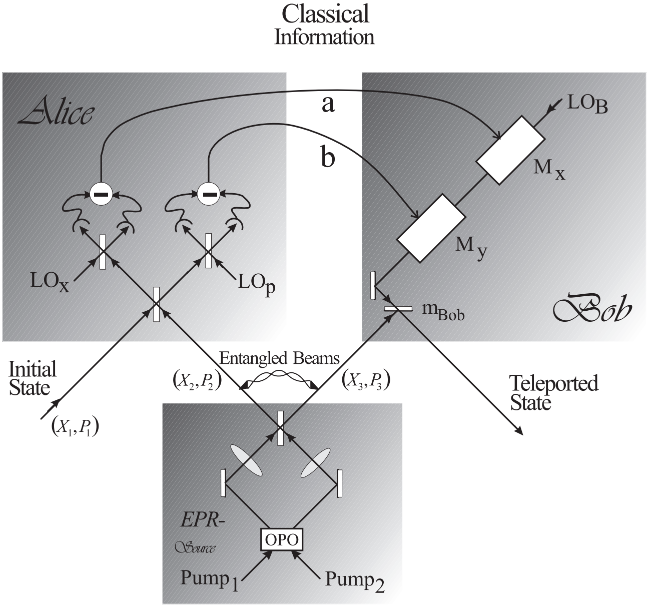


Fig. 12. Optical setup for continuous quantum variable quantum teleportation [41]. The EPR source consists of two light beams, one squeezed in X and the other in squeezed in Y, passing through a 50:50 BS. Alice performs homodyne detection on the input state (X_1, P_1), using local oscillators, LO_X and LO_Y, in order to measure $a = X_1 + X_2$, and $b = P_1 - P_2$. After Bob received the values of a and b via classical communication with Alice he can use modulators, M_X and M_P, to add the values a and b to field 3 which then transforms into field 1.

splitter. The difference in measured intensity as a function of the phase φ of the local oscillator is given by [48]

$$I(\varphi) = C\,(X \sin\varphi + P \cos\varphi)\ , \tag{77}$$

where C is an overall constant depending on the intensity of the local oscillator and on the properties of the detectors. Tuning the phase φ of the local oscillator, one can measure any superposition of the quadrature components.

Following the quantum teleportation scheme, Alice sends Bob the measured values a and b and Bob has to displace the light field at his side accordingly. Bob can achieve the displacement experimentally by reflecting his light field from a partially reflecting mirror (say 99% reflection and 1% transmission) and adding through the mirror a field that has been phase and amplitude modulated according to the values a and b. In principle, Bob ends up with an almost perfect replica of the light field that was initially in the hands of Alice. We refer to Ref. [41] for the experimental data obtained by the above scheme.

10 Quantum Error Detection and Correction

10.1 Introduction

The possibility to detect and correct errors in the evolution of a quantum system has been a most remarkable theoretical discovery [49,50,51,52,3]. This discovery, and the subsequent theoretical development of related ideas such as entanglement purification [53,54] the quantum repeater [55,56] and fault-tolerant quantum computation [57,58,59], turned the initial scepticism about implementing quantum computation and long distance quantum communication into optimism. In this section an optical scheme for the error-free transfer of quantum information through a noisy quantum channel is presented [60].

10.2 Quantum Error Detection

In order to explain the optical scheme we first point out the main ideas underlying classical and quantum error detection. A particularly simple classical error detection scheme uses the transmission of several copies of the bits to be transferred and requires that the probability of a bit-flip error during transmission is much smaller than unity. By comparing the copies of each initial bit after transmission one can determine the initial bits with high probability. Despite the fact that it is impossible to copy the state of an unknown quantum state, it is still possible to use a strategy similar to the classical one. Consider the state of a two-level system, a qubit, characterized by

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle \, .$$

In order to make comparison measurements after the state transmission, and thereby detect errors, we have to encode the initial qubit onto several particles. If we restrict our attention to the case that there is a small probability that a bit-flip error occurs, it is sufficient to encode the initial state onto the following three-particle entangled state:

$$|\Psi\rangle_{123} = \frac{1}{\sqrt{2}} \left(\alpha|000\rangle_{123} + \beta|111\rangle_{123}\right) \, . \tag{78}$$

The left-hand side of Fig. 13 indicates how this encoding is obtained using two controlled-NOT operations with the initial qubit as control qubit and two auxiliary particles initially prepared in state $|0\rangle$ as target qubits.

After transmission of the three-particle entangled state through a "noisy" quantum channel one can retrieve the initial qubit using the comparison measurements indicated on the right-hand side of Fig. 13. The measurements consists again of controlled-NOT operations, followed by detection of the two auxiliary particles in the $|0\rangle$, $|1\rangle$ basis. The detection acts as a parity check between the two particles on which the controlled-NOT operation acts: a $|0\rangle$ outcome indicates that in each term of the entangled state the two particles are the same, i.e. 00 or 11, a $|1\rangle$ outcome indicates that they are opposite, i.e. 01 or 10. If no error

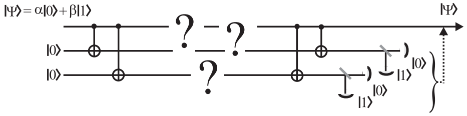


Fig. 13. Traditional scheme for the detection and correction of a bit-flip error. Using two controlled-NOT operations, an initial quantum state $|\Psi\rangle$ (the control qubit) is entangled with two auxiliary particles (the target qubits), each initially prepared in the state $|0\rangle$. After transmission of the three-particle entangled state through an area in which an error might occur, indicated by the question marks, each of the two auxiliary particles becomes the target particle of a controlled-NOT operation with the initial particle as the control particle. A final projection measurement on each of the two auxiliary particles onto the $|0\rangle$, $|1\rangle$ basis uniquely identifies a possible (single) error which can then be corrected.

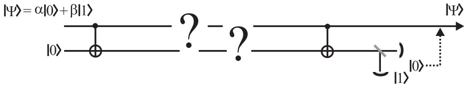


Fig. 14. Scheme for bit-flip error rejection. One auxiliary particle is sufficient in order to detect an error, without revealing on which particle the error occurred.

occurred during the transmission, both auxiliary particles should be detected in the state $|0\rangle$. However, if a bit-flip error occurred for the initial particle, and not for the other two, both auxiliary particles will be detected in the state $|1\rangle$. In the case that an error occurred on one of the auxiliary particles, and not on the remaining two, the corresponding particle will be detected in state $|1\rangle$ and the remaining auxiliary particle in state $|0\rangle$. After identification, a possible error can be corrected.

Crucial for the error detection/correction scheme is the fact that the parity-check measurements project the transmitted entangled state onto only four possible outcomes, namely no error, or one error on one of the three particles. Therefore, although during the transmission through the noisy quantum channel any qubit-rotation error can occur, the final state is quantized to contain either a full bit-flip error or no error.

If more than one error occurred the error-correction scheme is not useful. Therefore, it is crucial that the probability for an error on each particle is much smaller than unity ($P_{\text{error}} \ll 1$). Under this condition it is reasonable to consider for optical quantum communication purposes a simplified scheme that rejects transmissions that contain an error instead of identifying a specific error and correcting for it. Such a simplified scheme requires only one auxiliary particle as shown in Fig. 14. If the parity check measurement yields the $|0\rangle$ result, no error took place, or, with the very small probability P^2_{error}, a fatal double-error took place. If the measurement yields the $|1\rangle$ result, a single error occurred for one of the two particles and the transmission is invalidated.

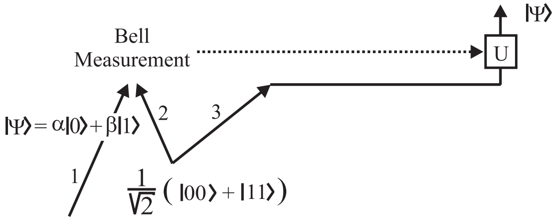


Fig. 15. Schematic drawing of the quantum teleportation protocol. The transmission of the unknown quantum state $|\Psi\rangle$ of particle 1 is broken down into the distribution of an auxiliary pair of entangled particles (2 and 3), a Bell-state measurement on particles 1 and 2 (*i.e.*, a projection onto a complete basis of maximally entangled particles), and the transfer of classical information (the outcome of the Bell-state measurement). After receiving the classical information, the relation of the state of particle 3 to the initial state $|\Psi\rangle$ is fully determined. The initial state can therefore be recovered by a well-defined unitary transformation U on particle 3.

10.3 Avoiding Controlled-NOT Operations

To present our error-free optical quantum communication scheme we note that the controlled-NOT (CNOT) operation in the preparation step of the schemes shown in Figs. 13 and 14 is used in order to encode an *arbitrary* initial quantum state onto a multi-particle entangled state. It is, however, not necessary to be able to encode an arbitrary input state. According to the teleportation scheme [31], illustrated in Fig. 15, the transmission of an arbitrary quantum state can be decomposed into the transmission of a *known* entangled state, a local Bell-state measurement and the transmission of classical information. Therefore, in order to establish error-free quantum communication, it is sufficient to be able to exclude erroneous transmission of one of the particles of a fixed entangled state.

Consider a pair of entangled photons in the state

$$|\Psi\rangle_{23} = \frac{1}{\sqrt{2}} \left(|0\rangle_2|0\rangle_3 + |1\rangle_2|1\rangle_3\right) . \tag{79}$$

To be able to detect errors on the transmission of, say, photon 2, the preparation scheme shown on the left-hand side of Fig. 14 would produce the state

$$|\Psi\rangle_{234} = \frac{1}{\sqrt{2}} \left(|0\rangle_2|00\rangle_{34} + |1\rangle_2|11\rangle_{34}\right) . \tag{80}$$

Since state (80) is a well-defined state, the use of the controlled-NOT operation is no longer necessary, as shown in the left-hand side of Fig. 16.

The right-hand side of Fig. 16 illustrates how the controlled-NOT operation for parity checking can also be avoided by using a polarizing beam splitter and a coincidence detection measurement in an appropriate basis. If a bit-flip error occurred for one of the two transmitted photons, both photons will exit the polarizing beam splitter in the same output arm. Therefore no coincidence will be observed between the detectors in arm a and b, and the transmission will be invalidated.

If no error occurred the state after the polarizing beam splitter will have one photon in each output arm, indicating that the two outgoing photons have the same polarization relation as when initially prepared, *i.e.*, the polarizations are parallel in each term of the entangled state. The detection scheme proceeds by detecting the particle in arm b in the basis

$$|0'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,, \quad |1'\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,. \tag{81}$$

The specific measurement outcome corresponds to a projection of the remaining particles, provided a particle is present in arm a, onto one out of two well-defined pure two-particle entangled state:

$$|0'\rangle_b \rightarrow \frac{1}{\sqrt{2}}\left(|0\rangle_2|0\rangle_a + |1\rangle_2|1\rangle_a\right) \tag{82}$$

$$|1'\rangle_b \rightarrow \frac{1}{\sqrt{2}}\left(|0\rangle_2|0\rangle_a - |1\rangle_2|1\rangle_a\right)\,. \tag{83}$$

The teleportation procedure can now be completed by a Bell-state measurement and the transfer of classical information as illustrated in Fig. 16.

10.4 Post-selection

One might be alarmed by the fact that the photon in arm a has still to be detected in order to complete the error-free transmission scheme. This will in

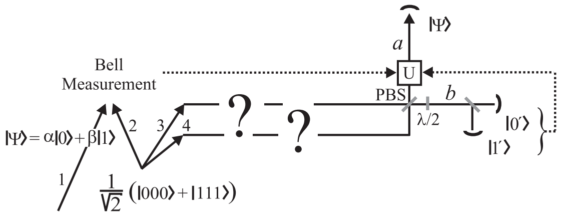


Fig. 16. Scheme for error-free quantum-state transmission without controlled-NOT operations. In order to transfer a quantum state it is sufficient to restrict the use of a quantum channel for the transmission of one of an entangled pair of particles (see Fig. 15). In order to reject erroneous transmissions, a three-particle entangled state is used. Two of the three entangled particles are sent through the "noisy" quantum channel. A parity check measurement on particles 2 and 3 identifies an error-free transmission and is obtained by using a polarizing beam splitter (PBS) followed by a coincidence detection of one particle in arm a and the other in arm b. The measurement in arm b must be such that the remaining two particles are projected onto a well-defined two-particle entangled state. This is achieved by performing the measurement in the linear basis rotated 45° with respect to the $|0\rangle$, $|1\rangle$ basis. After the result of the measurement on the particle in arm b is known, the remaining particles (one to be detected in arm a and particle 1) are guaranteed to be in a well-defined entangled state and can be used for error–free quantum teleportation or quantum cryptography.

practical applications imply the destruction of the photon, although absorption-free detection of single-photons has experimentally been demonstrated [61]. The anxiety to lose the photon before being able to use it is unjustified, at least for applications in quantum cryptography and other quantum communication protocols, since the detection of the photon is an integral part of all such applications. In fact, any realistic single-photon communication scheme needs a final verification step to guarantee that the fragile photon survived the transmission. The detection of the photon, therefore, plays the double role of enabling a projection onto a pure entangled state for photon 2 and the photon in arm a, as well as exploring this entanglement for quantum cryptography or for quantum communication purposes.

11 Stimulated Entanglement

11.1 Theory

Stimulated emission of radiation was one of the great breakthroughs of physics in the mid 20th century. Initially the applications of devices that used stimulated emission (e.g., MASER, LASER) were for basic research such as spectroscopy. However, laser's are now used in everything from the grocery checkout and fibre communication, to medical treatments, to nuclear fusion experiments. While, it is true that that the initial use of the acronym LASER meant visible Light Amplification via Stimulated Emission of Radiation, the term laser now has a broader interpretation of amplification via stimulated emission of any part of the electromagnetic spectrum. In fact, with the discovery that atoms of a Bose–Einstein Condensate (for a review see for example [63]) could be coherently and directionally discharged from the condensate, the term atom laser has even found its way into the literature. In this section, a new type of laser-like operation based on the stimulated emission of entangled photons will be discussed.

A conventional laser requires an inverted gain medium enclosed in a optical resonator. Numerous gain media have been developed over several decades, but all have basically the same function. When a gain medium is inverted the vast majority of quantum particles (e.g. atoms)in the medium, are in an excited quantum state. The quantum particles can spontaneously decay to a lower energy level and in the process emit a photon with no preferred direction. These spontaneously emitted photons impinge on the other excited quantum particles in the gain medium. In doing so, the spontaneously emitted photon stimulates the emission of additional photons. Unlike the spontaneously emitted photon, the stimulated photons are not randomly oriented, but are in the same direction and in phase with the original spontaneously emitted photon. Assuming no losses, an exponential gain in the number of photons occurs. This action is short lived though, because the gain medium is small and there is still no preferred direction. It is therefore important to make a preferred direction using a optical resonator. The resonator thus plays a crucial role of effectively lengthening the gain medium in a single direction. An exponential gain (until losses and saturation of the gain medium become significant) occurs along the mode defined by the cavity axis while a depletion of the other off-axis modes occurs.

Preferred feedback enhancement is not the only purpose of the highly reflecting cavity. The cavity also sets the boundary conditions. For highly reflecting mirrors, standing waves are set up with nodes at the mirrors. Thus, only frequencies of photons that satisfy these boundary conditions will experience gain in the medium. One of the mirrors has a lower reflectance than the other allowing the laser field to "leak" out. Using these techniques, one can obtain a well-collimated, narrow line-width source of photons.

It will be fruitful to discuss a little of the mathematics associated with laser operation. A laser generates a coherent state, which can be obtained from the single-mode Hamiltonian

$$\hat{H} = -\mathrm{i}\alpha\hat{a}^{\dagger} + h.c. \tag{84}$$

where $\hat{a}^{\dagger}$ is the photon creation operator, α is coupling constant and is a property of the gain medium, and $h.c.$ denotes Hermitian conjugate. The unitary evolution operator is given by $\hat{U} = e^{-\mathrm{i}\frac{\hat{H}t}{\hbar}}$. Acting on the vacuum state the output is an exponential function of the interaction time t. By reflecting the field off of a mirror back into the gain medium, the interaction time t is increased. If one expands the unitary operator acting on the vacuum in a series ($\hbar \equiv 1$) gives

$$\hat{U} = \sum_{N=0}^{\infty} \frac{(\alpha t\hat{a}^{\dagger})^N}{N!} \,. \tag{85}$$

Stimulated emission in a mathematical sense is to operate via a creation operator on a non-vacuum Fock state. Thus, terms of $(\hat{a}^{\dagger})^N$ with $N > 1$ can be considered stimulated emission.

The unique and important difference between an entangled photon laser as proposed in [64] and a standard laser is that the entangled photons are always created in pairs. The pair-wise amplification of the 2 and 4 photon states will be discussed and experimental results will be shown. This novel entanglement structure holds great promise in quantum information science where there is a strong demand for entangled states of increasing complexity.

Polarization entangled laser operation means that a (spontaneously created) photon pair in two polarization entangled modes stimulate, inside a non linear gain medium, the emission of additional pairs. The gain medium, a nonlinear beta barium borate crystal, is cut for type-II phase matching. A pump pulse enters the crystal and small probability exists that photons in the pump will down-convert or split into polarization entangled photons. A simplified interaction Hamiltonian for the nonlinear interaction between a classical pump field and two polarization-entangled modes a and b is given by

$$\hat{H}_{\text{int}} = e^{i\phi}\kappa\hat{K}^{\dagger} + e^{-i\phi}\kappa\hat{K} \,, \tag{86}$$

where $\hat{K}^{\dagger} \equiv (\hat{a}_h^{\dagger}\hat{b}_v^{\dagger} - \hat{a}_v^{\dagger}\hat{b}_h^{\dagger})$ and $\hat{K} \equiv (\hat{a}_h\hat{b}_v - \hat{a}_v\hat{b}_h)$ are the creation and annihilation operators of polarization entangled photon pairs in modes a and b. Horizontal and vertical polarization are represented by H and V, and κ is a real-valued coupling coefficient. As a note, the operator $\hat{K}^\dagger \equiv (\hat{a}_h^\dagger \hat{b}_v^\dagger - \hat{a}_v^\dagger \hat{b}_h^\dagger)$ when acting on the vacuum state yields the Ψ^- Bell state. The analogy between a standard laser and an entangled photon laser is now more apparent. The standard laser builds up the coherent state from an infinite superposition of polynomials of the single photon creation operator. An entangled photon laser builds up an entangled pair distribution from an infinite superposition of polynomials of the Ψ^- Bell state operator.

When acting on the vacuum state, the time evolution operator $\hat{U}{=}\exp(i\hat{H}t/\hbar)$ yields [65]

$$\hat{U}(\tau)|0\rangle = e^{-q} \sum_{n=0}^{\infty} \frac{r^n}{\sqrt{n!(n-1)!}} \times \left(\sum_{m=0}^{n} (-1)^m |(n-m), m; m, (n-m)\rangle \right) , \tag{87}$$

where $\tau \equiv \frac{\kappa t}{\hbar}$, $r \equiv \tanh \tau$, and $q \equiv 2 \ln(\cosh \tau)$. We used the shorthand notation $|i, j; k, l\rangle$ for $|i\rangle_{aH} |j\rangle_{aV} |k\rangle_{bH} |l\rangle_{bV}$ and $|0\rangle$ represents $|0, 0; 0, 0\rangle$. The first and second positions in the ket indicate the number of horizontal and vertical photons in mode a, respectively, and the third and fourth slot indicate the corresponding numbers for mode b (e.g., for $n = 2$ and $m = 0$, one obtains the quantum state $|2, 0; 0, 2\rangle$, which has two horizontally polarized photons in mode a and two vertically polarized photons in mode b). This state represents the general output of type-II parametric down-conversion, but for all experiments reported to date, τ is so small that mainly the first order term ($n = 1$) has been taken into account and only a few experiments and proposals addressed second order terms. Similar to a conventional laser, the idea of an entangled photon laser is to increase τ (the effective interaction length) using a resonator around the gain medium, which enhances the emission of the higher order terms in eqn. (87).

Interestingly, the quantum state described in (87) bears various noteworthy features. First of all, modes a and b are entangled in photon number since for any n the number of photons in each mode is identical. Several photon pair number distributions are shown in Fig. 17 for various levels of τ or average pair number. The shifting of the maximum and the broadening of the distribution for higher values of average pair number resembles the coherent state photon-number distribution as produced by conventional lasers. The one primary difference in the broadening of the distribution is that in a standard laser the distribution broadens as $\sqrt{n}$ while the entangled pair distribution broadens as n. These features are explicable upon recognizing that stimulated emission –originating from the boson statistics of photons– favors amplification of higher over lower photon-number terms.

The second important property of state (87) is that the set of terms for each n form a maximally entangled state in polarization. The normalised 1-pair term is the rotationally-symmetric Bell state (singlet spin-1/2):

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|1, 0; 0, 1\rangle - |0, 1; 1, 0\rangle) \tag{88}$$

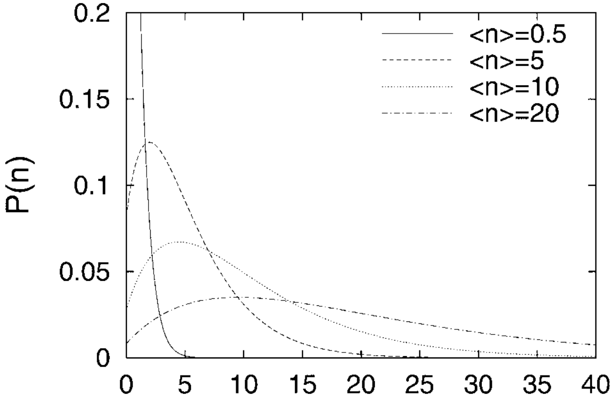


Fig. 17. The photon number (pair) distribution, $P(n)$, arising from stimulated parametric down-conversion shifts its peak and broadens as the mean number of photons increases. This indicates that for increasing interaction strength (gain), terms with higher numbers of photons obtain a larger amplification factor compared to lower terms, which is a familiar feature of laser operation.

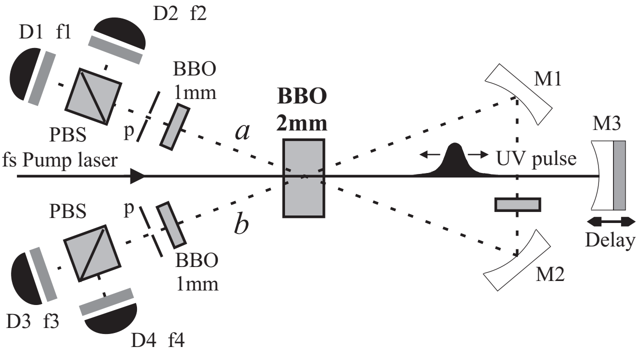


Fig. 18. Experimental setup. A frequency-doubled mode-locked Ti–Sapp laser (80 MHz repetition rate, $\lambda = 390$ nm) pumps a 2 mm BBO crystal. Pinholes p perform spatial selection of the entangled modes. The pump is reflected onto itself by mirror M3 that is mounted on a computer-controlled translation stage. Mirrors M1 and M2 form the feedback loop, including a polarization rotation element ($\lambda/2$), for the entangled photons. Photon detection of the $|1,1;1,1\rangle$ term in the H/V basis occurs at avalanche photo diodes D1–D4, after going through polarizing beam splitters (PBS) and 5-nm-bandwidth filters f1–f4. The role of the two extra 1-mm BBO crystals in modes a and b is to compensate for undesirable birefringent properties of the main crystal [11].

which is the well known Ψ^- Bell state. The normalized 2-pair term is given by :

$$|\Psi\rangle = \frac{1}{\sqrt{3}}(|2,0;0,2\rangle - |1,1;1,1\rangle + |0,2;2,0\rangle) \tag{89}$$

and represents the singlet spin-1 state. Similar to the spin-1/2 case, its rotational symmetry arises from the relative phase relations and the equal weights of the terms. In general, the n-pair term has the properties of a singlet spin-$n/2$ state.

The crucial role of stimulated emission is to provide for each n, equally weighted terms. In principle, a photon counting measurement on state (87) (either in mode a or b) performs a projection onto a certain singlet spin-$n/2$ state. Subsequently, this maximally entangled state can be explored for quantum information tasks. In practice, in quantum optics experiments where the fragile photons are typically destroyed by any measurement, the projection and the exploration of the state are performed simultaneously. This procedure, usually referred to as post-selection, has proven to be most useful, e.g. for demonstrations of quantum teleportation, quantum cryptography, and three particle GHZ correlations, and for novel optical quantum computation schemes [66]. Here, post-selection is used to demonstrate stimulated entanglement by measuring 2- and 4-photon properties of state (87) for increasing values of τ.

The setup used to demonstrate stimulated entanglement is illustrated in Fig. 18. A short pump pulse at 390 nm passes through a β–barium–borate (BBO) crystal and creates pairs of polarization entangled photons in spatially distinct modes a and b. The experimental parameters are chosen such that (to first order) the $|\Psi\rangle^-$ Bell state is created. Initially modes a and b are in the vacuum state and the photon pairs are spontaneously created. The fact that modes a and b geometrically diverge and that horizontally and vertically polarized photons experience different crystal parameters limits the useful crystal length and thereby prohibits an efficient stimulated emission process. To obtain significant stimulated emission the spontaneously created photon pairs are redirected into the crystal at the same time (tuned by a delay on mirror M3) as the reflected pump pulse passes through the crystal a second time. Thus, the mirrors, similar to a standard laser, set the boundary conditions and effectively lengthen the gain medium. Provided that the feedback loop for the photon pairs is polarization independent, which is obtained by using a bow-tie folded geometry including a waveplate that exchanges H and V polarizations, optimum conditions for stimulated emission of photon pairs can be established.

Stimulated emission can be seen as a constructive multi particle interference effect and is therefore sensitive to the phase of the pump. Hence, one expects to observe an oscillation between stimulation and suppression of emission as function of the pump-pulse delay (this is shown in Fig. 20). The period of this oscillation corresponds to the optical frequency of the pump laser. Clearly, in the region where the difference between the pump delay and the feedback loop is larger than the coherence length of the observed photons (determined by the 5 nm narrow-bandwidth filters in front of the single-photon detectors), no such interference pattern is expected.

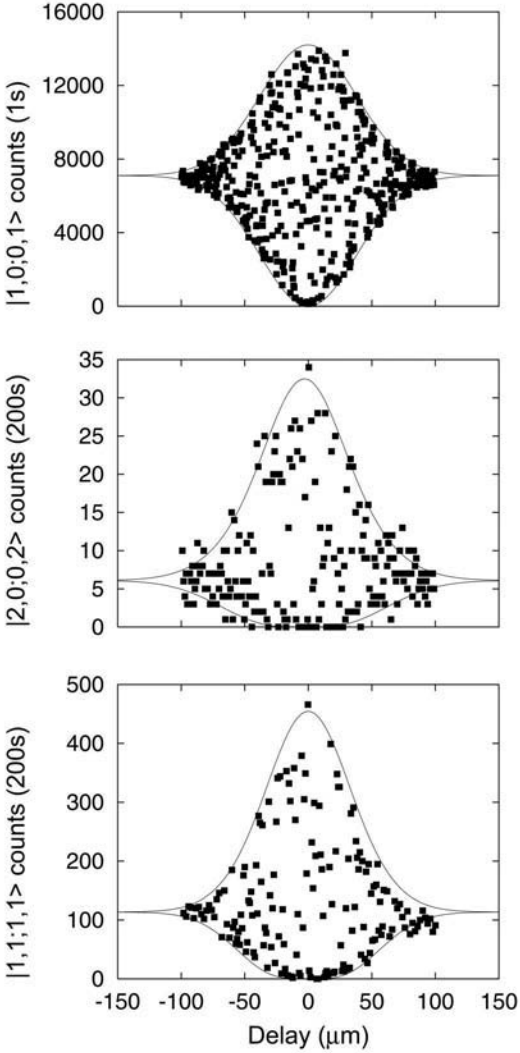


Fig. 19. Experimental demonstration of stimulated entanglement. The top graph shows the 2-fold coincidence rate corresponding to the detection of the $|1,0;0,1\rangle$ term in the $45°/-45°$ basis as function of the delay between the reflected pump and the entangled photons generated in the first pass through the crystal. The solid lines are theoretical fits to the envelope of the curve as the degree of overlap varies. Similarly the middle graph shows the 4-fold coincidence rate corresponding to the detection of the $|2,0;0,2\rangle$ term in the $45°/-45°$ basis and the bottom graph the $|1,1;1,1\rangle$ in the H/V basis. The effect of stimulated emission is apparent in the increase of the number of 4-fold coincidences at zero delay of a factor of 5.3 and 4.0 for the $|2,0;0,2\rangle$ and $|1,1;1,1\rangle$ terms (see text). The difference in rates between the two 4-photon graphs is due to the probabilistic detection and extra elements introduced to measure the $|2,0;0,2\rangle$ term.

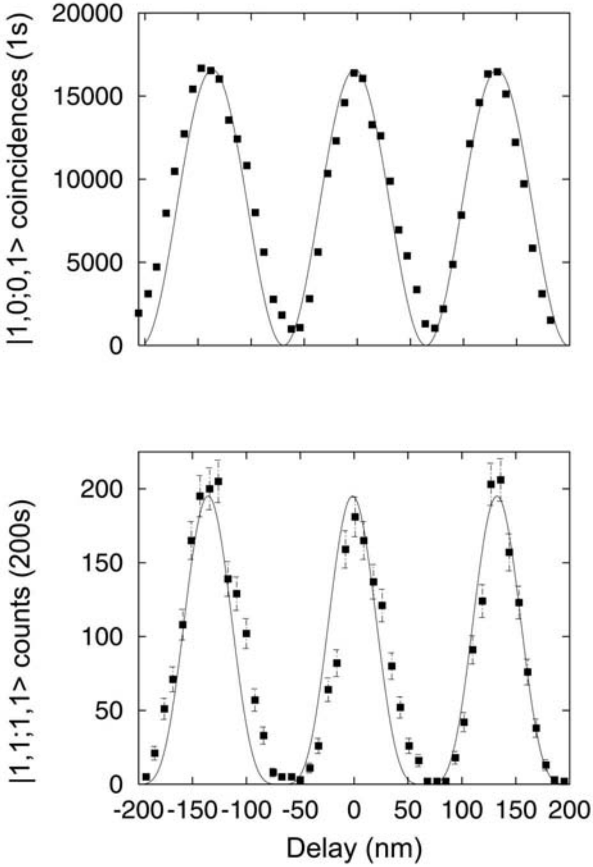


Fig. 20. Two- and four-photon interference due to stimulated emission. A fine scan of the $|1,0;0,1\rangle$ (top) and $|1,1;1,1\rangle$ (bottom) terms in the zero delay region shows optimum stimulation and suppression of the 2- and 4-fold coincidence probability.

To study the 2- and 4-photon entangled states we measure each term in (88) and (89) individually in two non orthogonal polarization bases. Quantitative predictions for the amplification of the individual terms in (88) and (89) can be obtained by expanding the double pass unitary evolution to second order

$$\hat{U} = \hat{U}_2\hat{U}_1 = 1 + e^{\mathrm{i}\theta}\tau\hat{K}_2^\dagger + \tau\hat{K}_1^\dagger + \frac{1}{2}e^{2\mathrm{i}\theta}\tau^2(\hat{K}_2^\dagger)^2 + \frac{1}{2}\tau^2(\hat{K}_1^\dagger)^2 + e^{\mathrm{i}\theta}\tau^2\hat{K}_1^\dagger\hat{K}_2^\dagger \quad (90)$$

where subscripts 1 and 2 refer to the first and second pass through the crystal. The relative phase θ between the first and second pass of the pump pulse is tunable via translation of mirror M3 in Fig. 18.

Two limits of interference require elucidation; namely, when $\hat{K}_2^\dagger = \hat{K}_1^\dagger$ and when $\hat{K}_2^\dagger$ is distinguishable from $\hat{K}_1^\dagger$. The first case applies at zero delay where efficient phase-sensitive stimulated emission occurs. From (refeq:stim) it follows that doubling the value of the interaction parameter results in an increase in probability for the 2-photon terms from τ^2 to $4\tau^2$ and in an increase for the 4-photon terms from τ^4 to $16\tau^4$. Note that the 4-photon state has a four times larger amplification than the 2-photon states, which is characteristic of stimulated emission. The second case, $\hat{K}_2^\dagger$ is distinguishable from $\hat{K}_1^\dagger$, occurs when the difference between the roundtrip distance of the reflected pump field and the down conversion field is larger than the single photon coherence length. In this case, there are simply two independent contributions to the 2-photon detection events but there are several distinct contributions to the 4-photon detection events. Each single pass has a small probability of τ^4 to create state (89). In addition, since current single photon detectors do not have a high enough time resolution to distinguish between photons arriving from the first or second pass, there are spurious 4-fold coincidences from a combination of 2-photon states created in both passes. The spurious contributions to the $|2,0;0,2\rangle$ and the $|0,2;2,0\rangle$ detections will be τ^4 and for the detection $|1,1;1,1\rangle$ it will be $2\tau^4$.

The transition between the two cases $\hat{K}_2^\dagger = \hat{K}_1^\dagger$ and $\hat{K}_2^\dagger \neq \hat{K}_1^\dagger$ is achieved by scanning the region where $d > \lambda_c$ to $d = 0$, while observing the intensity of the 2- and 4-photon terms. From the considerations above, one expects in the case of ideal stimulated emission that the terms in (88) show a two-fold increase and that the middle term in (89) shows a four-fold increase and the other two terms increase by a factor of 16/3=5.33. Due to the rotational symmetry of state (88) and (89), these predictions are basis independent.

Figure 19 shows the experimental data for the detection of the $|1,0;0,1\rangle$ (top) and the $|2,0;0,2\rangle$ (middle) terms measured in the 45° rotated basis, and the (bottom) term in the H/V basis. The solid curves are the envelopes of the oscillating functions giving the maximum and minimum theoretical values for the coincidence rates. The experimental data shows an increase of 1.95 ± 0.10 for $|1,0;0,1\rangle$, 5.3±0.6 for $|2,0;0,2\rangle$ and of 4.1±0.3 for $|1,1;1,1\rangle$. These results are in good agreement with the predictions discussed above. Similar results have been obtained in the other bases and for the $|0,1;1,0\rangle$ and $|0,2;2,0\rangle$ terms, demonstrating the rotational invariance, i.e. the spin-1/2 and spin-1 singlet structure, of states (88) and (89). Additional data indicates an amplification due to the second pass of 3.95 ± 0.10 for the 2-fold and of 17 ± 2 for the 4-fold coincidences. This demonstrates the shifting of the photon-number pair distribution towards terms with higher photon numbers, a characteristic of stimulated emission. A final proof of stimulated emission –seen as a constructive interference process– is the phase dependent emission probability shown in Fig. 20. This is a fine scan around the region of zero delay for the $|1,0;0,1\rangle$ (top) and the $|1,1;1,1\rangle$ (bottom) terms in the 45° rotated basis and H/V basis respectively. The solid lines are fits to the theoretical predictions, which vary as $1 + \cos(\theta)$ for the 2-photon case and as $(1+\cos(\theta))^2$ for the 4-photon case. The visibility of these interference fringes is in all cases above 97%.

12 Bohm-Type Spin-s Entanglements

In Sect. 3.4. we described Bell's inequality for two entangled spin-1/2 particles. In the early 90's Gisin and Peres showed that entangled particles with arbitrarily large spins still violated a Bell inequality [67]. This result implied that large quantum numbers are no guarantee of classical behavior. Apart from fundamental interest [67,68,69], entangled states of spin-s objects are also of clear interest for applications in quantum information due to the higher dimensional Hilbert space associated to these states (e.g. quantum cryptography, dense coding and bound entanglement [70]). In the previous section we showed that the polarization entangled four-photon fields (2-photons in each of two spatial modes) of pulsed parametric down-conversion are formally equivalent to two maximally entangled spin-1 particles [64,71]. Using postselection, it is possible to selectively measure the four-photon state. The rotationally invariant four-photon state is given by

$$\frac{1}{\sqrt{3}}\left(|2H,2V\rangle - |HV,VH\rangle + |2V,2H\rangle\right) \tag{91}$$

where the first term in the kets represent the polarization of the photons sent to Alice and the second term in the kets represent the polarization of the photons sent to Bob. For example, the $|2H,2V\rangle$ means that if Alice measures two horizontal photons, then Bob will measure two vertical photons. The photons sent to Alice have three possible measurement outcomes with equal probabilities, namely $|2H\rangle$, $|HV\rangle$ and $|2V\rangle$, which we will define as the $|1\rangle$, $|0\rangle$ and $|-1\rangle$ state respectively. Thus, it is *not* the photons that are the spin-1 particles, but the two-photon polarization entangled modes. These types of entanglements were theoretically realized by Drummond [72] when he described cooperative emission of wave packets containing N-bosons, proving that multiparticle states could violate Bell's inequalities. The connection between states produced in parametric down-conversion and the N-boson multiparticle states was also discussed by Reid *et al.* [73].

One can define a spin-1 local hidden variable measurement combination as

$$E^{HV}(a,b) \equiv \int \mathrm{d}\lambda f(\lambda)\overline{A}(a,\lambda)\overline{B}(b,\lambda), \tag{92}$$

where

$$\overline{A}(a,\lambda) = P(1|a,\lambda) - P(0|a,\lambda) + P(-1|a,\lambda), \tag{93}$$

$$\overline{B}(b,\lambda) = P(1|b,\lambda) - P(0|b,\lambda) + P(-1|b,\lambda)\,. \tag{94}$$

Because the signs (the value assignment) of the probabilities in both $\overline{A}(a,\lambda)$ and $\overline{B}(b,\lambda)$ are different, consequently $|\overline{A}(a,\lambda)| \leq 1$ and $|\overline{B}(b,\lambda)| \leq 1$. At this point it should be noted that there is a certain amount of freedom in the definitions $\overline{A}$ and $\overline{B}$ as well in the type of measurements performed on the two spin-1 particles. Here we consider Stern–Gerlach type of measurements and a value assignment of only $+1$ and -1 (rather than complex values). These choices do not maximally profit from the rotational symmetry of the spin-1 singlet state and therefore do not lead to the maximum violation possible. To optimize the Bell's inequalities for spin-s is not an easy task and only very recently has this issue been addressed in the literature [75,76,69].

The derivation of our modest spin-1 Bell inequality proceeds exactly as the spin-1/2 formalism [7,74], leading to

$$S = |E(a,b) - E(a,b') + E(a',b) + E(a',b')| \leq 2. \tag{95}$$

Hence, the maximum possible value that can be achieved, assuming locally explicable outcomes is 2. On the other hand, quantum mechanical measurement probabilities on both Alice's and Bob's side cannot be decoupled, which implies

$$\begin{aligned} E^{QM}(a,b) = {} & P(1,1|a,b) - P(1,0|a,b) \\ & + P(1,-1|a,b) - P(0,1|a,b) + P(0,0|a,b) \\ & - P(0,-1|a,b) + P(-1,1|a,b) \\ & - P(-1,0|a,b) + P(-1,-1|a,b). \end{aligned} \tag{96}$$

Using the Bell inequality in (95) we obtain a theoretical maximum violation of 2.552, which is in agreement with [67]. This prediction was obtained using analyzer rotations of $a = 0^o$, $a' = 22.5^o$, $b = 11.25^o$, and $b' = 33.75^o$. Experimentally a value of 2.27 ± 0.02 was achieved [71] which clearly violates the spin-1 Bell's inequality (95). The discrepancy between the quantum mechanical prediction and the experimental results can be understood from the imperfections in the set up for creating spin-1 singlet states. For details see [71].

Acknowledgements

We are very grateful to H. Howell for her support in preparing the manuscript and to B. Salcher for drawing Figures 1 and 2. This work was supported by the EPSRC GR/M88976 and the European QuComm (ISI-1999-10033) projects.

References

1. *Maxwell's Demon*, Eds. H.S. Leff and A.F. Rex, (Adam Hilger, Bristol 1990).
2. J. Preskill, *Quantum Information and Computation*, lecture notes, http://www.theory.caltech.edu/people/preskill/ph229.
3. *The Physics of Quantum Information*, Eds. D. Bouwmeester, A. Ekert, A. Zeilinger, Springer–Verlag, Berlin Heidelberg New York (2000).
4. A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
5. D. Bohm *Quantum Theory*, (Prentice–Hall, Engle–wood Cliffs, N.J. 1951).
6. J.S. Bell, Physics **1**, 195 (1964).
7. J.F. Clauser, M.A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett **23**, 880 (1969).
8. J.F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).
9. M.A. Rowe *et al.*, Nature **409**, 791 (2001).

10. A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981); Phys. Rev. Lett. **49**, 91 (1982); A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
11. P.G. Kwiat, *et al.* Phys. Rev. Lett. **75**, 4337 (1995).
12. G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger Phys. Rev. Lett. **81**, 5039 (1998).
13. W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, N. Phys. Rev. Lett. **81**, 3563 (1998).
14. P.G. Kwiat, A.M. Steinberg, and R.Y. Chiao, Phys. Rev. A **47**, R2472 (1993). PDC
15. C.K. Hong, Z.Y. Ou and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
16. J.D. Franson, Phys. Rev. Lett. **62**, 2205, (1989).
17. P.G. Kwiat, A.M. Steinberg, and R.Y. Chiao, Phys. Rev. A **47**, R2472 (1993).
18. P.G. Kwiat, *et al.* Phys. Rev. A **60**, R773 (1999). SLPITTER CLONING
19. W.K. Wootters and W.H. Zurek, Nature **299**, 802, (1982).
20. D. Dieks, Phys. Lett. A, **92**, 271 (1982).
21. L. Mandel, Nature **304**, 188 (1983).
22. P.W. Milonni and M.L. Hardies, Phys. Lett. A, **92**, 321 (1982).
23. V. Buzek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).
24. N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
25. C. Simon, G. Weihs and A. Zeilinger, Phys. Rev. Lett. **84**, 2993 (2000). CRYPTOGRAPHY
26. P. Shor in *Proc. of 35th Annual Symposium on the Foundations of Computer Science*, (IEEE Computer Society, Los Alamitos, 1994).
27. C.H. Bennett and G. Brassard, in *Proc. IEEE Int. Conference on Computers, Signals and Signal Processing*, IEEE, New York (1984).
28. A. Ekert, Phys. Rev. Lett. **67**, 661 (1991). CODING
29. C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett., **69**, 2881 (1992).
30. K. Mattle, H. Weinfurter, P.G. Kwiat, and A. Zeilinger Phys. Rev. Lett. **76**, 4656 (1996). TELEPORTATION
31. C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
32. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eible, H. Weinfurter, and A. Zeilinger, Nature **390**, 575 (1997).
33. M. Zukowski, A. Zeilinger and H. Weinfurter in *Fundamental Problems in Quantum Theory*, vol. 755 Annals of the New York Academy of Sciences, Eds. Greenberger and Zeilinger (1995).
34. J.-W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **80**, 3891 (1998).
35. M. Zukowski, A. Zeilinger, M.A. Horne, and A. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).
36. S. Bose, V. Vedral, and P.L. Knight, Phys. Rev. A **57**, 822 (1998).
37. S. Popescu, LANL, http://xxx.lanl.gov/quant-ph/9501020.
38. D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, Phys. Rev. Lett. **80**, 1121 (1998).
39. L. Vaidman, Phys. Rev. A **49**, 1473 (1994).

40. S.L. Braunstein and H.J. Kimble, Phys. Rev. Lett. **80**, 869 (1998).
41. A. Furusawa, J.L. Sørensen, S.L. Braunstein, C.A. Fuchs, H.J. Kimble, and E.S. Polzik, Science, **282**, 706 (1998).
42. R. Loudon, *The Quantum Theory of Light, 2nd edn*, (Clarendon Press, Oxford, 1983).
43. P.W. Milonni, *The Quantum Vacuum*, (Academic Press, San Diego, 1994).
44. D.F. Walls and G.J. Milburn, *Quantum Optics, second edition.* Springer, Berlin, Heidelberg (1994).
45. Z.Y. Ou, S.F. Pereira, H.J. Kimble, and K.C. Peng, Rhys. Rev. Lett. **68**, 3663 (1992).
46. H.J. Kimble, in *Fundamental Systems in Quantum Optics, Les Houches, 1990*, eds. J. Dalibard, J.M. Raimond, J. Zinn-Justin, (Elsevier Science Publishers, Amsterdam, 1992).
47. Ling-An Wu, H.J. Kimble, J.L. Hall, and Huifa Wu, Phys. Rev. Lett. **57**, 2520 (1986).
48. A. Yariv, *Quantum Electronics, third edition.*, (John Wiley& Sons, 1989).
CORRECTION
49. A. Steane, Phys. Rev. Lett. **77**, 793 (1995).
50. A. Steane, Proc. R. Soc. Lond. A, **452**, 2551 (1995).
51. P.W. Shor, Phys. Rev. A, **52**, R2493 (1995).
52. A.R. Calderbank and P.W. Shor, Phys. Rev. A, **54**, 1098 (1996).
53. C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
54. D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).
55. H.-J. Briegel, W. Dür, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 5932 (1998).
56. W. Dür, H.-J. Briegel, J.I. Cirac, and P. Zoller, Phys. Rev. A **59**, 169 (1999); ibid. **60**, 729 (1999).
57. P.W. Shor, *Proc. 37th Symp. on Foundations of Computer Science*, Los Alamitos, CA: IEEE Computer Society Press, (1996).
58. J. Preskill, Proc. Roy. Soc. Lond. A, **454**, 469 (1998).
59. A.M. Steane, Nature, **399**, 124 (1999).
60. D. Bouwmeester, Phys. Rev. A **63**, 040301 (2001).
61. G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J.M. Raimond and S. Haroche, Nature **400**, 239 (1999).
ENTANGLEMENT
62. J.R. Anglin and W. Ketterle, Nature **416**, 211 (2002).
63. M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman and E.A. Cornell, Science, **269**, 198 (1995).
64. A. Lamas-Linares, J. C. Howell and D. Bouwmeester, Nature **412**, 887 (2001).
65. P. Kok and S.L. Braunstein, Phys. Rev. A **61**, 042304-10 (2000).
66. E. Knill, R. Laflamme and G.J. Milburn, Nature **409**, 46 (2001).
INEQUALITIES
67. N. Gisin and A. Peres, Phys. Lett. A, **162**, 15 (1992).
68. B.S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980); A. Garg and N. D. Mermin, Phys Rev. Lett. **49**, 901 (1982); M. Ardehali, Phys. Rev. D **44**, 3336 (1991); K. Wodkiewicz, Acta. Phys. Pol. A **86** 223 (1994); D. Home, A.S. Majumdar, Phys. Rev. A **52**, 4959 (1995).
69. D. Kaszlikowski, P. Gnacinski, M. Zukowski, W. Miklaszewski, and A Zeilinger, Phys. Rev. Lett. **85**, 4418 (2000).

70. M. Horodecki, P. Horodecki and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
71. J.C. Howell, A. Lamas-Linares, and D. Bouwmeester, http://xxx.lanl.gov/quant-ph/0010356.
72. P. D. Drummond, Phys. Rev. Lett. **50**, 1407 (1983).
73. M.D. Reid, W.J. Munro and F. De Martini, http://xxx.lanl.gov/quant-ph/0104139 (2001).
74. J. S. Bell, *Speakable and unspeakable in quantum mechanics*, (Cambridge University Press, 1993).
75. D. Collins *et al.*, http://xxx.lanl.gov/quant- ph/0106024.
76. D. Kazslikowski, http://xxx.lanl.gov/quant-ph/0106010.

Quantum Information: Entanglement, Purification, Error Correction, and Quantum Optical Implementations

Juan Ignacio Cirac

1 Introduction

It is generally recognized that all the microscopic phenomena that we observe can be described and explained by the principles of Quantum Mechanics. These principles have been extensively tested, and some of them are commonly used in several technological applications. Other principles, like the ones related to the measurement process and the superposition principle, have only recently become important in some applications. In particular, they form the basis of what is called quantum communication and quantum computation. These two fields have been strongly developed during the last few years, and they may well give rise to a technological revolution in the fields of communication and computation [1].

The basic ideas behind quantum communication and computation are very simple (see the chapters by D. Bouwmeester and A. Ekert in this book). In the context of quantum communication, a sender (traditionally called Alice) tries to convey a secret message to a receiver (traditionally called Bob). The message (or more precisely, the random key) can be obtained by performing measurements on two quantum systems A and B owned by Alice and Bob, respectively, which are in some entangled state $|\Psi\rangle$. Due to the fact that the quantum state of a system is distorted when it is measured, if somebody tries to intercept the key, Alice and Bob will have a wrong state and will be able to detect this fact. This way of secret communication is usually called quantum cryptography, and it is the only provably secure way in which two partners can share secret messages. In the context of quantum computation, the existence of entangled states of several particles offers the possibility of performing certain computational tasks in times much shorter than the ones taken by common (classical) computers. By acting on a system entangled with other systems, one modifies the state of the whole system at the same time, which leads to an important speed up in obtaining the solution to several problems.

In reality, if one wants to build even a small quantum computer or to perform long distance quantum communication one will be confronted with the problem of decoherence (see the chapters by A. Leggett and W. Zurek in this book). Quantum systems are never perfectly isolated from the environment, which produces errors in the computation, as well as transforms the pure entangled states used in quantum cryptography into mixtures. There are two ways of dealing with this problem. The first one is error correction, which is a method that allows us

W. Dieter Heiss (Ed.): LNP 587, pp. 199–239, 2002.
© Springer-Verlag Berlin Heidelberg 2002

to recover the state of the quantum computer from these errors, provided they are small enough. The second one is the entanglement distillation (or purification) which allows us to distill almost pure entangled states out of mixed ones by just performing local operations (i.e. by Alice and Bob independently) and classical communication.

Entanglement plays an important role in most of the applications in the field of Quantum Information [1,2]. Thus, the characterization of this intriguing property of Quantum Mechanics is one of the central theoretical issues in Quantum Information. In fact, there are still many open questions regarding the entanglement properties of two or more quantum systems. Although for pure states of two systems, entanglement is well understood, for more systems we do not yet know how to quantify this property. The situation becomes much more complicated if the state of the system is mixed. In that case, we do not even know how to determine whether two systems are entangled or not.

On the other hand, there are very few systems in which one can implement a quantum computer. Many of the ideas in this respect come from the field of Quantum Optics. The reason is the spectacular experimental development of this field during the last years, which has allowed, sort to say, to dominate the quantum world. In particular, the internal quantum levels of atoms and ions can be manipulated very efficiently using lasers. Moreover, one can basically stop atoms using laser cooling techniques, and then manipulate their quantum state of motion by pushing them with laser light. These methods, when combined appropriately, allow, at least in principle, to perform quantum computations.

In these notes we briefly review some of the basic issues of Quantum Information Theory. First, in Section 2, after giving the conditions a physical system must fulfill if we want to use it to perform quantum computations, we explain the basic ideas of how to build a quantum computer using quantum optical systems. In particular, we will concentrate on the case of trapped ions, for which a specific method to build quantum gates will be given. In Section 3 we first review the process of decoherence and explain its consequences for quantum computing. Then we review the main ideas on which the error correction methods are based.

In the following sections we will study the problem of separability and distillability. First, in Section 4 we discuss the entanglement properties of two or more systems when they are in a pure state. Then, in Section 5, we review some of the problems related to the entanglement of mixed states, and in Section 6 we describe the process of entanglement distillation.

2 Quantum Computers: Quantum Optical Implementations

2.1 Introduction

This section is devoted to showing the basic ideas on which the quantum optical implementations for quantum computation are based. First, we will review the conditions required to build a quantum computer. Then, we will briefly mention some of the different set–ups that have been proposed to this aim, paying special attention to the quantum optical ones. Finally, we will discuss a specific method to perform quantum computations using trapped ions.

2.2 How to Construct a Quantum Computer

A quantum computer is composed of a set of qubits (quantum two-level systems) which can be manipulated in a controlled way. One should be able to initialize the state of the qubits (i.e. erase the state of the system) to some fixed state, for example $|0, 0, \ldots, 0\rangle$. One should also be able to read out the result after the computation. A computation corresponds to the evolution of the set of the qubits according to a specify unitary operator. A general operation of this sort can be decomposed into quantum gates. Those are quantum processes that transform the state of the qubits. For example, a general single–qubit gate has the form

$$|0\rangle \to \cos(\alpha)e^{\mathrm{i}\gamma/2}|0\rangle - \mathrm{i}e^{\mathrm{i}\phi}\sin(\alpha)|1\rangle, \tag{1a}$$

$$|1\rangle \to -\mathrm{i}e^{-\mathrm{i}\phi}\sin(\alpha)|0\rangle + \cos(\alpha)e^{-\mathrm{i}\gamma/2}|1\rangle. \tag{1b}$$

In particular, for $\alpha = \phi = \pi/2$ and $\gamma = 0$, we have the quantum version of the NOT gate. We can also have two-qubit gates, like the controlled-NOT gate,

$$|0,0\rangle \to |0,0\rangle, \tag{2a}$$

$$|0,1\rangle \to |0,0\rangle, \tag{2b}$$

$$|1,0\rangle \to |1,1\rangle, \tag{2c}$$

$$|1,1\rangle \to |1,0\rangle. \tag{2d}$$

It can be shown that any unitary operation acting on a set of qubits can be written as a sequence of controlled-NOT and single qubit gates (see the chapter by A. Ekert in this book). Note that the controlled-NOT gate requires interaction between the qubits (see Section 4.2). This can be shown, for example, noticing that the input product state $|0\rangle(|0\rangle + |1\rangle)$ is transformed into $|00\rangle + |11\rangle$, which is an entangled state.

Thus, in order to build a quantum computer one needs [3]:

- Qubits: A set of two-level systems.
- Quantum Gates: One has to be able to perform single qubit gates and the controlled-NOT gate (or an equivalent one).
- Erase: One has to be able to erase the state of the qubits. That is, to prepare the state $|0, 0, \ldots, 0\rangle$.
- Read out: One has to be able to perform local measurements on the qubits.

Apart from that, one needs to have isolation such that the operation is not spoiled due to the interaction with the environment (see next section). Moreover, the system has to be scalable such that one can, in principle, have as many qubits as possible interacting coherently without significantly increasing the error probability per gate (the increase cannot be exponential in the number of qubits).

2.3 Physical Implementations for Quantum Computation

For the moment, we know very few systems which fulfill the requirements to implement a quantum computer with them. Perhaps, the most important problem is related to the necessity of finding a quantum system which is sufficiently isolated, and for which the required controlled interactions can be produced. For the moment, there exist three kind of physical systems that fulfill, at least, most of the requirements:

- *Quantum optical systems [4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,2* Qubits are atoms, and the manipulation takes place with the help of a laser. These systems are very clean in the sense that with them it is possible to observe quantum phenomena very clearly. In fact, with them several groups have managed to prepare certain states which lead to phenomena that present certain analogies with the Schrödinger cat paradox, Zeno effect, etc. Moreover, those systems are currently used to create atomic clocks, and with them one can perform the most precise measurements that exist nowadays. For the moment, experimentalist have been able to perform certain quantum gates, and to entangle 3 or 4 atoms [19,12]. The most important difficulty with those systems is to scale up the models so that one can perform computations with many atoms.
- *Solid state system [25,26,27]:* There have been several important proposals to construct quantum computers using Cooper pairs or quantum dots as qubits (see the chapters by A. Leggett and D. Loss in this book). The largest difficulty in these proposals is to find the proper isolation of the system, since in a solid it seems hard to avoid interactions with other atoms, impurities, phonons, etc. For the moment, only single quantum gates have been experimentally reported. However, these systems possess the advantage that they are easily scalable.
- *Nuclear magnetic resonance systems [28,29]:* In this case the qubits are represented by atoms within the same molecule, and the manipulation takes place using the NMR technique. Initially, these systems seemed to be very promising for quantum computation, since it was thought that the cooling of the molecules was not required, which otherwise would make the experimental realization very difficult. However, it seems that without cooling, these systems loose all the advantages of quantum computation [30].

2.4 Quantum Optical Systems for Quantum Computation

In the following we will describe how to perform quantum computations with quantum optical systems. As we have mentioned before, the qubits are atoms, and the states $|0\rangle$ and $|1\rangle$ are two internal levels. In order to avoid spontaneous emission, those states must correspond to two stable electronic configurations. For example, in atoms with only one electron in the last shell, one can take two ground hyperfine levels with different magnetic numbers. In order to isolate them from the environment, one uses high-vacuum chambers, so that there are practically no other atoms or molecules that can collide with them. The initialization of the state is achieved using optical pumping, which consists of exciting the atom with a laser if it is in a different state than $|0\rangle$ (this is achieved by tuning appropriately the laser frequency and polarization). In this way, the atom will change the state via stimulated absorption and spontaneous emission until it decays in the state $|0\rangle$, in which the laser does not excite it anymore. The same method, with small modifications, can also be used to read out the state of each qubit. The idea is to us a laser in such a way that if the atom is in the state $|0\rangle$, then it does not absorb light; if it is in the state $|1\rangle$, then absorption–emission cycles occur such that the atom is excited by the laser and then it comes back to the state $|1\rangle$ via spontaneous emission. In this way, if at the end of the computation we switch the lasers on and observe light coming from some of the atoms, we will have measured the state $|1\rangle$ in that particular atom. If we do not observe light coming out, we will deduce that it is in the state $|0\rangle$. The single–qubit gates can be also carried out using a laser in such a way that it gives rise to stimulated absorption and emission (but not spontaneous emission, which would lead to decoherence). This is achieved by using two lasers whose frequencies are very far from resonance with respect to all atomic transitions. The absorption of one laser photon, followed by a stimulated emission of a photon in the other laser achieves the transition $|0\rangle \leftrightarrow |1\rangle$. By choosing appropriately the laser intensities and phases, one can carry out any arbitrary single–qubit operation. The controlled-NOT gate is usually the hardest part, since it requires the controlled interaction between the atoms. One way of achieving it is by manipulating the atoms in such a way that they exchange a photon (that is, one atom emits a photon and the other absorbs it). In order to do that, one needs either high-quality cavities [15,16,17,18,19,20], so that the photons emitted by the atoms always go to a single resonant cavity mode, and not in any other direction, or to use dipole–dipole interactions [22,23,24]. Another way consists of bringing the atoms together so that they interact (via a cold collision) [21]. Apart from that, one can also use the Coulomb repulsion between ions to perform this gate [4,5,6,7,8,9,10,11,12,13,14]. An example of this last method will be explained in the next section.

2.5 Quantum Computation with Trapped Ions

Ions confined in electric traps provide us with one of the most appropriate systems for quantum computation [31]. The ions can be easily trapped in a region of space in the following way. One heats up an oven filled with atoms (typically Be, Ca, Ba, Mg, In or Yb) in such a way that they leave the oven towards a region which contains some electric fields (trap). Since the atoms are neutral, they are not affected by those fields. However, if one targets the atoms with an electron beam, they may be ionized. As soon as this occurs, they start feeling the electric fields, which confine them in that region. Those fields are generated by some electrodes, whose parameters can be changed in such a way that the potential felt by the ions is harmonic, but in which the restoring forces along two directions (say x and y) are much stronger than in the other direction. In this way, and due to the Coulomb repulsion, the ions tend to align along the z axis. Once the ions are trapped, one can cool them (i.e. stop them) using laser light. The idea is to drive the atoms with a laser of frequency ω, which is quasi–resonant with some other ω_0 corresponding to a certain atomic transition. This happens in such a way that the atom absorbs photons from the laser and emits them spontaneously. Choosing appropriately the laser parameters ($\omega < \omega_0$), in each absorption–emission cycle the ions loose the energy $\hbar(\omega_0 - \omega)$, which is extracted from the ions motion. In this way one can achieve that the ions practically stop in space; well, in reality they end up in the ground state of the potential created by the trap and the Coulomb interaction. As a result we can have a set of ions, separated by a distance of the order of 15μm, which are basically stopped. As mentioned before, two internal states of each ion represent the qubit states $|0\rangle$ and $|1\rangle$. The single qubit gates can be performed as indicated in the previous subsection. There exist several methods to realize the controlled-NOT gate. Here we will mention one which is based on a conceptually simple effect [5].

Atoms, when they interact with light, apart from absorbing and emitting photons, feel pushed by the laser. In particular, if an atom absorbs a photon of energy $\hbar\omega$ and is transferred from some state $|g\rangle$ to some other $|e\rangle$, its momentum increases by $\hbar\boldsymbol{k}$, where $\boldsymbol{k}$ is the laser wave vector ($k = \omega/c$). On the contrary, if the atoms goes from $|e\rangle$ to $|g\rangle$, the momentum of the atom decreases by $\hbar\boldsymbol{k}$. If we have an atom practically stopped in the internal state $|g\rangle$ and we send a laser pulse propagating from left to right, then it will be transferred into the state $|e\rangle$ and will start moving to the right. Similarly, if it was in the state $|e\rangle$ then it will be transferred to $|g\rangle$ and will move to the left. Of course, if it is initially in the state $(|g\rangle + |e\rangle)/\sqrt{2}$ then the motional state will be a superposition of a state of the atom moving to the right and another moving to the left. In fact, this is the way in which some atomic interferometers operate, where the atomic wave function is split into two wave packets which are later on recombined to obtain an interference pattern.

Let us analyze how we can use this effect to produce a controlled-NOT gate between two ions. In order to simplify the argument, we will assume that we just have two ions (1 and 2) in the trap. After laser cooling, the first ion is located to the left of the other one. The quantum logic gate takes place in three steps:

- Pushing ion 1: Using a laser pulse, the first ion is pushed to the right or left depending on its internal state (for example, $|0\rangle = |g\rangle$ and $|1\rangle = |e\rangle$). Note that due to the Coulomb repulsion, if the first ion is pushed to the right (left) then the second one (pushed by the first ion) will be also moving to the right (left). That is, after this step the second ion will move to the right or to the left depending on the internal state of the first one.
- Transition in the second ion: Focalizing another laser beam to the right of the equilibrium position of the second ion one can change its internal state $|0\rangle \leftrightarrow |1\rangle$, but only provided the ion is there. Otherwise, if the ion was moving to the left, it will not be affected by this laser.
- Pushing ion 1 back: Due to the external electric potential, the ions will oscillate back to their original positions. At that moment, one can use the same laser

pulse as in the first step so that it is reversed. The ions will stop, and the first one will come back to its original internal state.

3 Decoherence and Error Correction

3.1 Introduction

Systems are never completely isolated. They interact with other degrees of freedom, what we call environment. The effects of these interactions are two fold: on the one hand, the system evolution is not the ideal one (nor even unitary); on the other hand, the state of the system becomes less pure, and loses the coherences responsible for interference phenomena (decoherence) and quantum parallelism (see the chapters by A. Legget and W. Zurek in this book). These effects are particularly inconvenient in quantum computation. The algorithms introduced so far are based on the fact that the state of the quantum computer is pure. Furthermore, small deviations from the ideal gates will produce wrong results in the calculation. Thus, the presence of decoherence eliminates all the advantages of quantum computation. In this section we will briefly review the process of decoherence, and then we will show how one can extend the classical methods of error correction to the quantum case in order to "undo" the effects of decoherence on a system.

3.2 Decoherence

Consider a two-level system that it is coupled to the environment. Let us denote by $|E\rangle$ the initial state of the environment. The interaction of the system with the environment is in all generality described by a unitary operator, which can be characterized as follows:

$$|0\rangle \otimes |E\rangle \to |0\rangle \otimes |E_{00}\rangle + |1\rangle \otimes |E_{01}\rangle, \tag{3a}$$

$$|1\rangle \otimes |E\rangle \to |0\rangle \otimes |E_{10}\rangle + |1\rangle \otimes |E_{11}\rangle, \tag{3b}$$

where $|E_{ij}\rangle$ are unnormalized states of the environment and their scalar products ensure that the evolution is unitary. Since we cannot measure all the degrees of freedom of the environment, all the information of the system will be in the reduced density operator defined after tracing over the environment degrees of freedom. Thus, the state of the system changes due to the coupling to the environment. In general, the state of the system will not be pure anymore. Consider the simple case where $|E_{01}\rangle = |E_{10}\rangle = 0$, and the initial state of the qubits is $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle$. We have for the reduced density operator

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1| + |0\rangle\langle 1|\langle E_{11}|E_{00}\rangle + |1\rangle\langle 0|\langle E_{00}|E_{11}\rangle). \tag{4}$$

If $\langle E_{11}|E_{00}\rangle = 0$, the coherences disappear in the density operator, and therefore the state becomes impure [the purity $\mathrm{Tr}(\rho^2)$ goes down to 1/2]. This is what in reality occurs: due to the interaction with the environment it may happen, for instance, that $\langle E_{11}|E_{00}\rangle \to e^{-\gamma t}$, so that after a time $\tau_c \simeq 1/\gamma$ the quantum behavior is lost. The time τ_c is called *decoherence time.*[1]

Consider N two-level systems, each of them interacting with its own environment. We denote by $|E^i\rangle$ the initial state of the environment for system i, and assume that each of the qubits interacts with an environment in the form (3a). For simplicity let us assume again that $|E^i_{01}\rangle = |E^i_{10}\rangle = 0$, and $\langle E^i_{11}|E^i_{00}\rangle \to e^{-\gamma t}$. Then, one can readily check that

$$\langle 0,0,\dots,0|\rho|1,1,\dots,1\rangle = \frac{1}{2^N} e^{-\gamma N t}. \tag{5}$$

If initially one has all qubits in the so-called GHZ state

$$\frac{1}{\sqrt{2}}(|0,0,\dots,0\rangle - |1,1,\dots,1\rangle), \tag{6}$$

then after a time $t \sim 1/(N\gamma)$, the state will have lost its purity (it will have "decohered"). Thus, we see that the coherences when one has N qubits in entangled states may decay N times faster than for a single qubit. This can be understood by the fact that if something happens to a single qubit, then the state may be significantly changed; the probability for this to happen is p^N, where p is the probability that something happens if we just have a single qubit.

3.3 Error Correction

In any computation (classical and quantum) or during storing of information there will be errors. One way to fight against these errors is to improve the hardware and make it better. However, this is expensive and not always possible. Shannon realized that instead of trying to avoid the errors it is much better to correct them. This is done by giving redundant information, and using this extra information to find out if an error occurred.

One can distinguish two kinds of errors:

- *Memory errors:* Those that occur to the information that is stored, regardless of whether an operation takes place or not.
- *Operation errors:* Those that occur during an operation.

Here we will concentrate on memory errors, since the corresponding correction procedures are easier to understand. On the other hand, they play an important role not only in quantum computing, but also in quantum communication and information. Once one knows how memory errors can be corrected, (with some modifications) one can understand how to correct operation errors. We will first revise the most straightforward way of correcting errors in a classical computer, and then we will show how to do it in a quantum computer.

[1] The term decoherence comes from a process which makes the coherences (non–diagonal elements of the density operator in a given basis) vanish. However, since this definition depends on the basis some authors prefer to call decoherence any process which is not describable by a unitary operator, i.e. which comes from the interaction of the system with some other system.

Classical Error Correction

Imagine that one wants to store a single bit for a time t (we will call this bit a *logical bit*). Let us denote by P_τ the probability that one error occurs in a time interval τ; that is, the probability that the bit flips (if it was 0 then it changes to 1 and vice versa). If $P_\tau \simeq 1$ there will be problems in achieving the goal.

One way to correct the errors is based on what is called *redundant coding.* This consists of using three bits to store the logical bit. That is, we *encode* the information such that if the logical bit is 0 the three bits are 0, and if it is 1, the three bits are 1: $0_L \equiv 000$, and $1_L \equiv 111$. These logical qubits are called *code words.*

After at time τ, we will have

- Probability of no errors: $(1 - P_\tau)^3$ (for example, if we had initially 000, after the time τ it is 000).
- Probability of error in one bit: $3\ P_\tau(1 - P_\tau)^2$ (for example, if we had initially 000, after the time τ it is 100, 010 or 001).
- Probability of error in two bits: $3\ P_\tau^2(1 - P_\tau)$ (for example, if we had initially 000, after the time τ it is 011, 101 or 110).
- Probability of error in three bits: P_τ^3 (for example, if we had initially 000, after the time τ it is 111).

The error correction consists of measuring if the three bits are in the same state or not. If they are in the same state, then we do nothing. If they are in a different state, we use majority vote to change the bit that is different. For example, if we have that the first and the third bit are equal and the second is different (010 or 101), we flip the second bit (000 and 111, respectively) .

After the correction we will have the correct state with a probability $P_\tau^c = (1 - P_\tau)^3 + 3P_\tau(1 - P_\tau)^2 = 1 - 3P_\tau^2 + 2P_\tau^3$. Thus, one gains if $P_\tau^c < 1 - P_\tau$, that is, if (roughly) $P_\tau < 1/3$. If one wants to keep the state for very long times t, one has to perform many measurements. More precisely, assume that $P_\tau = 1 - e^{-\gamma\tau} \simeq \gamma\tau$ for times τ sufficiently short. Let us divide t in N intervals of duration $\tau = t/N$. For N sufficiently large, the probability of having the correct state after performing the correction after the time t will be

$$P_t^c \geq \left[1 - 3\left(\frac{\gamma t}{N}\right)^2\right]^N . \tag{7}$$

For $N \gg 3(\gamma t)^2$ this probability can be made as close to one as desired.

One can generalize this method to the case in which one wants to store k logical bits and allow for errors in t bits. For example, encoding $0_L \equiv 00000$, $1_L \equiv 11111$, one can allow for two errors.

Quantum Error Correction

Imagine that one wants to store a single quantum bit in an unknown state $c_0|0\rangle + c_1|1\rangle$ for a time t (we will call this qubit a *logical qubit*). Let us assume that after a time τ with a probability $1 - P_\tau$ the qubit remains intact and that with a probability P_τ it changes to $|\psi\rangle = c_0|1\rangle + c_1|0\rangle$. This error is called spin flip, and it can be represented by the action of σ_x onto the state of the qubit. As before, if $P_\tau \simeq 1$ there will be problems in achieving the goal.

One can correct the above error by using *redundant coding* [32,33] . For example, one can *encode* the state of the logical qubit in 3 qubits as $|0\rangle_L = |000\rangle$, $|1\rangle_L = |111\rangle$ (code words). The subspace spanned by these states is called subspace of code words.

After at time τ, we will have

- Probability of no errors: $(1 - P_\tau)^3$ (the state will be $|\Psi\rangle_L$).
- Probability of error in one bit: $3\ P_\tau(1-P_\tau)^2$ (the state may be $\sigma_x^1|\Psi\rangle_L$, $\sigma_x^2|\Psi\rangle_L$, or $\sigma_x^3|\Psi\rangle_L$).
- Probability of error in two bits: $3\ P_\tau^2(1 - P_\tau)$ (the state may be $\sigma_x^1\sigma_x^2|\Psi\rangle_L$, $\sigma_x^1\sigma_x^3|\Psi\rangle_L$, or $\sigma_x^2\sigma_x^3|\Psi\rangle_L$).
- Probability of error in three bits: P_τ^3 (the state may be $\sigma_x^1\sigma_x^2\sigma_x^3|\Psi\rangle_L$.

Note that in order to correct the errors, we cannot do the same as in the classical case, since measuring the state of the qubit will collapse it in a different state (for example $|000\rangle$), and therefore the superposition will be destroyed. What we can do is to detect whether the three bits are in the same state or not, without disturbing the state. If the qubits are in the same state, then we do nothing. If they are a different state, we use majority vote to change the bit that is different. All these measurements have to be performed without destroying the superposition. This can be done as follows: first we measure the projector $P = |000\rangle\langle 000| + |111\rangle\langle 111|$ (which corresponds to an incomplete measurement). If we obtain 1, then we leave the qubits as they are. If we obtain 0 then we measure the projector $P_1 = |100\rangle\langle 100| + |011\rangle\langle 011|$: if we obtain 1 we apply the local unitary operator σ_z^1 and if not we proceed. We measure $P_2 = |010\rangle\langle 010| + |101\rangle\langle 101|$; if we obtain 1 we apply the local unitary operator σ_z^2 and if not we apply the operator σ_z^3 (note that if we measure the operator P_3 we would obtain 1 with probability 1). As a result, if there was either no error or one error, it will be corrected. If there were two or more errors, they will not be corrected. Using this method, we achieve the same results as in the classical correction method, namely, by correcting very often we can keep the unknown state of a qubit for as long as we want.

The idea of the method for quantum error correction is based on designing the code words in such a way that every possible error (in the first, second, or third qubit) transforms the subspace of code words onto another subspace which is orthogonal to it, but without modifying its internal structure. Then, by performing an incomplete measurement, we can detect in which subspace our state is, and therefore we know how to correct the error. This method can be generalized to the case in which other kinds of errors can occur. For example, imagine that with a small probability we can have errors consisting of applying the operator σ_α ($\alpha = x, y, z$) to a qubit. We want to preserve the state of k qubits against arbitrary errors in t different qubits. We will denote by E the possible operators corresponding to the errors that we want to correct. For example, $\sigma_x^1 \otimes \sigma_y^4$. We encode the k *logical qubits* in n qubits. The subspace of code words $\mathcal{H}_L$ has dimension 2^k, whereas the Hilbert space $\mathcal{H}$ of all the qubits has dimensions 2^n. Each of the possible error operators (consisting of up to t tensor products of Pauli operators) transform $\mathcal{H}_L$ into a subspace of dimension 2^k (Note that the E's are unitary and therefore they conserve the dimension of the subspace on which they are applied). The subspace of code words has to be such that all these subspaces are mutually orthogonal. This condition imposes a minimum bound (the quantum Hamming bound) to the number of qubits needed, since all these orthogonal subspaces have to fit in $\mathcal{H}$. Let us calculate this bound. To do that we have to count the number of different E operators that have t or less Pauli operators acting on the n qubits. We count first the number of those operators that contain l Pauli operators, and then sum these numbers for $l = 0$ up to $l = t$. For the first part of the problem we find that there are $3^l n!/[l!(n-l)!]$ operators , since there are $n!/[l!(n-l)!]$ combinations of l qubits within n qubits, and in each of the l qubits there are three possible Pauli operators. Thus, the quantum Hamming bound is

$$2^k \sum_{l=0}^{t} 3^l \binom{n}{l} < 2^n. \tag{8}$$

For $k = 1$ the minimum n is 5. Methods have been devised to construct codewords for each of these cases [34,35].

On the other hand, one can take into account the errors that are produced while errors are being corrected, as well as the ones produced during operations. There is a whole theory dealing with the so–called fault tolerant error correction [39], which basically shows that this is always possible provided the error per gate is smaller than some error threshold.

3.4 Error Correction and Decoherence

The above error correction schemes work in the presence of (undesired) coupling to the environment which leads to decoherence. In order to show that, one can expand the operator that describes the evolution of the i–th qubit with its local environment as

$$U^i = \alpha^i 1^i \otimes E_0^i + \epsilon_1^i \sigma_x^i \otimes E_1^i + \epsilon_2^i \sigma_y^i \otimes E_2^i + \epsilon_3^i \sigma_z^i \otimes E_3^i, \tag{9}$$

where the E's are operators acting on the environment, and α^i and $\epsilon_{1,2,3}^i$ are constant numbers. Note that we can always use this expansion given the fact that the Pauli operators (plus the identity) form a basis in the space of operators acting on a qubit. We will consider that the time is sufficiently short so that all $\alpha^i \simeq 1$ and $\epsilon_{1,2,3}^i \ll 1$.

The state of all the qubits after some interaction time can be expanded in terms of the ϵ_k^i as follows:

$$\begin{aligned} U|\psi\rangle|E\rangle &= \prod_{i=1}^{n} U^i|\psi\rangle|E\rangle \qquad (10) \\ &= \Bigg[\prod_{i=1}^{n} \alpha^i 1^i U_0^i + \sum_{j=1}^{n} \epsilon_1^j \sigma_x^j U_1^j \prod_{j\neq i} \alpha^i 1^i U_0^i + \sum_{j=1}^{n} \epsilon_2^j \sigma_y^j U_2^j \prod_{j\neq i} \alpha^i 1^i U_0^i \\ &\quad + \sum_{j=1}^{n} \epsilon_3^j \sigma_z^j U_3^j \prod_{j\neq i} \alpha^i 1^i U_0^i + o(\epsilon^2)\Bigg] |\psi\rangle|E\rangle . \end{aligned}$$

The error correction explained in the previous sections will project the state onto only one of the terms of the expression (10). The state of the environment will therefore factorize, and therefore all the analysis made before remains valid.

4 Entanglement of Pure States

4.1 Introduction

Generally, in Physics we associate physical quantities and situations with mathematical concepts. These mathematical concepts can then be processed using a series of rules (or axioms), which allow us to make predictions back on the physical systems. In particular, in Quantum Mechanics we associate different situations (states) of a physical system with the elements of a complex Hilbert space H. This fact has important consequences, like, for example, the superposition principle: If a system can be in two different states (associated to the vectors $|0\rangle, |1\rangle \in H$) then it can also be in the state described by a linear superposition $\alpha|0\rangle + \beta|1\rangle$. This implies the existence of states in which properties are not well defined. For example, if $|0\rangle$ and $|1\rangle$ represent two states in which an object is situated at different locations, the superposition principle implies that there is a state in which the object does not have a well defined position. This is a striking property of Quantum Mechanics which has raised many debates and paradoxes since the appearance of such a theory [41].

The consequences of the mathematical structure of Quantum Mechanics are even more intriguing when we have a composite system. For example, let us consider two subsystems A and B whose states are associated with the elements of two Hilbert spaces, H_{A} and H_{B}, respectively. We will assume that these systems are located at different places, although for most of our treatment this condition is unimportant. Of course, the states of the whole system comprising A and B must be mathematically described by a Hilbert space H. Let us consider states of the whole system in which one subsystem is in certain state $|i\rangle_{\mathrm{A}}$ and the other in $|j\rangle_{\mathrm{B}}$. One denotes those states as $|i\rangle_{\mathrm{A}} \otimes |j\rangle_{\mathrm{B}} \in H$ or simply $|i,j\rangle \in H$. For example, we can have the systems in the state $|0,0\rangle$ or $|1,1\rangle$. But again, the fact that H is a Hilbert space imposes that any superposition of these stats must also be possible,[2] i.e. the state represented by $|\Psi\rangle \equiv \alpha|0,0\rangle + \beta|1,1\rangle \in H$. A state of this form cannot be described as a certain state for system A and some other state for system B; that is, there exist no pair of vectors $|\phi_{1,2}\rangle_{A,B} \in H_{A,B}$ such that $|\Psi\rangle = |\phi_1, \phi_2\rangle$. States of this form are called entangled states and play a fundamental role in Quantum Information. Note that their existence arises from the fact that the states of the whole system must be described as elements of a Hilbert space themselves. The only way of obtaining H starting from $H_{A,B}$ is by defining it as the span of all vectors in the form $|\phi_1, \phi_2\rangle$, where $|\phi_{1,2}\rangle_{A,B} \in H_{A,B}$. One says that H is the tensor product of H_A and H_B, and writes $H = H_\mathrm{A} \otimes H_\mathrm{B}$. Thus, the existence of entangled states is a direct consequence of the tensor product structure of the Hilbert space describing composite systems.

In this section we will discuss under which conditions a state of two or more subsystems is entangled. We will also quantify the entanglement of such subsystems. We will consider most of the time two systems A and B with corresponding Hilbert spaces H_A and H_B, respectively. We will restrict ourselves to finite-dimensional Hilbert spaces, i.e. $d_{A,B} \equiv \dim(H_{A,B}) < \infty$. Given the fact that all finite Hilbert spaces of the same dimension are isomorphic we can simply take $H_{A,B} = \mathbb{C}^{d_{A,B}}$. We will denote by $\{|k\rangle\}_{k=1}^{d_{A,B}}$ an orthonormal basis in $H_{A,B}$. Although the definitions and results apply for general dimensions, for most of the examples we will consider qubits, i.e. systems where $d_\mathrm{A} = d_\mathrm{B} = 2$. In that case we will take as a basis $\{|0\rangle, |1\rangle\}$. We will use the Pauli operators

$$\sigma_x = |1\rangle\langle 0| + |0\rangle\langle 1|, \tag{11a}$$

$$\sigma_y = -\mathrm{i}(|1\rangle\langle 0| - |0\rangle\langle 1|), \tag{11b}$$

$$\sigma_z = |1\rangle\langle 1| - |0\rangle\langle 0|, \tag{11c}$$

and $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$.

4.2 2-Partite Entanglement: Definition

Let us consider two systems A and B. We say that $|\Psi\rangle \in \mathcal{H}_\mathrm{A} \otimes H_\mathrm{B}$ is a *product state* if there exist $|\phi_{1,2}\rangle_{A,B} \in H_{A,B}$ such that $|\Psi\rangle = |\phi_1, \phi_2\rangle$. Otherwise we say that $|\Psi\rangle$ is an *entangled state.*

Example 1 *The states* $|0,1\rangle$ *and* $|1,0\rangle$ *are product states, whereas the Bell states*

$$|\Phi^\pm\rangle \equiv \frac{1}{\sqrt{2}}(|0,0\rangle \pm |1,1\rangle), \tag{12a}$$

$$|\Psi^\pm\rangle \equiv \frac{1}{\sqrt{2}}(|0,1\rangle \pm |1,0\rangle), \tag{12b}$$

are entangled.

[2] Except for the case in which there are superselection rules, but we will not consider those systems here.

The most important property of entangled states is that they carry correlations. That is, if we measure an observable in A and another in B the outcomes will be, in general, correlated. For example, if we have the state $|\Psi^-\rangle$ and measure the observable σ_z in both systems we will obtain the opposite result. Actually, if we measure any observable $\boldsymbol{\sigma}\cdot\boldsymbol{n}$ we will always obtain opposite results in A and B, the reason being that $|\Psi^-\rangle$ is invariant under global rotations [i.e. $U\otimes U|\Psi^-\rangle = |\Psi^-\rangle$ for all unitary operators $U \in su(2)$]. Note that for all entangled states there always exist some correlations. For product vectors, however, the outcomes in A are independent of the outcomes in B. This can be also viewed by noting that if A and B are two observables, then $\langle A\otimes B\rangle = \langle A\rangle\langle B\rangle$ for product vectors, but not (in general) for entangled states.

The existence of correlations, by itself, is not a property of entangled states. For example, if somebody provides us with two boxes A and B in which there are either two black or two white balls, when we open the boxes we will see correlations. However, the correlations carried by entangled states are, in some sense, different than those, since they occur for any pair of observables. In fact, classical correlations like the ones displayed by the balls in the boxes are restricted by Bell's inequalities, whereas the ones corresponding to entangled states may violate them [42]. This is why with the correlations contained in entangled states we can perform things that are not possible using classical correlations.

In order to create entangled states out of product states we need interactions. This can be easily understood as follows. If we do not have interactions, the Hamiltonian describing the evolution of systems A and B will be written as $H = H_\mathrm{A}\otimes \mathbb{1}_\mathrm{B} + \mathbb{1}_\mathrm{A}\otimes H_\mathrm{B}$, where $\mathbb{1}$ is the identity operator. Since $H_\mathrm{A}\otimes\mathbb{1}_\mathrm{B}$ and $\mathbb{1}_\mathrm{A}\otimes H_\mathrm{B}$ commute with each other, we have that the evolution operator can be always written as $U(t) = U_\mathrm{A}(t)\otimes U_\mathrm{B}(t)$, and the product state $|\Psi(0)\rangle = |\phi_1\rangle_\mathrm{A}\otimes|\phi_2\rangle_\mathrm{B}$ will evolve into $|\Psi(t)\rangle = [U_\mathrm{A}(t)|\phi_1\rangle_\mathrm{A}]\otimes[U_\mathrm{B}(t)|\phi_2\rangle_\mathrm{B}] = |\phi_1(t)\rangle_\mathrm{A}\otimes|\phi_2(t)\rangle_\mathrm{B}$ which is a product state. Operators of the form $U = U_\mathrm{A}\otimes U_\mathrm{B}$ are called local operators. Similarly, we cannot get entangled states by measuring observables in A and B independently since the state after the measurement will be changed by local operators. One says that entanglement cannot be created by local operations (operations meaning any action on the systems). Note, however, that product states can be obtained by local operations (in particular, by measurements).

4.3 2-Partite Entanglement: Qualifying and Quantifying

In this subsection we will determine how we can tell whether a state is entangled or not, and how much. We consider a state of the form

$$|\Psi\rangle = \sum_{i=1}^{d_\mathrm{A}}\sum_{j=1}^{d_\mathrm{B}} c_{i,j}|i,j\rangle. \tag{13}$$

All the information of the state is in the coefficients $c_{i,j}$ which form a $d_\mathrm{A}\times d_\mathrm{B}$ matrix that we will call C. Note that we could have chosen another orthonormal bases in $H_{A,B}$ to express this state. In fact, there is a particular basis in which the matrix of the coefficients is diagonal and positive. If we choose such a basis to write the state, it will have the simple form

$$|\Psi\rangle = \sum_{k=1}^{d} d_k |u_k, v_k\rangle, \tag{14}$$

where $d = \min(d_\mathrm{A}, d_\mathrm{B})$ and

$$\sum_{k=1}^{d} d_k^2 = 1. \tag{15}$$

The form (14) is called *Schmidt decomposition*. Its existence directly follows from the singular value decomposition of the matrix C, i.e., the existence of two unitaries U and V and a diagonal one D whose diagonal elements are the d's such that $C = UDV$. Substituting $c_{i,j} = \sum_k d_k U_{i,k} V_{k,j}$ we obtain the Schmidt decomposition. The coefficients d_k are called Schmidt coefficients and the bases $\{|u_k\rangle\} \in H_\mathrm{A}$ and $\{|v_k\rangle\} \in H_\mathrm{B}$ are called Schmidt bases.

Once we have expressed the state in the form (14), it is very simple to obtain some other information. For example, if we are interested in predicting expectation values or probabilities of outcomes if we only measure system A (or B), all the information about them is in the reduced density operator $\rho_\mathrm{A} = \mathrm{tr}_\mathrm{B}(|\Psi\rangle\langle\Psi|)$ (analogously for ρ_B). Using the Schmidt decomposition we obtain

$$\rho_\mathrm{A} = \sum d_k^2 |u_k\rangle\langle u_k|, \tag{16a}$$

$$\rho_\mathrm{B} = \sum d_k^2 |v_k\rangle\langle v_k|. \tag{16b}$$

Conversely, the Schmidt coefficients and the corresponding bases can be easily found by simply diagonalizing both reduced density operators.

Example 2 *The Schmidt decomposition of a product state is trivial ($|\phi_1, \phi_2\rangle$ directly). The Schmidt decomposition of $|\Phi^+\rangle$ and $|\Psi^+\rangle$ is given directly by their definitions (12a), whereas for $|\Phi^-\rangle$ we have $d_1 = d_2 = 1/\sqrt{2}$, $|u_1\rangle = |v_1\rangle = |0\rangle$ and $|u_2\rangle = -|v_2\rangle = |1\rangle$. Note that the Schmidt decomposition is unique except when the reduced density operators are degenerate.*

For a product state $|\phi_1, \phi_2\rangle$, the reduced density operators are rank-one projectors, i.e. $\rho_{A,B} = |\phi_{1,2}\rangle\langle\phi_{1,2}|$. This means that there is only one Schmidt coefficient which is different from zero. Conversely, if we have a state with only one Schmidt coefficient then it must be a product state. Equivalently, $|\Psi\rangle$ is a product state if and only if the corresponding reduced density operators correspond to pure states. This means that if we have an entangled state, the corresponding reduced density operators must correspond to mixed states, or, equivalently, that there must be more than one nonzero Schmidt coefficient. Thus, we see that the entanglement of a state is directly related to the mixedness of the reduced density operators. This is intuitively clear since, as we mentioned above, entangled states give rise to correlations and if we only observe one of the systems we lose information about these correlations which results in the fact that we will effectively have a mixed state. This suggests that we can measure the degree of entanglement by the degree of mixedness of the reduced density operators.

There are several measures of mixedness of density operators; perhaps the most popular one is the von Neumann entropy $S(\rho) = -\mathrm{tr}(\rho \log_2 \rho)$. For a pure state this entropy is zero, whereas for a maximally mixed state (described by the identity operator, properly normalized) it gives $\log_2 d$, where d is the dimension of the Hilbert space. The entropy is convex, i.e. for $p \in [0,1]$, $S[p\rho_1 + (1-p)\rho_2] \geq pS(\rho_1) + (1-p)S(\rho_2)$, which means that it always increases by mixing (i.e by losing information). This motivates the following definition:

Given a state $|\Psi\rangle$, we define the *entropy of entanglement* $E(\Psi)$ as the von Neumann entropy of the reduced density operator [43]. Using (16a) we have

$$E(\Psi) = S(\rho_{\mathrm{A}}) = S(\rho_{\mathrm{B}}) = -\sum_{k=1}^{d} d_k^2 \log_2(d_k^2). \tag{17}$$

For the moment this is just a definition to quantify the amount of entanglement contained in a state, which is based on the idea that the more mixed the reduced density operator is, the more entangled the original state is. Later on, when we discuss the process of distillation, we will give a definite physical meaning to this definition. On the other hand, note that this definition only applies to pure states. As we will see in the next section, for mixed states the definition of entanglement is not so obvious and its quantification is a very complicated problem.

The entropy of entanglement only depends on the Schmidt coefficients, but not on the corresponding basis. This means that it is invariant under local unitary operations. That is, if $|\Psi'\rangle = U_{\mathrm{A}} \otimes U_{\mathrm{B}} |\Psi\rangle$, then $E(\Psi') = E(\Psi)$. This is clear since the Schmidt decomposition of $|\Psi'\rangle$ can be directly calculated from the one of $|\Psi\rangle$, and it is clear that the Schmidt coefficients are the same. On the other hand, one can show that it cannot increase in average by local operations [44]. That is, if we perform (independent) measurements in A and B and obtain the state $|\Psi_k\rangle$ after the measurement with probability p_k, we have that

$$E(\Psi) \geq \sum_k p_k E(\Psi_k). \tag{18}$$

Note, however, that the previous inequality does not imply that none of the $E(\Psi_k)$ can be larger than $E(\Psi)$, or even the maximum allowed $\log_2 d$. In fact this is what occurs in the process of distillation, namely that with certain probability one is able to increase the entanglement.

Note also that the entanglement so defined is additive in the sense that if we have two states $|\Phi_1\rangle_{AB}$ and $|\Phi_2\rangle_{AB}$, the total entanglement is $E(\Phi_1) + E(\Phi_2)$. States in $\mathbb{C}^d \otimes \mathbb{C}^d$ for which $E(\Psi) = \log_2(d)$ are called *maximally entangled states* in d dimensions.

4.4 Multipartite Entanglement

Let us now consider more systems $A_1, A_2, \ldots, A_N$. Now, we can have entangled states and product states of the different systems [46]. For example, we can have a state of the form $|\Psi\rangle = |\phi_1\rangle_{A_1A_3} \otimes |\phi_2\rangle_{A_2A_5A_6} \otimes |\phi_3\rangle_{A_4}$, where $|\phi_1\rangle$ and $|\phi_2\rangle$ cannot be written as product states. It is clear that in a state like that, the parties A_1 and A_3 are entangled with each other, but not to the rest; similarly, the parties A_2, A_5, and A_6 are entangled among themselves, and the party A_4 is completely disentangled.

In general we can consider all possible partitions of those systems in which we group certain ones of them. For example, we can consider the partition (A_1A_3), $(A_2A_5A_6)$, (A_4). We can classify the entangled states according to the different partitions. That is, a state is entangled according to some partition if it can be written as a product state of the corresponding disjoint elements of the groups, but not within each of the groups. In order to determine the partition corresponding to a particular state we can calculate all possible reduced density operators and look whether they correspond to mixed states or not.

The quantification of the multipartite entanglement is a more complicated question which can be illustrated by the following example.

Example 3 *Let us consider three parties and the states [45]*

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0,0,0\rangle - |1,1,1\rangle), \tag{19a}$$

$$|W\rangle = \frac{1}{\sqrt{3}}(|0,0,1\rangle + |0,1,0\rangle + |1,0,0\rangle). \tag{19b}$$

Those are entangled states according to the partition $(A_1A_2A_3)$. However it is hard to say which one is more entangled. Certainly, the first one possesses a sticking non–local behavior, in the sense that it can be used to prove Bell's theorem without using inequalities [45]. However, it is very weak in the sense that if one party does not participate in the measurement (or is lost), then all the entanglement disappears. However, the second one retains some entanglement even if one particle is lost (in fact it is the most robust against particle losses) [48].

5 Entanglement of Mixed States

5.1 Introduction

The states that we have considered in the previous section are idealized. In reality, all systems interact with some sort of environment. Thus, we should include the state of the environment in our description in order to be consistent. In fact, due to the interaction system–environment and according to the discussion of Section 3, they will become entangled even if initially they were in a product state:

$$|\Psi_\mathrm{S}(0)\rangle_\mathrm{S} \otimes |\Psi_\mathrm{E}(0)\rangle_\mathrm{E} \xrightarrow{t} |\Psi(t)\rangle_\mathrm{SE}. \tag{20}$$

Since we are only interested in our system, all the information that we can acquire (without performing measurements in the environment) is contained in the reduced density operator

$$\rho_\mathrm{S}(t) = \mathrm{tr}_\mathrm{E}[|\Psi(t)\rangle_\mathrm{SE}\langle\Psi(t)|], \tag{21}$$

which will correspond to a mixed state as it was discussed in Section 3. Note that density operators can always be written in the form

$$\rho = \sum p_k |\phi_k\rangle_{\mathrm{A}}\langle\phi_k|, \tag{22}$$

where the $p_k \geq 0$ add up to one. One particular decomposition of the form (22) is the spectral decomposition, in which, additionally, the vectors $\{|\phi_k\rangle\}$ form an orthonormal basis. In general, except for pure states (rank-one density operators) there are infinitely many decompositions of the form (22).

Note that a decomposition like (22) tells us one way of creating a state described by ρ. We simply have to prepare the system in state $|\phi_k\rangle$ with probability p_k. For example, we could take a large number N of systems, $n_k \sim Np_k$ prepared in the state $|\phi_k\rangle$ and then choose one of these systems randomly. The fact that there exist infinitely many decompositions of a state means that it can be prepared in infinitely many different forms. For example, the state $\rho = \mathbb{1}/2$ can be prepared by choosing randomly one among the states $\{|0\rangle, |1\rangle\}$ or one among the states $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, or even selecting randomly one state $\cos(\theta/2)|0\rangle + e^{\mathrm{i}\phi}\sin(\theta/2)|1\rangle$ by choosing the angles $(\theta, \phi) \in ([0,\pi], [0, 2\pi))$ according to the probability density $\sin(\theta)$. It is also worth stressing that even though the systems are prepared in different forms, they are completely indistinguishable. The reason is that the probability of any outcome after a measurement is completely determined by the density operator, so that if two systems have the same density operator they cannot be distinguished by performing any measurement (and therefore by any means).

Density operators are linear and self–adjoint ($\rho = \rho^\dagger$), have trace one [$\mathrm{tr}(\rho) = 1$], and are positive ($\rho \geq 0$). A self–adjoint operator X is positive if for all $|\Psi\rangle \in H$, $\langle\Psi|X|\Psi\rangle \geq 0$.[3] Equivalently, if all its eigenvalues are ≥ 0. The set of linear self–adjoint operators acting on H will be denoted by $\mathcal{A}(H)$. This set has itself the structure of a real Hilbert space with the scalar product $(A, B) = \mathrm{tr}(AB)$. We will denote the set of positive operators by $\mathcal{P}(H) \subset \mathcal{A}(H)$. This set is not a vector space but a convex set; that is, if $\rho_{1,2} \in \mathcal{P}(H)$, and $p \in [0,1]$, then $p\rho_1 + (1-p)\rho_2 \in \mathcal{P}(H)$. We will also consider linear maps $\epsilon : \mathcal{A}(H) \to \mathcal{A}(H')$. The set of all linear maps of this form will be denoted by $\mathcal{M}(H, H')$.

If we write

$$\rho = \sum_{i,j} \rho_{i,j} |i\rangle\langle j|, \tag{23}$$

we can represent the density operator in terms of a matrix whose elements are $\rho_{i,j}$. In particular for two systems, we can write this matrix as

$$\rho = \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,d_{\mathrm{A}}} \\ A_{2,1} & A_{2,2} & \dots & A_{2,d_{\mathrm{A}}} \\ & & \dots & \\ A_{d_{\mathrm{A}},1} & A_{d_{\mathrm{A}},2} & \dots & A_{d_{\mathrm{A}},d_{\mathrm{A}}} \end{pmatrix}, \tag{24}$$

where the A's are $d_{\mathrm{B}} \times d_{\mathrm{B}}$ matrices with $(A_{i_i,i_2})_{j_i,j_2} = \langle i_1, j_1|\rho|i_2, j_2\rangle$.

[3] Strictly speaking, we should say positive semidefinite.

We thus have to define entanglement for mixed states. In Quantum Information a state is called entangled if it cannot be prepared by local operations (and classical communication) out of a product state [49]. This definition makes sense since: (i) it reduces to the one given in the previous section for pure states; (ii) it directly shows that the entanglement can only be produced by interactions; (iii) it does not consider classical correlations (like the ones discussed in the previous section in the example of boxes with black and white balls) to be related to the phenomenon of entanglement; (iv) as we will see, this definition is equivalent to imposing that mixtures of product states are not entangled.

Example 4 *The state described by* $\rho = |0,0\rangle\langle 0,0|$ *is not entangled since it is already a product state.*

Example 5 *Any density operator of the form* $\rho = \rho_{\mathrm{A}} \otimes \rho_{\mathrm{B}}$ *is not entangled since the states* $\rho_{A,B}$ *can be prepared locally out of the state* $|0\rangle$*. To see this, if we write* ρ *as in (22) we can simply transform the state* $|0\rangle$ *into the* $|\phi_k\rangle$ *with probability* p_k*. For example, the state* $\rho = (\mathbb{1}/2) \otimes (\mathbb{1}/2)$ *can be prepared if we prepare each of the systems in the states* $|0\rangle$ *or* $|1\rangle$*.*

Example 6 *The state*

$$\rho = \frac{1}{2}(|0,0\rangle\langle 0,0| + |1,1\rangle\langle 1,1|), \tag{25}$$

is not entangled since it can be locally prepared as follows. We choose randomly 0 or 1. If we have 0, we prepare both A and B in state $|0\rangle$ *and otherwise in* $|1\rangle$*. Obviously, in this way we do not need any interaction between the systems. We just need classical communication between the location of A and B so that the corresponding prepares can agree on the state they prepare.*

It turns out that, in contrast to pure states, it is very hard in general to determine if some given density operator ρ is entangled or not [50,51]. Indeed, this is still an open question which has only been partially solved in certain cases. Thus, even though entanglement seems to be the essential ingredient of Quantum Information we still do not even know how to qualify it. In this section we will summarize some of the results that are known for this problem, and will present some of the questions that still remain open.

5.2 2-Partite Entanglement: Definition

We say that $\rho \in \mathcal{P}(H_{\mathrm{A}} \otimes H_{\mathrm{B}})$ is *entangled* (equivalently, inseparable) if it cannot be prepared locally with the help of classical communication. Otherwise we say that it is *separable.* This definition is equivalent to the following mathematical characterization: ρ is separable if and only if there exist $p_k \geq 0$ and $\{|a_k\rangle\} \in H_{\mathrm{A}}$ and $\{|b_k\rangle\} \in H_{\mathrm{B}}$ such that

$$\rho = \sum_k p_k |a_k, b_k\rangle\langle a_k, b_k|. \tag{26}$$

Otherwise it is entangled. This follows from the discussion of the introduction of this section. Indeed, if ρ is separable it means that it can be prepared locally, and this means that we can transform the state $|0,0\rangle$ into the state $|a_k, b_k\rangle$ by local operations with probability p_k. Conversely, if we can write the state as in (26) this means that we can prepare it locally following this procedure.

Thus, the problem we want to address first is: given ρ, is it entangled? In order to show that this problem is by no means trivial, let us consider some examples for the simplest case in which our systems are qubits.

Example 7 *The state*

$$\rho = \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|), \tag{27}$$

is separable even though it is written as a mixture of two entangled states. The reason is that the same state can be written as (25) which can be directly seen if we substitute the definition of the Bell states (12a). Note that this means that the same state can be created by two different methods, one which involves interactions and another one which does not. Our definition of entanglement implies that this state is not entangled since there is a way of preparing it without any interactions. In some sense, the first way of preparing is very inefficient, since creating entangled state usually requires a lot of effort and we could obtain exactly the same state without effort.

Example 8 *We take the matrix corresponding to ρ as (see [24])*

$$\frac{1}{20}\begin{pmatrix} 7 & 1 & 2 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 2 & 3 & 1 \\ 2 & 2 & 1 & 7 \end{pmatrix}. \tag{28}$$

One can readily see that this matrix has no degenerate eigenvalues and that all the eigenvectors correspond to entangled states. That is, the spectral decomposition contains states that cannot be created locally. However, it can be easily shown that

$$\rho = \frac{1}{5}(|0,0\rangle\langle 0,0| + |1,1\rangle\langle 1,1|) + \frac{2}{5}|+,+\rangle\langle +,+| + \frac{1}{10}(|0,-\rangle\langle 0,-| + |1,-\rangle\langle 1,-|), \tag{29}$$

and therefore is separable.

These examples illustrate the so–called separability problem. A given density operator has infinitely many decompositions. If there exists one which has the form (26) then the state is separable. Otherwise it is entangled. However, in practice it is impossible to check all possible decompositions of a density operator and therefore is impossible to see whether it is entangled or not.

5.3 2-Partite Entanglement: Entanglement Witnesses

In this subsection we will review the concept of entanglement witnesses (EW) [52]. These are observables that permit to detect the presence of entangled states.

An operator $W \in \mathcal{A}(H_\mathrm{A} \otimes H_\mathrm{B})$ is an *entanglement witness* if it fulfills

(a) $W \not\geq 0$, and
(b) $\langle a, b|W|a, b\rangle \geq 0$ for all $|a\rangle \in H_\mathrm{A}$ and $|b\rangle \in H_\mathrm{B}$.

That is, they are non–positive operators but they are positive on product states.

The properties of EWs automatically give us a necessary criterion for separability:

Proposition 1 *(Necessary condition for separability). If ρ is separable then $\mathrm{tr}(W\rho) \geq 0$ for all EW.*

Proof: If ρ is separable, then it can be written as (26), and therefore

$$\mathrm{tr}(W\rho) = \sum_k p_k \langle a_k, b_k|W|a_k, b_k\rangle \geq 0. \tag{30}$$

■

The usefulness of this proposition is based on the fact that, given ρ, if we find an EW such that $\mathrm{tr}(W\rho) < 0$ then we automatically know that the state is entangled. This is why these operators are called EW, since they witness the presence of entanglement. If $\mathrm{tr}(W\rho) < 0$, we say that W detects the state ρ.

Example 9 *For qubits, the operator $W = \mathbb{1} - 2|\Phi^+\rangle\langle\Phi^+|$ is an EW. First, it is clear that $\langle\Phi^+|W|\Phi^+\rangle = -1 < 0$. Besides, for any pair of states $|a\rangle = |0\rangle + \alpha|1\rangle, |b\rangle = |0\rangle + \beta|1\rangle \in \mathbb{C}^2$ we have $\langle a, b|W|a, b\rangle = (1 + |\alpha|^2)(1 + |\beta|^2) - |1 + \alpha^*\beta^*|^2 \geq 0$.*

Example 10 *For qubits, let us consider the operator considered by Clauser, Horn, Shimony, and Holt [53] in order to derive a Bell-like inequality that proves Bell's theorem and can be experimentally tested,*

$$S = \sigma_0^A \sigma_{\pi/4}^B + \sigma_{\pi/2}^A \sigma_{\pi/4}^B + \sigma_{\pi/2}^A \sigma_{3\pi/4}^B - \sigma_0^A \sigma_{3\pi/4}^B, \tag{31}$$

where $\sigma_\theta \equiv \cos\theta\sigma_z + \sin\theta\sigma_x$. If we take any normalized product vector $|a, b\rangle$ we have

$$\langle a, b|S|a, b\rangle = (\alpha + \gamma)\beta + (\alpha - \gamma)\delta, \tag{32}$$

where $\alpha = \langle a|\sigma_0|a\rangle$, $\gamma = \langle a|\sigma_{\pi/2}|a\rangle$, $\beta = \langle b|\sigma_{\pi/4}|b\rangle$, and $\delta = \langle b|_{3\pi/4}|b\rangle$. Note that the possible values of these parameters lie in the interval $[-1, 1]$. In fact, it is simple to check that $|\langle a, b|S|a, b\rangle| \leq 2$. Certain entangled states, though, fulfill $|\langle\Psi|S|\Psi\rangle| > 2$. Those are precisely the states that violate the CHSH inequalities. In particular, $\langle\Psi^-|S|\Psi^-\rangle = -2\sqrt{2}$. We can now construct the following EW, $W = 2 + S$. According to the properties of S we have just mentioned, it is clear that for product vectors $\langle a, b|W|a, b\rangle \geq 0$, whereas it is not positive since there exist a vector $|\Psi^-\rangle$ such that $\langle\Psi^-|W|\Psi^-\rangle = -2(\sqrt{2} - 1) < 0$.

For the moment we have that given ρ, if we find an EW, W, that detects ρ then we know that it corresponds to an entangled state. However, what about the existence of such an EW for any arbitrary entangled state ρ? In fact, the following proposition ensures that [52]:

Proposition 2 *(Sufficient condition for separability). If ρ_0 is entangled then there exists and EW, W, such that* $\mathrm{tr}(W\rho_0) < 0$.

Proof: Here we sketch the proof of this result. We consider the Hilbert space of self–adjoint operators $X = \mathcal{A}(H_\mathrm{A} \otimes H_\mathrm{B})$, which is locally convex (since it is a Hilbert space). We also consider the following two subsets. First, $\mathcal{S}$ is the set of all separable density operators. Second, the set formed by a single element, the entangled density operator ρ_0. These two sets are clearly convex, disjoint, and non–empty. Moreover, the second one is compact (it is closed and bounded) whereas the first one is closed (any converging sequence of separable density operators converges to a separable density operator). Thus, Hahn–Banach's theorem [54] ensures that there exists $\Lambda \in X^*$ (the dual of X) and $\gamma \in \boldsymbol{R}$ such that $\Lambda(\rho_0) < \gamma_1 \leq \Lambda(\rho)$ for all $\rho \in \mathcal{S}$. On the other hand, Riesz–Fréchet's theorem [54] ensures that for any $\Lambda \in X^*$ there exists $W_\Lambda \in X$ such that $\Lambda(\rho) = \mathrm{tr}(W_\Lambda \rho)$ for all $\rho \in X$. Thus, defining $W = W_\Lambda - \gamma_1$ we have that W is an EW that detects ρ_0. ■

Combining these two propositions we have that ρ is entangled if and only if there exists an EW that detects it. Note that this powerful result does not solve the problem of separability, since we do not know how to construct all possible EW's. If that were possible, we could check with all of them whether ρ is entangled or separable. However, the effort of constructing these EW's is comparable to checking all possible decompositions of ρ, and therefore is not practical.

5.4 2-Partite Entanglement: Positive Maps

In this subsection we relate the entanglement witnesses to the concept of positive maps, which has been thoroughly studied by mathematicians [55,56]. Thus, we can use all the results obtained by them in the problem of separability.

A linear map $\epsilon \in \mathcal{M}(H, H')$ is called a *positive map* (PM) if for all $0 \leq \rho \in \mathcal{A}(H)$, $\epsilon(\rho) \geq 0$.

Example 11 *All physical actions correspond to positive maps since they transform density operators into density operators (which are positive).*

Example 12 *Transposition in a given basis is also a map $\in \mathcal{M}(H, H)$. Given an orthonormal basis $\{|k\rangle\} \in H$, we can write any operator $\rho \in \mathcal{A}(H)$ as $\rho = \sum_{i,j} \rho_{i,j} |i\rangle\langle j|$. We define the transpose of ρ in that basis, $\rho^T = \tau(\rho) \equiv \sum_{i,j} \rho_{i,j} |j\rangle\langle i|$. It is clear that since the eigenvalues of a self–adjoint matrix and its transpose are the same, transposition is a PM.*

When we have tensor product spaces, it is convenient to define the extensions of maps as follows. Given a linear map $\epsilon \in \mathcal{M}(H,H')$ and another Hilbert space H'', we define the extension of ϵ as the linear map $\epsilon \otimes 1 \in \mathcal{M}(H \otimes H'', H' \otimes H'')$ such that

$$(\epsilon \otimes 1)\left(\sum_k A_k \otimes B_k\right) = \sum_k \epsilon(A_k) \otimes B_k, \tag{33}$$

where $A_k \in \mathcal{A}(H)$ and $B_k \in \mathcal{A}(H'')$. Analogously, we can define the extension $1 \otimes \epsilon$. We say that a linear map is a *completely positive map* (CPM) if all its extensions are PM. The physical motivation for defining extensions of maps is illustrated by the following example.

Example 13 *Physical actions. Imagine we have two subsystems, A and B, and that we apply a physical action to system A alone. It may happen that the state of A and B is initially entangled, in which case the action on system A will modify the state of the whole system. Thus, such a physical action will be described mathematically by an extension of a PM to include the Hilbert space of subsystem B, i.e. an element of* $\mathcal{M}(H_{\mathrm{A}} \otimes H_{\mathrm{B}}, H_{\mathrm{A}} \otimes H_{\mathrm{B}})$*. Moreover, such action must keep the density operator describing the whole system positive, whatever B we have, and therefore it must be described by a CPM.*

Example 14 *Partial transposition. Not all PM's are CPM's. For example, if we extend the transposition, it is not a CPM. Given A and B, and an orthonormal basis in A,* $\{|k\rangle\}$*, we have* $\rho^{T_{\mathrm{A}}} \equiv (\tau \otimes 1)(\rho) = \sum_{i,j} \langle i|\rho|j\rangle |j\rangle_{\mathrm{A}}\langle i|$*. Note that* $\langle i|\rho|j\rangle |j\rangle_{\mathrm{A}}\langle i|$ *is an operator acting on* H_{B}*. If we use the matrix representation (24), we have that*

$$\rho^{T_{\mathrm{A}}} = \begin{pmatrix} A_{1,1} & A_{2,1} & \dots & A_{d_{\mathrm{A}},1} \\ A_{1,2} & A_{2,2} & \dots & A_{d_{\mathrm{A}},2} \\ & & \dots & \\ A_{1,d_{\mathrm{A}}} & A_{2,d_{\mathrm{A}}} & \dots & A_{d_{\mathrm{A}},d_{\mathrm{A}}} \end{pmatrix}, \tag{34}$$

This extension of transposition in this form is usually referred to as partial transposition with respect to A. In order to see that partial transposition is not a PM (and therefore transposition is not a CPM) we consider the following example: $\rho = |\Phi^+\rangle\langle\Phi^+|$*. The corresponding matrix is given by*

$$\rho = \frac{1}{2}\begin{pmatrix} 1\,0\,0\,1 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 1\,0\,0\,1 \end{pmatrix}, \quad \rho^{T_{\mathrm{A}}} = \frac{1}{2}\begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}, \tag{35}$$

which is obviously non positive (it has a negative eigenvalue). We can analogously define the partial transposition with respect to B.

The importance of CPM is that they represent physical actions. All CPM can be written as (see below)

$$\epsilon(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{36}$$

where A_k are operators (not necessarily self–adjoint) acting on H. If additionally we have

$$\sum_k A_k^\dagger A_k = \mathbb{1}, \tag{37}$$

we have a trace preserving CPM, since $\mathrm{tr}(\rho) = \mathrm{tr}[\epsilon(\rho)]$ for all $\rho \in \mathcal{A}(H)$. In fact, any physical action corresponds to a trace preserving CPM and viceversa.

There exists no formula similar to (36) for the general action of a PM. In some sense, PM's are much more difficult to classify and characterize than CPM's. In fact, except for some special cases, there exist no characterization of PM's. However, we can distinguish among two kinds of PM's. The decomposable PM's are those which can be written as

$$\epsilon = \epsilon_1 + \epsilon_2 \circ \tau, \tag{38}$$

where $\epsilon_{1,2}$ are CPM's, τ is the transposition, and $\circ$ denotes composition (i.e., application of one map after the other). All other PM's are called non–decomposable. An important result in this context is the following [55,56]

Lemma 1 *(all positve maps are decomposable in low dimensions) If* $\dim(H)$ $\dim(H') \leq 6$ *then all the PM's are decomposable, and otherwise there also exist non–decomposable PM's.*

Decomposable maps, in some sense, are very simple since they just involve CPM's [which are characterized by (36)] and transposition. However, non–decomposable maps are practically unknown even in the mathematical context. Examples of these maps have been given in [57,58]. Their general form has recently been derived [59].

Now, we can relate the problem of separability to PM's. We can do that via the EW's defined in the preceding section. In order to do that, we consider the existing isomorphism between $\mathcal{A}(H_\mathrm{A} \otimes H_\mathrm{B})$ and $\mathcal{M}(H_\mathrm{A}, H_\mathrm{B})$ [60]. Given $W_{AB} \in \mathcal{A}(H_\mathrm{A} \otimes H_\mathrm{B})$, and an orthonormal basis $\{|k\rangle\} \in H_\mathrm{A}$, we can construct a linear map $\epsilon \in \mathcal{M}(H_\mathrm{A}, H_\mathrm{B})$ which acts as follows

$$\epsilon_W(\rho_\mathrm{A}) = \mathrm{tr}_A(W_{AB}\rho_A^T), \tag{39}$$

where the trace is taken in H_A, and ρ_A^T denotes the transpose of ρ_A in the basis $\{|k\rangle\}$. The correspondence $W \to \epsilon_W$ is an isomorphism. First, it is clearly linear. On the other hand, it is one-to-one. This can be seen by showing that to each linear map we can associate an operator through the inverse correspondence of the one defined in (39). In particular, given $\epsilon \in \mathcal{M}(H_\mathrm{A}, H_\mathrm{B})$ and an orthonormal basis $\{|k\rangle\} \in H_\mathrm{A}$ we can construct $W \in \mathcal{A}(H_\mathrm{A} \otimes H_\mathrm{B})$ as

$$W_\epsilon = (1 \otimes \epsilon)(|\Phi\rangle\langle\Phi|), \tag{40}$$

where

$$|\Phi\rangle = \sum_k |k,k\rangle \in H_\mathrm{A} \otimes H_\mathrm{A}. \tag{41}$$

It can be easily checked by substitution that $W_\epsilon = W$.

Under this isomorphism, positive operators are mapped into CPM's. This fact allows us to easily show that all CPM's can be expressed in the form (36). On the other hand, EW's are mapped into positive (but not completely positive) maps. Thus, with the help of this isomorphism, we can establish the link between entanglement and PM's [52]:

Proposition 3 *(Entanglement and positive maps) $\rho \in \mathcal{P}(H_\mathrm{A} \otimes H_\mathrm{B})$ is separable if and only if $(\epsilon \otimes 1)(\rho) \geq 0$ for all PM $\epsilon \in \mathcal{M}(H_\mathrm{A}, H_\mathrm{B})$.*

Proof: ($\Rightarrow$) If ρ is separable we can write it as (26). Thus we have

$$(\epsilon \otimes 1)(\rho) = \sum p_k \epsilon(|a_k\rangle\langle a_k|) \otimes |b_k\rangle\langle b_k|) \geq 0 \tag{42}$$

since $\epsilon(|a_k\rangle\langle a_k|) \geq 0$ given the fact that ϵ is a PM. ($\Leftarrow$) Let us consider any EW, W. Using the isomorphism (39) we have that there exists a PM ϵ_W such that

$$\mathrm{tr}_{AB}(W\rho) = \mathrm{tr}_\mathrm{A}[(1 \otimes \epsilon)(\rho)] \geq 0, \tag{43}$$

since for all PM ϵ, $(1 \otimes \epsilon)(\rho) \geq 0$. Thus, using Proposition 2 we have that ρ is separable. ■

Once this link is established, we can use the characterization of PM's to characterize entanglement. In particular, according to Lemma 1 we have that for low dimensions all PM's are decomposable. Thus, we have the following result due to Peres [61] and the Horodecki family [52].

Proposition 4 *$\rho \in \mathcal{A}(\mathbb{C}^2 \otimes \mathbb{C}^N)$ with $N \leq 3$ is separable if and only if $\rho^{T_\mathrm{A}} \geq 0$.*

Proof: ($\Rightarrow$) Follows directly from Proposition 3 and the fact that transposition is a PM. ($\Leftarrow$) We have $\dim(H_\mathrm{A}) = 2$ and $\dim(H_\mathrm{B}) \leq 3$. Thus, according to Lemma 1 any PM $\epsilon \in \mathcal{M}(H_\mathrm{A}, H_\mathrm{B})$ is decomposable and can be written as (38). Thus,

$$(\epsilon \otimes 1)(\rho) = (\epsilon_1 \otimes 1)(\rho) + (\epsilon_2 \circ \tau \otimes 1)(\rho) \geq 0, \tag{44}$$

since $\epsilon_{1,2}$ are CPM's and $(\tau \otimes 1)(\rho) = \rho^{T_\mathrm{A}} \geq 0$. ■

Example 15 *The state $\rho = |\Phi^+\rangle\langle\Phi^+|$ has non–positive partial transposition, as shown in the Example 14 and thus it is not separable.*

Example 16 *One can easily check that the density operator of Example 8 has positive partial transposition, and therefore is separable.*

In higher dimensions, however, the existence of non–decomposable PM's precludes the existence of a simple criterion like this one. In fact, positive partial transposition is a necessary condition (but not sufficient) for separability [61].

5.5 2-Partite Entanglement: Entangled States with Positive Partial Transposition

There are several examples of entangled states with positive partial transposition [62,63,64,65,66]. Here we give one of them.

Example 17 *Let us consider the bound entangled state ρ_b introduced in Ref. [63], where $H_{\mathrm{A}} = H_{\mathrm{B}} = \mathbb{C}^3$. It is defined as $\rho_b = P_b/4$, where P_b is a projector operator onto the orthogonal complement to the subspace spanned by the following vectors:*

$$\begin{aligned}
&|0\rangle \otimes (|0\rangle + |1\rangle), \\
&(|0\rangle + |1\rangle) \otimes |2\rangle, \\
&|2\rangle \otimes (|1\rangle + |2\rangle), \\
&(|1\rangle + |2\rangle) \otimes |0\rangle, \\
&(|0\rangle - |1\rangle + |2\rangle) \otimes (|0\rangle - |1\rangle + |2\rangle).
\end{aligned}$$

It is clear that $P_b^{T_{\mathrm{A}}} = P_b$ and therefore $\rho_b^{T_{\mathrm{A}}} = \rho_b \geq 0$. Now, in order to show that it is entangled we will use the following fact [62,67]: For any decomposition of ρ,

$$\rho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|, \tag{45}$$

the vectors $|\Psi_k\rangle$ must be in the range of ρ. [4] *Otherwise we would have that there would exist some $|\Psi\rangle \in \ker(\rho)$ such that $\langle\Psi|\rho|\Psi\rangle \neq 0$. In our case, the kernel of ρ_b is precisely spanned by the product vectors (45). One can easily check that there exist no product vector $|a, b\rangle$ which is orthogonal to all these vectors, and therefore there is no product vector orthogonal to $\ker(\rho_b)$. Equivalently, there is no product vector in $R(\rho_b)$. Since if ρ_b was separable there should be a decomposition in terms of product vectors, and they should be in its range, we arrive at a contradiction, which indicates that ρ_b is entangled.*

Finally, it is quite remarkable that for Gaussian states in infinite dimensions a general criterion which is more powerful than partial transposition has been recently derived [47].

5.6 Multipartite Entanglement

In the case of more than two parties, the situation is even more complicated than for two parties. As for pure states, one can consider all partitions and define separability with respect to them [68]. For example, a state is separable

[4] Given an operator $\rho \in \mathcal{A}(H)$ the $\ker(\rho) \in H$ is the subspace such that for all $|\Psi\rangle \in \ker(\rho)$ then $\rho|\Psi\rangle = 0$. The range $\mathrm{R}(\rho) \in H$ is the subspace such that for all $|\Psi\rangle \in \mathrm{R}(\rho)$ there exists some $|\Psi'\rangle$ such that $|\Psi\rangle = \rho|\Psi'\rangle$. We have $\ker(\rho) \perp \mathrm{R}(\rho)$ and $H = \ker(\rho) \oplus \mathrm{R}(\rho)$

with respect to the partition (A_1A_3), $(A_2A_5A_6)$, (A_4) if it can be written in the form

$$\rho = \sum_k p_k |\phi_k^1\rangle\langle\phi_k^1| \otimes |\phi_k^2\rangle\langle\phi_k^2| \otimes |\phi_k^3\rangle\langle\phi_k^3|, \tag{46}$$

where $|\phi_k^1\rangle \in H_{A_1} \otimes H_{A_3}$, $|\phi_k^2\rangle \in H_{A_2} \otimes H_{A_5} \otimes H_{A_6}$, and $|\phi_k^3\rangle \in H_{A_4}$. Obviously, there is no general way of knowing whether a state is separable with respect to a given partition or not. As before, the partial transposition criterion provides us with a necessary condition for separability, but it is not sufficient in general.

One idea to study the separability properties of multipartite states is based on a process which is called depolarization [69]. If ρ' can be obtained by local operations out of ρ, and ρ' is separable with respect to a given partition, then ρ is too. This property follows from the fact that we cannot create entanglement by local operations. It suggests to find a family of states to which any state can be transformed by local operations. Indeed, such a family exists. Thus, by studying the separability properties of such a family we can learn about the separability properties of general states. A word of caution here is in order. Note that if ρ' can be obtained from ρ by local operations and ρ' is separable, this does not imply that ρ is separable. For example, we could have an entangled state and transform it into a product vector by a simple measurement. Thus, the family of states provides us with necessary conditions for separability of the original density operator, but not sufficient in general.

Example 18 *(Werner–like states) We consider two systems A and B with corresponding Hilbert spaces of dimension d both. Given a state ρ we depolarize it locally by applying the same random unitary operator to system A and system B. The state after this process will be*

$$\rho' = \int \mathrm{d}\mu(U)(U \otimes U)\rho(U \otimes U)^\dagger, \tag{47}$$

where $d\mu$ is the standard Haas measure for $su(d)$. Applying Schur's Lemma one can easily show that

$$\int \mathrm{d}\mu(U)(U \otimes U)\rho(U \otimes U)^\dagger = \frac{P_s}{d_s}\mathrm{tr}(P_s\rho) + \frac{P_a}{d_a}\mathrm{tr}(P_a\rho), \tag{48}$$

$P_s = (1 + \Pi_{AB})/2$ and $P_a = (1 - \Pi_{AB})/2$ are the projectors onto the symmetric and antisymmetric subspaces (Π_{AB} is the permutation operator and $P_s + P_a = \mathbb{1}$), and $d_s = d(d+1)/2$ and $d_a = d(d-1)/2$ the corresponding dimensions. Thus, we can always depolarize any state to the form

$$\rho_F = F\frac{P_a}{d_a} + (1 - F)\frac{P_s}{d_s}, \tag{49}$$

where $F = \mathrm{tr}(P_a\rho)$. ρ_F is called Werner–like state, since Werner was the first who introduced them for the case of qubits [49]. One can easily show that ρ_F is entangled if and only if $\rho_F^T \not\geq 0$. The 'if' part is a direct consequence of the results of the previous section. In order to prove the 'only if' part, one can simply take

$\rho = p|0,1\rangle\langle 0,1| + (1-p)|0,0\rangle\langle 0,0|$ *and depolarize it according to (47) to obtain a state* ρ' *of the Werner form. Using (48) one can easily see that if* ρ' *has positive partial transposition, then* $p \in [0,1]$, *which means that* ρ' *is separable (since it is the result of depolarizing locally a separable state,* ρ*).*

Example 19 *(multi–qubit states) For the case of qubits, we consider the family of states [68]*

$$\rho_N = \sum_{\sigma=\pm} \lambda_0^\sigma |\Psi_0^\sigma\rangle\langle\Psi_0^\sigma| + \sum_{j=1}^{2^{(N-1)}-1} \lambda_j (|\Psi_j^+\rangle\langle\Psi_j^+| + |\Psi_j^-\rangle\langle\Psi_j^-|). \tag{50}$$

where

$$|\Psi_j^\pm\rangle \equiv \frac{1}{\sqrt{2}}(|j\rangle \otimes |0\rangle \pm |(2^{N-1} - j - 1)\rangle \otimes |1\rangle), \tag{51}$$

and j *is understood in binary notation (e.g, for* $N = 5$, $|j = 5\rangle \otimes |0\rangle$ *stands for* $|01010\rangle$*). Using spin flip and phase shift operations one can depolarize any state of* N *qubits into this form [68]. One can readily check that the partial transpose of this operator with respect to the qubit* A_N *is positive iff* $\Delta \equiv \lambda_0^+ - \lambda_0^- \leq 2\lambda_{2^{N-1}-1}$ *and similarly for the rest of the qubits. Let us describe some of the separability properties of the states (50). We have that if* $\rho_N^{T_{A_k}} \geq 0$ *then it can be written in the form*

$$\rho_N = \sum_i |a_i\rangle_{A_k}\langle a_i| \otimes |\varphi_i\rangle_{\text{rest}}\langle\varphi_i|. \tag{52}$$

On the other hand, if considering all possible partitions of the qubits in two sets it turns out that if for each partition the partial transpose with respect to one of the sets is positive then ρ_N *is fully separable.*

6 Entanglement Distillation

6.1 Introduction

Most of the applications in the field of Quantum Information are based on the use of superpositions of pure states. However, in practice, the state that one has at one's disposal are mixed. For example, if one would like to perform quantum cryptography over long distances using entangled photons, when they arrive at the final location their state will also be entangled with the environment and therefore mixed. The longer the distance the photons have to travel, the more mixed they will become. Unfortunately, if they are significantly mixed, the security of the corresponding cryptographic protocol will not longer be ensured. This fact considerably limits the distances over which one can perform secure quantum cryptography. Fortunately, there is a method that allows us to make the states more pure, and even more entangled. The idea is to use several copies of a state that is not useful for the applications of Quantum Information, but that is still entangled. Using local operations and classical communication it is sometimes possible to obtain fewer copies of particles in a state which is close to a maximally entangled states, for example the state $|\Phi^+\rangle$. This process is called entanglement distillation, and will be the subject of the present section [69,70].

We thus consider the following scenario. Two partners, Alice and Bob, are situated at different locations and share a large number N of pairs of systems, each pair in a state ρ. We will denote the total state of all pairs as $\rho^{\otimes N}$. Then, by using local operations and classical communications they will be able to create another state of $2M$ systems (they probably have to discard some systems), ρ'_M. The goal is to see which operations are needed in such a way that the state ρ'_M is "close" to a maximally entangled state of M particles, $|\Phi\rangle\langle\Phi|^{\otimes M}$. In fact, one would like to find the optimal protocol in the sense that it is the one for which $\rho'_M \to |\Phi\rangle\langle\Phi|^{\otimes M}$ in the limit $N \to \infty$ and with (the averaged) M as big as possible. This will give a way of measuring the entanglement of a state by the (averaged) amount of maximally entangled states that can be distilled from the state in the limit $N \to \infty$.

We will start with the distillation of pure states, and then consider the distillation of mixed states. For some of the distillation protocols we will need to introduce the concept of generalized measurements (see, e.g. [41]).

6.2 Generalized Measurements

Given a system, the measurement process in Quantum Mechanics is described as follows. Each observable O is represented mathematically by a self–adjoint operator $O = O^\dagger$. Let us write it as

$$O = \sum_k o_k P_k, \tag{53}$$

the spectral decomposition of O. Here o_k are real numbers, which correspond to the possible outcomes of the measurements, whereas P_k are orthogonal projector operators onto the (possibly degenerate) eigenspace corresponding to the eigenvalue o_k. They fulfill

$$P_k P_{k'} = \delta_{k,k'} P_k, \quad \sum_k P_k = \mathbb{1}. \tag{54}$$

The probability p_k of obtaining the outcome o_k and the state ρ_k of the system after this outcome (for a filtering measurement) if the initial density operator is ρ will be

$$p_k = \mathrm{tr}(P_k \rho P_k), \tag{55a}$$

$$\rho_k = \frac{1}{p_k} P_k \rho P_k. \tag{55b}$$

This probability and the state after the measurement are completely determined by the operators P_k and the density operator of the system. Thus, we could describe all possible measurements in terms of these operators [that is, we do not need to consider observables but just projection operators fulfilling (54)].

Actually, there are some types of measurements on a system which cannot be described in this way. Those are the so–called generalized measurements. Physically, they can be carried out as follows. We take an auxiliary system (called ancilla) in a well known state $|a\rangle$, then perform an interaction between our system and the ancilla, and then perform a measurement on the ancilla in the form described above (i.e. described by a set of orthogonal projectors). The probability of obtaining certain outcome o_k and the state of the system after this outcome are given by

$$p_k = \mathrm{tr}(A_k \rho A_k^\dagger), \tag{56a}$$

$$\rho_k = \frac{1}{p_k} A_k \rho A_k^\dagger. \tag{56b}$$

The only condition that the operators A_k have to fulfill is

$$\sum_k A_k^\dagger A_k = \mathbb{1}, \tag{57}$$

which is equivalent to imposing that the sum of the probabilities for any outcome, given an arbitrary state of the system, must add up to one. It can be shown that any measurement of the sort described above (with the help of an ancilla) can be described in terms of a set of operators $\{A_k\}$ fulfilling (57). Conversely, given a set of operators fulfilling that property, there always exist a measurement procedure using an ancilla that gives (56a) [41].

Example 20 *Let us consider a qubit in a state $|\Psi\rangle = c|0\rangle + s|1\rangle$, where $|c|^2 + |s|^2 = 1$. We take as ancilla another qubit in the state $|0\rangle_a$. Then we apply the unitary operation which transforms*

$$|0\rangle \otimes |0\rangle_a \to |0\rangle \otimes (\alpha|0\rangle_a + \beta|1\rangle_a), \tag{58}$$

$$|1\rangle \otimes |0\rangle_a \to |0\rangle \otimes (\beta|0\rangle_a + \alpha|1\rangle_a), \tag{59}$$

where $|\alpha|^2 + |\beta|^2 = 1$, and measure σ_z on the ancilla (equivalently, we perform a measurement characterized by the projectors $P_0 = |0\rangle_a\langle 0|$ and $P_1 = |1\rangle_a\langle 1|$). It can be easily shown that the probability of obtaining $o_k = -1, 1$ and the state after the measurement are given by (56a), with

$$A_0 = \alpha|0\rangle\langle 0| + \beta|1\rangle\langle 1|, \tag{60}$$

$$A_1 = \alpha|1\rangle\langle 1| + \beta|0\rangle\langle 0|. \tag{61}$$

These operators fulfill (57). Moreover, for $|\alpha| \neq 0, 1$ it is impossible to find some projection operators which give the same states after the measurement.

Generalized measurements are very important in Quantum Information. This can be understood as follows. Imagine that we have a state $|\Psi\rangle$ and that we would like to have a different one, say $|\Psi'\rangle$. If there exists an operator A fulfilling $|\Psi'\rangle \propto A|\Psi\rangle$ then this means that we can always obtain $|\Psi'\rangle$ out of $|\Psi\rangle$ by performing a generalized measurement in which one of the operators A_k is proportional to A. Note that if we impose additional conditions to the operator A, such for example that it is local (i.e. only acts on one subsystem), it may be not possible to find it such that it accomplishes the desired task. One example of this is precisely the one mentioned in a previous section, in which $|\Psi'\rangle$ is entangled whereas $|\Psi\rangle$ is a product vector.

6.3 Distillation of Pure States

We consider that Alice and Bob share N copies of a state $|\Psi\rangle \in H_\mathrm{A} \otimes H_\mathrm{B}$. We will consider the case of qubits, although one can generalize the methods discussed here to higher dimensions. They are allowed to perform measurements and operations on their particles and communicate the results. We will consider different scenarios. In the first one, Alice and Bob are allowed to perform measurements on a single copy at a time [71,72,73,74,75]. If they succeed, then they will have produced a maximally entangled state (of two qubits). If they fail, they will produce a product state. If they repeat the same procedure with N copies, the number of maximally entangled state that they will obtain in average will be $M = pN$, where p is the success probability. Thus, the efficiency of the process will be

$$D^{(1)}(\Psi) = \lim_{N\to\infty} \frac{M}{N} = p. \tag{62}$$

We will consider the optimal protocol in the sense that p is maximal. Then we will consider that they are allowed to act on two copies at the same time. With a probability p_d they will be able to obtain a maximally entangled state in $d \leq 4$ dimensions. If we impose that the entanglement of pure states must be additive, we have that two maximally entangled states of two qubits are equivalent to one maximally entangled state of one four–level system, i.e. the entanglement of a maximally d–level system must be $\log_2 d$. The efficiency of the process will be

$$D^{(2)}(\Psi) = \frac{1}{2} \sum_{k=2}^{4} p_k \log_2(k). \tag{63}$$

This quantity measures the number of qubits in a maximally entangled state that we obtain per copy of $|\Psi\rangle$. The optimal procedure is the one that maximizes this figure of merit. We can proceed in the same way with three, four, ..., n copies to obtain $D^{(n)}(\Psi)$. Obviously, $D^{(n)}(\Psi) \geq D^{(n-1)}(\Psi)$. In fact, one can show that

$$\lim_{n\to\infty} D^{(n)}(\Psi) = E(\Psi), \tag{64}$$

so that the quantification of the entanglement that we did in Section 4 acquires a physical meaning [43]. It is the maximum amount of qubits in a maximally entangled state per qubit that we can obtain using the optimal distillation protocol. Actually, the formula (64) also holds for any dimensions, and not only for qubits.

One can also show that, in the above limit, the optimal procedure becomes reversible in the following sense [44]. If we have M copies of the maximally entangled state $|\Phi^+\rangle$ by local operations we can obtain N copies of a state $|\Psi_M\rangle$ such that in the limit $M \to \infty$ we have $N \to M/E(\Psi)$ and $|\langle\Psi_M|\Psi\rangle^{\otimes M}| \to 1$. That is, we can dilute entanglement without losing it in the asymptotic limit of an infinite number of copies. This means, for example, that pure-state entanglement can be stored in the form of maximally entangled states of qubits and then transformed into the form that is needed for applications.

We consider a state of two qubits whose Schmidt decomposition is $|\Psi\rangle = c|u_0, v_0\rangle + s|u_1, v_1\rangle$, where $c \geq s \geq 0$ and $c^2 + s^2 = 1$. Since we are allowed to use local operations, we can always transform this state into the state

$$|\Psi_0\rangle = c|0,0\rangle + s|1,1\rangle \tag{65}$$

by simply applying a unitary transformation which produces $|u_0\rangle \to |0\rangle$, and $|u_1\rangle \to |1\rangle$ (and similarly for Bob's qubit). Thus, without loss of generality we can take the state in the form (65).

One Copy

We can use generalized measurements to distill the state (65). We just have to find a local operator A_0 that transforms it into the state $|\Phi^+\rangle$. One easily finds

$$A_0 = k\left(|1\rangle\langle 1| + \frac{s}{c}|0\rangle\langle 0|\right). \tag{66}$$

Imposing (57) we obtain that $k \leq 1$ and that $A_1 = U(\mathbb{1} - A_0^\dagger A_0)^{1/2}$, where U is an arbitrary unitary operator. The state after the measurement can be easily determined using (56a) obtaining that the maximal probability of obtaining $|\Phi^+\rangle$ occurs for $k = 1$ (and it is independent of U), and it is

$$p = 2s^2. \tag{67}$$

Thus, we have that $D^{(1)}(\Psi_0) = 2s^2$. Clearly, the more entangled the initial state is, the most efficient the process becomes. As it should occur, if the initial state is not entangled ($s = 0$), no entanglement can be produced.

In this simple case we can also analyze the entanglement dilution. If we had initially the state $|\Phi^+\rangle$ and we want to obtain $|\Psi\rangle$, we can use the (local) generalized measurements with

$$A_0 = (c|0\rangle_{\mathrm{A}}\langle 0| + s|1\rangle_{\mathrm{A}}\langle 1|) \otimes (|0\rangle_{\mathrm{A}}\langle 0| + |1\rangle_{\mathrm{A}}\langle 1|), \tag{68}$$

$$A_1 = (s|1\rangle_{\mathrm{A}}\langle 0| + c|0\rangle_{\mathrm{A}}\langle 1|) \otimes (|1\rangle_{\mathrm{A}}\langle 0| + |0\rangle_{\mathrm{A}}\langle 1|). \tag{69}$$

It is easy to check that whatever the outcome of the measurement is, the produced state is the desired one. This means that we can always dilute entanglement of qubits with unit probability. Note also that for the measurement local communication is required. Actually, the measurement can be performed by Alice, and if she obtains the outcome 1, then she and Bob change the qubit according to $|0\rangle \leftrightarrow |1\rangle$.

Two Copies and More

Defining $|\tilde{0}\rangle = |0,0\rangle, \dots |\tilde{3}\rangle = |1,1\rangle$ we can write

$$|\Psi\rangle^{\otimes 2} = c^2|\tilde{0},\tilde{0}\rangle + sc(|\tilde{1},\tilde{1}\rangle + |\tilde{2},\tilde{2}\rangle) + s^2|\tilde{3},\tilde{3}\rangle. \tag{70}$$

The best generalized measurement to start with is the one for which Alice measures and [73]

$$A_0 = |\tilde{3}\rangle\langle\tilde{3}| + \frac{s}{c}(|\tilde{2}\rangle\langle\tilde{2}| + |\tilde{1}\rangle\langle\tilde{1}|) + \frac{s^2}{c^2}|\tilde{0}\rangle\langle\tilde{0}|, \tag{71}$$

$$A_1 = (\mathbb{1} - A_0^\dagger A_0)^{1/2}. \tag{72}$$

If Alice obtains 0 she has a maximally entangled state in $d = 4$ dimensions. If she obtains 1, the produced state is still entangled and has a similar form to the one (70), but where $|\tilde{3}\rangle$ is absent. We can also find the best generalized measurement such that again we obtain with the maximal probability a maximally entangled states in $d = 3$ dimensions. If the opposite outcome is obtained, another generalize measurement can be found so that with certain probability the state $|\Phi^+\rangle$ is produced. The determination of the corresponding probabilities is tedious, but one obtains [71] that $D^{(2)}(\Psi) > D^{(1)}(\Psi)$ provided $s \neq 0, 1/\sqrt{2}$.

The above procedure can be easily extended to the case in which one can act on n copies simultaneously. The value of $D^{(n)}$ obtained in this form is given by (64) [43].

6.4 Distillation of Mixed States

We consider that Alice and Bob share N copies of a mixed state ρ of two qubits. The goal is to produce a state which is close to a maximally entangled state of a high-dimensional system by using local operations and classical communication. As before, we can consider different scenarios. The difference now will be that Alice and Bob will not be able to obtain a pure state by processing a finite number of copies, in general. But still, we can characterize, as before, the efficiency of the process by the yield in the limit $N \to \infty$, when they are able to manipulate one copy, two copies, etc, at a time. The distillable entanglement $D(\rho)$ is defined as the optimal yield over all possible strategies involving manipulations of an arbitrary number of copies (see [76]).

As one could expect after our experience with the problem of separability, this problem is much more difficult for mixed states than for pure states. In fact, except for some very special cases [77], we do not know how to find the optimal way of distilling mixed states, i.e. we do not know how to determine the distillable entanglement. Here we will describe two distillation protocols [69,70] which, although they are not optimal and although they only apply to certain classes of states, are the most relevant ones in the sense that other protocols can be obtained by generalizing them.

One Copy: Filtering [70]

We consider the family of states

$$\rho(F) = F|\Psi^-\rangle\langle\Psi^-| + (1-F)|1,1\rangle\langle 1,1|. \tag{73}$$

Here, $F \in [0,1]$ is the fidelity, i.e. the overlap of the density operator with the maximally entangled states $|\Psi^-\rangle$,

$$F = \langle\Psi^-|\rho(F)|\Psi^-\rangle. \tag{74}$$

One can readily check, using the partial transposition criterion, that this state is entangled unless $F = 0$.

In order to distill, Alice and Bob perform the generalized measurement described by the operators given in Example 20

$$A_0 = \alpha|0\rangle\langle 0| + \sqrt{1-\alpha^2}|1\rangle\langle 1|, \tag{75}$$
$$A_1 = \alpha|1\rangle\langle 1| + \sqrt{1-\alpha^2}|0\rangle\langle 0|, \tag{76}$$

where $\alpha \in (0,1)$. If both obtain the outcome '0', they keep the state, and otherwise they throw it away (and take another pair). We have

$$A_0 \otimes A_0|\Psi^-\rangle = \alpha\sqrt{1-\alpha^2}|\Psi^-\rangle, \tag{77}$$
$$A_0 \otimes A_0|1,1\rangle = (1-\alpha^2)|1,1\rangle, \tag{78}$$

so that the probability that both of them obtain '0' is

$$p_{00} = (1-\alpha^2)[F\alpha^2 + (1-F)(1-\alpha^2)], \tag{79}$$

whereas the state after this outcome is $\rho'(F) = \rho(F')$ with

$$F' = \frac{F\alpha^2}{F\alpha^2 + (1-F)(1-\alpha^2)}. \tag{80}$$

We see that for $\alpha \to 1$, we have that $F' \to 1$ and therefore we can have a fidelity as close to one as we wish. Note that $p_{00} \to 0$ as $\alpha \to 1$ meaning that if we want a higher fidelity, the probability becomes smaller. In fact, we see that this purification protocol produces zero yield since in order to obtain 1 maximally entangled state we will need to use infinitely many copies. Nevertheless, this method can be very useful in practice. Imagine that in order to perform secure quantum cryptography we need a fidelity of 0.9, whereas we have pairs in a state of the form (73) with $F = 0.6$ we can choose $\alpha = 0.93$ and obtain $F' > 0.9$ with a probability larger than 1/12.

Two Copies: Using Controlled–NOTS [69,78]

Let us consider the two-qubit Werner state (49),

$$\rho_F = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}(|\Psi^+\rangle\langle\Psi^+| + |\Phi^-\rangle\langle\Phi^-| + |\Phi^+\rangle\langle\Phi^+|) \tag{81}$$

where all the vectors appearing here are Bell states (12a).

In the present scenario, Alice and Bob share two pairs of qubits in the state (81). Let us denote by A_1 and A_2 Alice's particles and by B_1 and B_2 Bob's, so that their state is $\rho_F \otimes \rho_F$. The distillation procedure proceeds as follows. First, Alice applies the unitary transformation σ_y to her two qubits. This transforms in each pair $\varPsi^{\pm} \leftrightarrow \varPhi^{\mp}$. Thus, the state after this rotation will be

$$\rho'_F = F|\varPhi^+\rangle\langle\varPhi^+| + \frac{1-F}{3}(|\varPhi^-\rangle\langle\varPhi^-| + |\varPsi^+\rangle\langle\varPsi^+| + |\varPsi^-\rangle\langle\varPsi^-|). \tag{82}$$

The state (81) that had the maximum contribution coming from $\varPsi^-$ becomes now one with maximum contribution of $\varPhi^+$. Then, both apply locally a controlled not operation to their two particles, where A_1 and B_1 act like sources, and the other two (A_2 and B_2) as targets. The control–NOT operation acts as follows:

$$|0\rangle_{A_1}|0\rangle_{A_2} \to |0\rangle_{A_1}|0\rangle_{A_2}, \tag{83a}$$
$$|0\rangle_{A_1}|1\rangle_{A_2} \to |0\rangle_{A_1}|1\rangle_{A_2}, \tag{83b}$$
$$|1\rangle_{A_1}|0\rangle_{A_2} \to |1\rangle_{A_1}|1\rangle_{A_2}, \tag{83c}$$
$$|1\rangle_{A_1}|1\rangle_{A_2} \to |1\rangle_{A_1}|0\rangle_{A_2}, \tag{83d}$$

and similarly with B. The state of the particles after this operation can be derived using the following table:

Initial state		Final state	
Sources	Targets	Sources	Targets
$\varPhi^{\pm}$	$\varPhi^{+}$	$\varPhi^{\pm}$	$\varPhi^{+}$
$\varPhi^{\pm}$	$\varPhi^{-}$	$\varPhi^{\mp}$	$\varPhi^{-}$
$\varPsi^{\pm}$	$\varPsi^{+}$	$\varPsi^{\pm}$	$\varPhi^{+}$
$\varPsi^{\pm}$	$\varPsi^{-}$	$\varPsi^{\mp}$	$\varPhi^{-}$
$\varPhi^{\pm}$	$\varPsi^{+}$	$\varPhi^{\pm}$	$\varPsi^{+}$
$\varPhi^{\pm}$	$\varPsi^{-}$	$\varPhi^{\mp}$	$\varPsi^{-}$
$\varPsi^{\pm}$	$\varPhi^{+}$	$\varPsi^{\pm}$	$\varPsi^{+}$
$\varPsi^{\pm}$	$\varPhi^{-}$	$\varPsi^{\mp}$	$\varPsi^{-}$

Alice and Bob then measure the state of their target particle in the basis $\{|0\rangle, |1\rangle\}$ (i.e., they measure $\sigma_z^{A_2}$ and $\sigma_z^{B_2}$) and broadcast their results. If the results are the same, they keep the source particles, and otherwise they discard them. This amounts to taking only the states which had as a result $\varPhi^{\pm}$ in the target bits (i.e. only considering the first four rows of the above table). Using the table, one sees that this is equivalent to projecting the initial states onto the subspace in which either both the sources and the targets are $\varPhi$ states or both are $\varPsi$ states. Let us calculate the projection of the new density operator onto the state $|\varPhi^+\rangle$ in the case the measurements were successful. We have the following possibilities:

- With probability F^2 the initial state was $|\varPhi^+\rangle_1|\varPhi^+\rangle_2$. In this case the final state of the source will be the desired one.

- With probability $(1-F)^2/9$ the initial state was $|\Phi^-\rangle_1|\Phi^-\rangle_2$. In this case the final state of the source will be the desired one.
- With probability $F(1-F)/3$ the initial state was $|\Phi^-\rangle_1|\Phi^+\rangle_2$. With the same probability the initial state was $|\Phi^-\rangle_1|\Phi^+\rangle_2$. In both cases, the final state of the source will not be the desired one.
- The other 4 possible initial state, products of $|\Psi^\pm\rangle_1$ with $|\Psi^\pm\rangle_2$ have a probability $(1-F)^2/9$ each.

Thus, the probability of having at the end of the process the state $|\Phi^+\rangle_1$ is

$$F' = \frac{F^2 + (1-F)^2/9}{F^2 + 2F(1-F)/3 + 5(1-F)^2/9}. \tag{84}$$

For $1 > F > 1/2$, we have that $F' > F$. Therefore, the fidelity after this operation increases. To finish the process and leave the states in a Werner state, so that they can continue this process Alice applies the operation σ_y to her source particle, which transforms $|\Phi^+\rangle \to |\Psi^-\rangle$. In summary, if the process is successful, Alice and Bob are left with a single pair in a Werner state but with fidelity F'. Then, they can take two successful pairs and repeat the same procedure to obtain a higher fidelity. By proceeding in this way they can reach a fidelity as close to one as they wish, but at the expenses of waisting many pairs. Again, the asymptotic yield of this procedure is zero. In Fig. 1 we have ploted F' as a function of F and show how the fidelity increases as one repeates it with the successful pairs.

Many Copies

It is possible to device purification protocols which process n copies and such that when $n \to \infty$ give a finite yield [69]. However, it is not know what is the

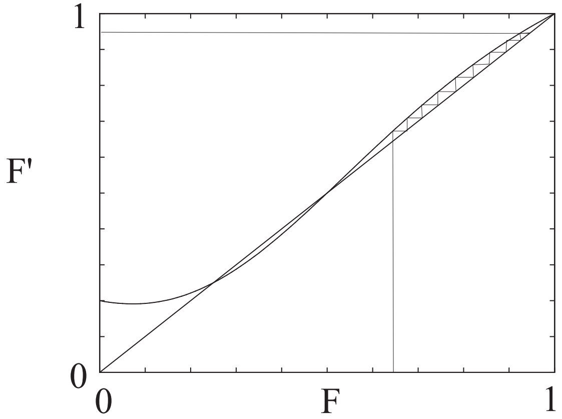


Fig. 1. New fidelity in terms of the old fidelity for the purification protocol based on controlled–NOTS. Successive applications lead to a fidelity as close to one as one wishes.

distillable entanglement for general states since it is not known how to optimize the procedures.

Distillability

If a state cannot be distilled, then it will most likely be useless for the applications in Quantum Information, since they require states close to maximally entangled states. Thus, a relevant question in this context is: which states can be distilled? Or, equivalently, given many copies of a state ρ, can we distill maximally entangled states? This is the so–called distillability problem, and it is as relevant (and difficult) as the separability problem studied in Section 5.

In principle, one could suspect that all entangled states can be distilled. However, this is not the case in general. The entangled states that cannot be distilled are called bound entangled states [79].

Example 21 *(Bound entangled states [79]) If $\rho^T \geq 0$ and ρ is entangled then ρ is a bound entangled state. This can be shown by contradiction. If ρ is distillable, it means that there exist some physical action on Alice's and Bob's sides such that a state ρ' close to a maximally entangled state is produced out of N copies of ρ. Mathematically, this means that there exist two operators A and B acting on $H_{\mathrm{A}}^{\otimes N}$ and $H_{\mathrm{B}}^{\otimes N}$, respectively such that*

$$(A \otimes B)\rho^{\otimes N}(A \otimes B)^\dagger \propto \rho'. \tag{85}$$

But since ρ' is close to a maximally entangled state, it must have a non–positive partial transposition. Thus, the lhs of (85) must have it too. But

$$[(A \otimes B)\rho^{\otimes N}(A \otimes B)^\dagger]^{T_{\mathrm{A}}} = (A^{T_{\mathrm{A}}} \otimes B)^\dagger(\rho^{T_{\mathrm{A}}})^{\otimes N}(A^{T_{\mathrm{A}}} \otimes B) \geq 0 \tag{86}$$

which leads to a contradiction.

This shows that entangled states with positive partial transposition, as those studied in Section 5.5 are always bound entangled states. What about those which do not fulfill this property? This question has not been solved yet, but let us show that we can restrict ourselves to the family of states (49) [79,3,82]. Let us assume that $\rho^T \ngeq 0$. That is, there exists some $|\Psi\rangle \in H_{\mathrm{A}} \otimes H_{\mathrm{B}}$ such that $\langle\Psi|\rho^T|\Psi\rangle < 0$. Let us write the Schmidt decomposition (14) as $|\Psi\rangle = A \otimes B|\Phi^+\rangle$, where

$$A = \sqrt{d_k}|u_k\rangle_{\mathrm{A}}\langle k|, \tag{87}$$

$$B = \sqrt{d_k}|v_k\rangle_{\mathrm{B}}\langle k|, \tag{88}$$

$$|\Phi^+\rangle = \sum_k |k, k\rangle. \tag{89}$$

Thus we have that $\langle\Phi^+|\tilde\rho^T|\Phi^+\rangle < 0$, where $\tilde\rho = (A^T \otimes B)^\dagger\rho(A^T \otimes B)$ can be obtained from ρ by local operations. Now, using the fact that

$$(U \otimes U^*)|\Phi^+\rangle\langle\Phi^+|(U \otimes U^*)^\dagger = |\Phi^+\rangle\langle\Phi^+|, \tag{90}$$

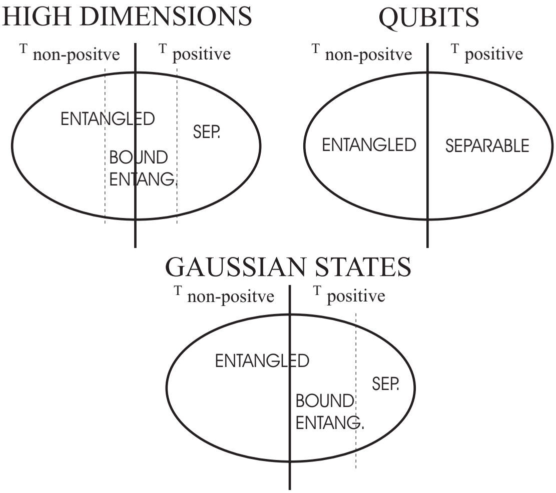


Fig. 2. Diagrammatic representation of the current status of the problems of separability and distillability. The dashed lines represent that we do not know where the border between the corresponding sets lies.

we have

$$0 > \langle \Phi^+|\tilde{\rho}^T|\Phi^+\rangle = \int d\mu(U)\langle \Phi^+|(U\otimes U^*)\rho^T(U\otimes U^*)^\dagger|\Phi^+\rangle = \langle \Phi^+|\rho_F|\Phi^+\rangle, \quad (91)$$

where ρ_F is of the Werner form (49) and we have used (48). What we have shown is that if ρ has non–positive partial transposition, it can always be depolarized to a Werner state which still has non–positive partial transposition. Thus, in order to study whether all states with non–positive partial transposition are distillable we can restrict ourselves to Werner states.

For the case of finite dimensions in which one of the systems is a qubit [62,82], and in the case of infinite dimensions with Gaussian states it turns out that a state is distillable if and only if it has a non–positive partial transposition [80]. For systems in higher (but finite) dimensions, however, although nothing has been rigorously proved, there is a strong evidence that there exist Werner states that despite having non–positive partial transposition, are not distillable [81,82].

Thus, what is known up to now (April 2001) about separability and distillability of two quantum system is summarized in Fig. 2

6.5 Multipartite Distillation

Finally, just to mention that the problem of multipartite distillation is even more complicated than the one of two parties. Again, one can use the depolarization method to show that a state is distillable [83].

Acknowledgments

I thank B. Kraus for carefully reading this manuscript. This work was supported by the Austrian Science Foundation (SFB "Control and measurement of coherent quantum systems"), the European Union project EQUIP (contract IST-1999-11053), the TMR european network ERB–FMRX–CT96–0087, the ESF under the PESC program "Quantum Information", and the Institute for Quantum Information GmbH.

References

1. See, for example, Nielsen and Chuang *Quantum computation and quantum information*, (Cambridge University Press, Cambridge, 2000); *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*, Edited by D. Bouwmeester, A. Ekert and A. Zeilinger (Springer–Verlag, Berlin, 2000).
2. Most of the papers on quantum information can be found in http: //xxx.lanl.gov/ archive / quant-ph.
3. See, for example, D. P. DiVincenzo, Fortschr. Phys. **48**, 771 (2000).
4. J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74** 4091 (1995).
5. J. F. Poyatos, J. I. Cirac and P. Zoller, Phys. Rev. Lett. **81** 1322 (1998).
6. J. I. Cirac and P. Zoller, Nature **404** 579 (2000); T. Calarco, J.I. Cirac and P. Zoller, Phys. Rev. A (in press), quant-ph/0010105.
7. A. Sorensen and K. Molmer, Phys. Rev. Lett. **82**, 1971 (1999).
8. K. Molmer and A. Sorensen, Phys. Rev. Lett. **82**, 1835 (1999).
9. S. Schneider, D. F. V. James, and G. J. Milburn, J. Mod. Opt. **47**, 499 (2000).
10. D. Jonathan and M. B. Plenio, quant-ph/0103140.
11. C. Monroe *et al*, Phys. Rev. Lett. **75** 4714 (1995).
12. C. A. Sackett *et al.*, Nature **404**, 256 (2000).
13. M. A. Rowe *et al.*, Nature **409**, 791 (2001).
14. Ch. Roos *et al.*, Phys. Rev. Lett. **83**, 4713 (1999).
15. T. Sleator and H. Weinfuhrter, Phys. Rev. Lett. **74**, 4087 (1995)
16. T. Pellizari, *et al.*, Phys. Rev. Lett. **75**, 3788 (1995).
17. Q. A. Turchette *et al*, Phys. Rev. Lett. **75** 4710 (1995).
18. X. Maître *et al*, Phys. Rev. Lett. **79** 769 (1997).
19. Rauschenbeutel, A. *et al.*, *Science* **288**, 2024 (2000).
20. S. Brattke *et al.*, Phys. Rev. Lett. **86**, 3534 (2001).
21. D. Jaksch *et al.*, Phys. Rev. Lett. **82**, 1975 (1999); T. Calarco *et al.*, Phys. Rev. A **61**, 022304 (2000).
22. D. Jaksch *et al.*, Phys. Rev. Lett. **85**, 2208 (2000).
23. G. K. Brennen *et al.*, Phys. Rev. Lett. **82**, 1060 (1999).
24. A. Beige *et al.*, J.Mod.Opt. **47**, 401 (2000).
25. B. E. Kane, Nature **393** 133 (1998).
26. D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).
27. Y. Makhlin and G. Schön, Nature **398**, 305 (1999).
28. D. G. Cory, A. F. Fahmy and T. F. Havel, Proc. Natl. Acad. Sci. **94** 1634 (1997).
29. N. A. Gershenfeld and I. L. Chuang, Science **275** 350, (1997).
30. S. L. Braunstein *et al.*, Linden, ... Phys. Rev. Lett. **83**, 1054 (1999)

31. See, for example, D. F. James, Fortschr. Phys. **48**, 823 (2000); see also, A. M. Steane and D. M. Lucas, Fortschr. Phys. **48**, 839 (2000).
32. P. W. Shor, Phys. Rev. A **52** 2493 (1995).
33. A. M. Steane, Phys. Rev. Lett. **77** 793 (1996).
34. R. Laflamme *et al.*, Phys. Rev. Lett. **77**, 198-201 (1996).
35. D. Gottesman, Phys. Rev. A **54**, 1862-1868 (1996)
36. P. Shor, in 37th Symposium on Foundations of Computing, IEEE Computer Society Press, pp. 56–65 (1996).
37. J. I. Cirac, T. Pellizari and P. Zoller, Science **273**, 1207 (1996).
38. D. Gottesman, Phys. Rev. A **57**, 127 (1998)
39. J. Preskill, Proc. R. Soc. **454** 385 (1998).
40. A. Steane, Fortsch.Phys. **46**, 443 (1998).
41. A very good introduction to the main concepts of quantum information can be found in A. Peres, *Quantum Theory: Concepts and Methods* Edited by Kluwer Academic, 1993.
42. J. S. Bell, Physics **1**, 195 (1964).
43. C. H. Bennett *et al.*, Phys. Rev. A **53**, 2046 (1996).
44. S. Popescu and D. Rohrlich, Phys. Rev. A **56** 3319 (1997).
45. D. M. Greenberger, M. Horne, and A. Zeilinger, in *Bell's theorem, Quantum Theory and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht, The Netherlands, 1989), pp 69.
46. N. Linden and S. Popescu, Fortsch.Phys. **46**, 567 (1998).
47. G. Giedke *et al.*, quant-ph/0104050.
48. W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 62314 (2000).
49. R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
50. M. Lewenstein *et al.*, J. Mod. Opt. **77**, 2481 (2000).
51. B. M. Terhal, quant-ph/0101032.
52. R. Horodecki, P. Horodecki, M. Horodecki, Phys. Lett. A **210** 377 (1996).
53. J.F. Clauser *et al.*, Phys. Rev. Lett. **23**, 880 (1969).
54. See, for example, W. Rudin, *Functional analysis* (Mac Graw–Hill, New York, 1991).
55. E. Stromer, Acta Math. **110**, 233 (1963).
56. S. L. Woronowicz, Rep. Math. Phys. **10**, 165 (1976).
57. B. M. Terhal, quant-ph/9810091; see also quant-ph/9911057.
58. M. Lewenstein *et al.*, Phys. Rev. A **62**, 52310 (2000).
59. M. Lewenstein *et al.*, quant-ph/0005112.
60. A. Jamiolkowski, Rep. of Math. Phy. No. 4, **3** (1972).
61. A. Peres, Phys. Rev. Lett **77**, 1413 (1996).
62. P. Horodecki, Phys. Lett. A **232**, 333 (1997).
63. C. H. Bennett *et al.*, **82**, 5385 (1999).
64. P. Horodecki and M. Lewenstein, Phys. Rev. Lett. **85**, 2657 (2000).
65. R. F. Werner and M. M. Wolf, Phys. Rev. Lett. **86**, 3658 (2001).
66. D. Bruss and A. Peres, Phys. Rev. A **61**, 030301 (2000).
67. B. Kraus *et al.* Phys. Rev. A **61**, 062302 (2000).
68. W. Dür, J. I. Cirac, and R. Tarrach, Phys. Rev. Lett. 83, 3562 (1999); W. Dür and J. I. Cirac, Phys. Rev. A **61**, 042314 (2000).
69. C. H. Bennett et al., Phys. Rev. Lett. **76**, 722 (1996); C. H. Bennett, et al., Phys. Rev. A **54**, 3824 (1996).
70. N. Gisin, Phys. Lett. A **210**, 151 (1996).
71. H.-K. Lo and S. Popescu, Phys. Rev. A **63**, 022301 (2001).
72. M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

73. G. Vidal, Phys.Rev.Lett. **83**, 1046 (1999).
74. D. Jonathan, M. B. Plenio, Phys. Rev. Lett. **83**, 1455 (1999); see also erratum Phys. Rev. Lett. **84**, 4781 (2000).
75. L. Hardy, Phys. Rev. A **60**, 1912 (1999)
76. E. M. Rains, Phys. Rev. A **60**, 173 (1999).
77. E. M. Rains, Phys. Rev. A **60**, 179 (1999).
78. David Deutsch *et al.*, Phys. Rev. Lett. **77**, 2818 (1996).
79. M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
80. G. Giedke *et al.*, quant-ph/0104072.
81. D. DiVincenzo *et al.*, Phys. Rev. A **61**, 62312 (2000).
82. W. Dür *et al.*, Phys. Rev. A **61**, 62313 (2000).
83. W. Dür and J. I. Cirac, Phys. Rev. A **62**, 22302 (2000).

Spintronics, Quantum Computing, and Quantum Communication in Quantum Dots [*]

Guido Burkard, Hans-Andreas Engel, and Daniel Loss

Abstract. The coherent manipulation, filtering, and measurement of the electron spin in solid-state nanostructures has potential applications for both conventional and quantum computation as well as for quantum communication. This article is intended as a review of our proposal to use electron spins in quantum confined structures as quantum bits (qubits). The physical requirements for implementing a quantum computer, including single- and two-qubit gate operations, phase coherence, initialization, and read-out, will be discussed. In addition, we also present recently proposed schemes for using a single quantum dot as spin-filter and spin-memory device. In the context of spintronics, it is quite natural to consider spin-entangled electron pairs as a basic resource for quantum communication; we show that the entanglement of such EPR pairs can be detected in mesoscopic transport measurements using metallic as well as superconducting leads attached to the dots.

1 Introduction

Recent spin-related experiments [1,2,3,4,5,6,7] show that electron spins are potentially useful for information processing and transmission. In quantum-confined nanostructures (e.g. quantum wells), unusually long spin dephasing times (approaching microseconds) have been reported [2,3,4], as well as phase-coherent spin transport over long distances (up to 100 μm) [2]. Besides the intrinsic theoretical interest in spin-related phenomena, two main areas are promising for future applications: conventional devices based on the electron spin [1] as well as spin-based quantum computer hardware [8,9]. It can be expected that the electron spin can also enhance the performance in conventional computers in various ways, examples being spin-transistors (based on spin-currents and spin injection), non-volatile memories, single spin as the ultimate limit of information storage [1]. These devices are not yet available, and experimental progress as well as theoretical investigations are needed to provide guidance and support in the search for realizable implementations. On the other hand, the emerging field of quantum computing [10,11] and quantum communication [11,12] requires a radically new approach to the design of the necessary hardware. It was first pointed out in [8] that the spin of the electron is a most natural candidate for the qubit–the fundamental unit of quantum information. These spin qubits [8], when located in quantum-confined structures (e.g. semiconductor quantum dots,

[*] The present article contains the basic material of the lectures given by D. Loss at the 13th Chris Engelbrecht Summer School, Stellenbosch, South Africa, 24 January to 2 February 2001.

W. Dieter Heiss (Ed.): LNP 587, pp. 241–265, 2002.
© Springer-Verlag Berlin Heidelberg 2002

atoms, or molecules) satisfy all requirements needed for a scalable quantum computer. Being attached to an electron with orbital degrees of freedom, spin qubits can in principle even be transported along conducting wires between different subunits in a quantum network [11]. In particular, spin-entangled electrons can be created in coupled quantum dots, or as recently suggested, by connecting a superconductor to a normal metal [13] or to a Luttinger liquid [14]. These spin-entangled electrons can then act as mobile Einstein-Podolsky-Rosen (EPR) pairs [11], providing the necessary resources for quantum communication.

The abundance of proposals for solid state implementations of quantum computers can be explained by the versatility of solid state physics, in that almost any phenomenon possible in physics can be embodied in an appropriately designed condensed matter system [15]. Furthermore, the interplay between solid state physics and (computer) technology has brought about an amazing progress in the fabrication of small artificial structures and devices. It appears natural to expect that this versatility will also extend to the creation of solid-state quantum computers. However, only the experimental progress will tell whether any of these proposals will actually provide a successful route to a quantum computer.

This review is intended to give an overview of the current status of our theoretical efforts towards the goal of implementing quantum computation and quantum communication with electron spins in quantum-confined nanostructures. The topics presented here have been discussed in greater detail in various research articles, to which we refer the interested reader.

1.1 Quantum Computing

The recent interest in quantum computing has is mainly due to the possibility of outperforming classical computation through new and more powerful quantum algorithms such as the Shor algorithm [16] for factoring large integers and the Grover algorithm [17] for searching unsorted databases. In addition to these two algorithms and their modifications, there is a growing list of other "quantum tasks" [11,12] such as quantum key distribution for cryptography, error correcting schemes, quantum teleportation, etc., indicating even more the desirability of experimental implementations of a quantum information processor. The basic unit of information in a quantum computer is the quantum bit (qubit) which is physically represented as the state of a quantum two-level system (e.g. a spin 1/2). It is known that quantum algorithms can be implemented by concatenating one- and two-qubit gates [18]. There is a growing number of proposed physical implementations of qubits and quantum gates e.g. trapped ions [19], cavity QED [20], nuclear spins [21,22,23], superconducting devices [24], and our qubit proposal [8] based on the spin of the electron in quantum-confined nanostructures.

1.2 Quantum Communication

Coupled quantum dots provide a powerful source of deterministic entanglement between qubits of localized but also of delocalized electrons [11,8]. For example, with such quantum gates it is possible to create a singlet state out of two electrons and subsequently separate (by electronic transport) the two electrons spatially with the spins of the two electrons still being entangled–the prototype of an EPR pair. Implementations of an electron-spin entangler that make use of the s-wave (spin singlet) nature of conventional superconductors were proposed in [13,14]. The availability of an entangler for electron spins would make it possible to study a new class of quantum phenomena in electronic nanostructures [11] such as the entanglement and non-locality of electronic EPR pairs, tests of Bell inequalities, quantum teleportation [25], and quantum cryptography [26] which promises secure information transmission.

1.3 General Remarks on Quantum Dots

Semiconductor quantum dots are structures where charge carriers are confined in all three spatial dimensions, the dot size being of the order of the Fermi wavelength in the host material, typically between 10 nm and 1 μm [27]. The confinement is usually achieved by electrical gating of a two-dimensional electron gas (2DEG), possibly combined with etching techniques, see Fig. 1. Precise control of the number of electrons in the conduction band of a quantum dot (starting from zero) has been achieved in GaAs heterostructures [28]. The electronic spectrum of typical quantum dots can vary strongly when an external magnetic field is applied [27,28], since the magnetic length corresponding to typical laboratory fields $B \approx 1\,\mathrm{T}$ is comparable to typical dot sizes. In coupled quantum dots Coulomb blockade effects [29], tunneling between neighboring dots [27,29], and magnetization [30] have been observed as well as the formation of a delocalized single-particle state [31].

2 Phase Coherence

Quantum computation is essentially a quantum-mechanical (hence phase coherent) time evolution of a set of qubits (the quantum register or quantum memory). Therefore, the very difference between quantum and classical computing lies in the phase coherence. The issue of decoherence, i.e. the loss of phase coherence, is thus of high importance for quantum computing. Fundamental research in mesoscopic physics has traditionally been focusing on characterizing and understanding the decoherence of electrons in small structures. However, most theoretical and experimental work (say, in weak localization studies or the Aharonov-Bohm effect) is devoted to the *orbital* coherence of electron states, that is, the preservation of the relative phase of superpositions of spatial states of the electron (e.g., in the upper and lower arm of an Aharonov-Bohm ring). The coherence times seen in these investigations are almost completely irrelevant to the *spin* coherence times which are important in our quantum computer proposal. Some relation between the two coherence times exists if there are strong spin-orbit effects, but our intention is that conditions and materials should be chosen such that these effects are weak.

For weak spin-orbit interaction the spin coherence times (the time over which the phase of a superposition of spin-up and spin-down states is well-defined) can be completely different from the charge coherence times (typically a few nanoseconds at low temperatures), and it is known that they can be orders of magnitude longer. Given the importance of phase coherence for quantum information processing, the rather long spin decoherence times in semiconductor structures were among the main motivations for proposing spin [8] rather than charge as the qubit in these structures.

Long spin coherence times were recently reported in doped GaAs in the bulk and in a 2DEG from magneto-optical experiments based on time-resolved Faraday rotation measurements [2]. In the limit of a vanishing magnetic field and at a temperature of $T = 5\,\mathrm{K}$, a transverse spin lifetime (decoherence time) T_2^* exceeding 100 ns was measured, with experimental indications that this time is a single-spin effect [2]. Since T_2^* still includes inhomogeneous effects (e.g. g-factor variations in the material, leading to spins rotating with slightly different frequencies and thus reducing the total magnetization) it represents only a lower bound of the transverse lifetime of a *single* spin, $T_2 \geq T_2^*$, which is relevant for using spins as qubits. With the same optical pump-probe technique, spin lifetimes in semiconductor quantum dots have been measured [32], with at most one spin per dot. The relatively small T_2^* decoherence times (a few ns at vanishing magnetic field), which have been found in these experiments, probably originate in a large inhomogeneous broadening due to a strong variation of g-factors [32].

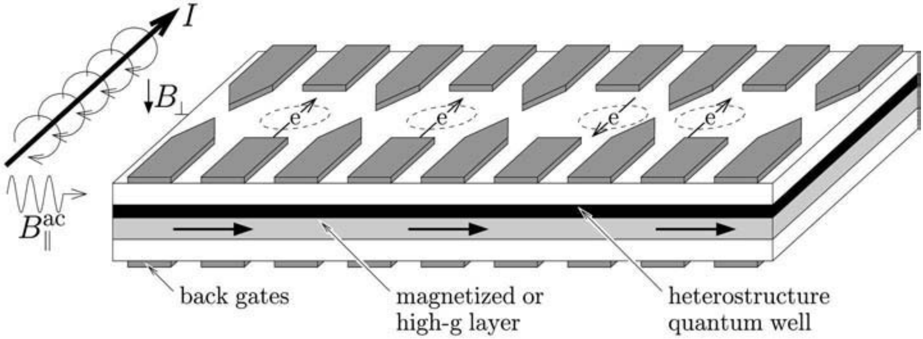


Fig. 1. All-electrically controlled quantum dot array (schematic). Gate electrodes (dark gray) confine single electrons to the dot regions (circles). Electrons can be shifted by electrical gating into the magnetized or high-g layer to produce locally different Zeeman splittings. Alternatively, such local Zeeman fields can be produced by magnetic field gradients as, e.g., induced by a current wire (left). Since every dot-spin is subject to a different Zeeman splitting, the spins can be addressed individually, e.g. through ESR pulses of an additional in-plane magnetic ac field with the corresponding Larmor frequency ω_{L}. Such mechanisms can be used for single-spin rotations and the initialization step. The exchange coupling between the dots is controlled by electrically lowering the tunnel barrier between the dots (as indicated here for the two rightmost dots).

Nevertheless, the fact that many coherent oscillations were observed [32] provides strong experimental support to the idea of using electron spin as qubits.

Since in GaAs semiconductors, both Ga and As possess a nuclear spin $I = 3/2$ (and no Ga/As isotopes are available with $I = 0$) it turns out that a serious source of possible qubit errors using semiconductors such as GaAs is the hyperfine coupling between the electron spin (qubit) and nuclear spins in the quantum dot [34]. Silicon-based structures would be more advantageous from this aspect, but at present the control over nanostructures such as quantum dots in Si is not as advanced as in GaAs. The hyperfine coupling between the electron spin $\mathbf{S}$ and the nuclear spins $\mathbf{I} = \sum_{n=1}^{N} \mathbf{I}^{(n)}$, is given by $A\,\mathbf{S} \cdot \mathbf{I}$, where A is the hyperfine coupling constant. Due to this coupling, a flip of the electron spin with a concomitant change of one nuclear spin may occur, causing an error in the quantum computation. One can analyze this error in the presence of a magnetic field B_z [34], and find in time-dependent perturbation theory that the total probability for a flip of the electron spin oscillates in time (a more detailed theoretical study of electron spin decoherence due to nuclear spins in a quantum dot in the absence of a magnetic field, going beyond perturbation theory, was presented in [33]). The amplitude of the oscillations of the electron spin is

$$P_i \approx \frac{1}{N} \left(\frac{B_{\mathrm{n}}^*}{B} \right)^2 , \tag{1}$$

where B is defined below and $B_{\mathrm{n}}^* = NAI/g\mu_{\mathrm{B}}$ is the maximal magnitude of the effective nuclear field (Overhauser field). In typical quantum dots we have $N \sim 10^5$. If $B_z = 0$ and with a polarization $p \neq 0$, $-1 \leq p \leq 1$ of the nuclear spins, an effective nuclear field $B = pB_{\mathrm{n}}^*$ is produced and the transition probability becomes suppressed with $P_i \approx 1/p^2 N$. Such a polarization p can be established by dynamically spin-polarizing the nuclear spins, e.g. by optical pumping [39] or by spin-polarized currents at the edge of a 2DEG [40]. For these methods, nuclear Overhauser fields are reported as large as $pB_{\mathrm{n}}^* = 4$ T in GaAs (corresponding to $p = 0.85$) [40] and which can have a lifetime on the order of minutes [39]. Alternatively, for unpolarized nuclei, the amplitude of P_i can be suppressed by an external field $B = B_z$ (1). Thus, the decoherence of an electron spin due to hyperfine interaction can be suppressed drastically, either by dynamically polarizing the nuclear spins in the host material or by applying an external magnetic field. It would be highly desirable to test this prediction by measuring the electron-spin T_2 time with and without Overhauser field.

3 Pulsed Switching and Adiabaticity

The Hamiltonian of a set of electron spins (qubits) localized in a coupled array of quantum dots (Fig. 1) can be written as (see also Sects. 5 and 6 below),

$$H(t) = \sum_{i<j} J_{ij}(t)\, \mathbf{S}_i \cdot \mathbf{S}_j + \sum_i \mu_B g_i(t)\, \mathbf{B}_i(t) \cdot \mathbf{S}_i \,, \tag{2}$$

where we assume the exchange coupling to be local, i.e. $J_{ij} \neq 0$ only for neighboring qubits i and j. Quantum gate operations will be controlled via some external control fields $v(t)$, i.e. $J_{ij}(t) = J_{ij}(v(t))$, $g_i(t) = g_i(v(t))$, and $\mathbf{B}_i(t) = \mathbf{B}_i(v(t))$. Note that in cavity-QED systems, there is also a long-range coupling of qubits as some of us have described in [35]. However, nearest-neighbor (local) exchange coupling is sufficient for performing operations on nonneighboring qubits because one can swap the state of two qubits with the help of the exchange interaction (as we will show in Sect. 5), i.e. qubits can be moved around in an array of quantum dots.

The one- and two-qubit operations described in Sects. 5 and 6 do not depend on the details of the function $P(v(t))$, where P stands for the exchange coupling J or the Zeeman interaction. It is only the time integral $\int_0^\tau P(v(t))\mathrm{d}t$ that needs to assume a certain value (modulo 2π). The exchange interaction $J(t)$ should be switched adiabatically, i.e. such that $|\dot{v}/v| \ll \delta\varepsilon/\hbar$, where $\delta\varepsilon$ is the energy scale on which excitations may occur. Here, $\delta\varepsilon$ should be taken as the energy-level separation of a single dot (if spin is conserved). A rectangular pulse leads to excitation of higher levels, whereas an adiabatic pulse with amplitude v_0 is e.g. given by $v(t) = v_0 \,\mathrm{sech}(t/\Delta t)$ where Δt controls the width of the pulse. We need to use a switching time $\tau_s > \Delta t$, such that $v(t = \tau_s/2)/v_0$ becomes vanishingly small. We then have $|\dot{v}/v| = |\tanh(t/\Delta t)|/\Delta t \leq 1/\Delta t$, so we need $1/\Delta t \ll \delta\varepsilon/\hbar$ for adiabatic switching. The Fourier transform $v(\omega) = \Delta t v_0 \pi \,\mathrm{sech}(\pi\omega\Delta t)$ has the same shape as $v(t)$ but width $2/\pi\Delta t$. In particular, $v(\omega)$ decays exponentially in the frequency ω, whereas it decays only with $1/\omega$ for a rectangular pulse.

The rotation of single spins, i.e. single qubit operations, can be performed for example in g-factor-modulated materials, as proposed in Sect. 6. A spin can be rotated by a relative angle of $\phi = \Delta g_{\mathrm{eff}} \mu_B B\tau/2\hbar$ through changing the effective g-factor by Δg_{eff} for a time τ. Thus, a typical switching time for an angle $\phi = \pi/2$, a field $B = 1$T, and $\Delta g_{\mathrm{eff}} \approx 1$ is $\tau_s \approx 30$ps. If slower operations are required, they are easily implemented by choosing a smaller Δg_{eff}, reducing the magnitude of the field B, or by replacing ϕ by $\phi + 2\pi n$ with integer n, thus "overrotating" the spin. Next we consider two exchange-coupled spins, which perform a square-root-of-swap gate for the integrated pulse $\int_0^{\tau_s} J(t)\mathrm{d}t/\hbar = \pi/2$, as described in Sect. 5. We apply a pulse $J(t) = J_0 \,\mathrm{sech}(t/\Delta t)$ with $J_0 = 80\,\mu$eV ≈ 1 K and $\Delta t = 4$ps. Again, we calculate a switching time $\tau_s \approx 30$ps, while the adiabaticity criterion is $\hbar/\Delta t \approx 150\,\mu\mathrm{eV} \ll \delta\varepsilon$. Once more, the switching time can be easily increased by adding $2\pi n$ with integer n to the integrated pulse $\int_0^{\tau_s} \mathrm{d}t J(t)/\hbar$, i.e. by "overswapping" the two spins. This increased switching time allows a slower switching of $J(t)$ if required.

We note here that the total time consumed by an algorithm can be optimized considerably by simultaneously switching different parameters of the Hamiltonian, i.e. producing parallel instead of serial pulses. As an example, we have shown that for an error-correcting algorithm using only three qubits, a speed-up of a factor of two can be achieved [36]. For algorithms handling a larger number of qubits, a more drastic optimization can be expected.

4 Initialization of a Quantum Register

Commonly, quantum algorithms and error correcting schemes require as an input a properly initialized register of qubits in a well defined state such as spin up, $|\uparrow\rangle$. Single spins can be polarized by exposing them to a large magnetic field $g\mu_B B \gg kT$ and letting them relax to the ground state. A magnetic field could be applied locally or realized by forcing the electrons (via external gates) into a magnetized layer, into a layer with a different effective g-factor [8,11] or into a layer with polarized nuclear spins (Overhauser effect) [34] etc., see also Fig. 1 and Sect. 6. If a spin-polarized current can be produced, such as by spin-polarizing materials [3,4] or by spin-filtering with the help of another dot [41] (see Sect. 7.3), polarized electrons can be injected into an empty quantum dot, i.e. the dot is filled with an already initialized spin.

Sometimes, it may be favorable to start with a given initial state, such as $|0110\ldots\rangle$, instead of $|0000\ldots\rangle$. This can be readily implemented with spins as qubits using standard electron spin resonance (ESR) techniques [34]: We start with a ground state $|0000\ldots\rangle$ and then produce a Zeeman splitting by applying a static local magnetic field for these spins, which should be initialized into state $|1\rangle$. An ac magnetic field is then applied perpendicularly to the first field with a resonant frequency that matches the Larmor frequency $\omega_{\mathrm{L}} = g\mu_B B/\hbar$. Due to paramagnetic resonance [42], this causes spin flips in the quantum dots with the corresponding Zeeman splitting, thus producing the desired state. We note that since we do not want to affect the other spins (having a different Zeeman splitting) the amplitude of the ac field must be switched adiabatically, see Sect. 3. Of course, spin precession can also be used to perform single-spin rotations (see Sect. 6).

5 Two-Qubit Operations in Coupled Quantum Dots

In combination with single-qubit operations, two-qubit gates are sufficient for doing arbitrary quantum computations [43] (i.e. they can form a *universal* set of quantum gates). We can therefore focus on a mechanism that couples pairs of spin-qubits. A mechanism of this type exists in coupled quantum dots, resulting from the combined action of the Coulomb interaction and the Pauli exclusion principle. Two coupled electrons in absence of a magnetic field have a spin-singlet ground state, while the first excited state in the presence of strong Coulomb repulsion is a spin triplet. Higher excited states are separated from these two lowest states by an energy gap, given either by the Coulomb repulsion or the single-particle confinement. The low-energy dynamics of such a system can be described by the effective Heisenberg spin Hamiltonian

$$H_{\mathrm{s}}(t) = J(t)\ \mathbf{S}_1 \cdot \mathbf{S}_2, \tag{3}$$

where $J(t)$ denotes the exchange coupling between the two spins $\mathbf{S}_1$ and $\mathbf{S}_2$, i.e. the energy difference between the triplet and the singlet. After a pulse of $J(t)$ with $\int_0^{\tau_s} \mathrm{d}t J(t)/\hbar = J_0\tau_s/\hbar = \pi \pmod{2\pi}$, the time evolution $U(t) =$

$T\exp(\mathrm{i}\int_0^t H_\mathrm{s}(\tau)\mathrm{d}\tau/\hbar)$ corresponds to the "swap" operator U_sw, whose application leads to an interchange of the states in qubit 1 and 2 [8]. While U_sw is not sufficient for quantum computation, any of its square roots $U_\mathrm{sw}^{1/2}$, say

$$U_\mathrm{sw}^{1/2}|\phi\chi\rangle = (|\phi\chi\rangle + \mathrm{i}|\chi\phi\rangle)/(1+\mathrm{i}) \tag{4}$$

turns out to be a *universal* quantum gate. Thus, it can be used, together with single-qubit rotations, to assemble any quantum algorithm. This is shown by constructing the known universal gate XOR [44], through combination of $U_\mathrm{sw}^{1/2}$ and single-qubit operations $\exp(\mathrm{i}\pi S_i^z/2)$, applied in the sequence [8],

$$U_\mathrm{XOR} = e^{\mathrm{i}(\pi/2)S_1^z} e^{-\mathrm{i}(\pi/2)S_2^z}\, U_\mathrm{sw}^{1/2}\, e^{\mathrm{i}\pi S_1^z}\, U_\mathrm{sw}^{1/2}. \tag{5}$$

Knowing about the universality of these gates, we can reduce the study of general quantum computation to the study of single-spin rotations (see Sect. 6) and the *exchange mechanism*, in particular how $J(t)$ can be controlled experimentally. The central idea is that $J(t)$ can be switched by raising or lowering the tunneling barrier between the dots. In the following, we shall review our detailed calculations to describe such a mechanism. We note that the same principles can also be applied to other spin systems in quantum-confined structures, such as coupled atoms in a crystal, supramolecular structures, and overlapping shallow donors in semiconductors [23,45] etc., using similar methods as explained below. We point out that, beyond the mechanisms described in Sect. 5.1 and Sect. 5.2, spins in quantum dots can also be coupled on a long distance scale by using a cavity-QED scheme [35] or by using superconducting leads to which the quantum dots are attached [46], see Sect. 8.4.

5.1 Laterally Coupled Dots

Two quantum dots can be coupled in a two-dimensional electron gas (2DEG), containing one (excess) electron each, as described in Sect. 1.3. The dots are arranged in a plane, at a sufficiently small distance $2a$, such that the electrons can tunnel between the dots (for a lowered barrier) and an exchange interaction J between the two spins is produced. We model this system of coupled dots with the Hamiltonian

$$H = \sum_{i=1,2} h_i + C + H_\mathrm{Z} = H_\mathrm{orb} + H_\mathrm{Z}, \tag{6}$$

where the single-electron dynamics in the 2DEG (xy plane) is described through

$$h_i = \frac{1}{2m}\left(\mathbf{p}_i - \frac{e}{c}\mathbf{A}(\mathbf{r}_i)\right)^2 + V(\mathbf{r}_i), \tag{7}$$

with m being the effective mass and $V(\mathbf{r}_i)$ the confinement potential as given below. A magnetic field $\mathbf{B} = (0,0,B)$ is applied along the z axis, which couples to the electron spin through the Zeeman interaction H_Z and to the charge through the vector potential $\mathbf{A}(\mathbf{r}) = \frac{B}{2}(-y,x,0)$. In almost depleted regions, like few-electron quantum dots, the screening length λ can be expected to be much larger than the screening length in bulk 2DEG regions (where it is 40 nm for GaAs). Thus, for small quantum dots, say $\lambda \gg 2a \approx 40\,\mathrm{nm}$, we need to consider the bare Coulomb interaction $C = e^2/\kappa|\mathbf{r}_1 - \mathbf{r}_2|$, where κ is the static dielectric constant. The confinement and tunnel-coupling in (7) for laterally aligned dots is modeled by the quartic potential

$$V(x,y) = \frac{m\omega_0^2}{2}\left[\frac{1}{4a^2}\left(x^2 - a^2\right)^2 + y^2\right], \tag{8}$$

with the inter-dot distance $2a$ and $a_{\mathrm{B}} = \sqrt{\hbar/m\omega_0}$ the effective Bohr radius of the dot. Separated dots ($a \gg a_{\mathrm{B}}$) are thus modeled as two harmonic wells with frequency ω_0. This is motivated by the experimental evidence that the low-energy spectrum of single dots is well described by a parabolic confinement potential [28].

We now consider only the two lowest orbital eigenstates of H_{orb}, leaving us with one symmetric (spin-singlet) and one antisymmetric (spin-triplet) orbital state. The spin state for the singlet is

$$|S\rangle = (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}, \tag{9}$$

while the triplet spin states are

$$|T_0\rangle \;= (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}, \tag{10}$$

$$|T_+\rangle = |\uparrow\uparrow\rangle, \quad \text{and} \quad |T_-\rangle = |\downarrow\downarrow\rangle. \tag{11}$$

For temperatures with $kT \ll \hbar\omega_0$, higher-lying states are frozen out and H_{orb} can be replaced by the effective Heisenberg spin Hamiltonian (3). The exchange energy $J = \epsilon_{\mathrm{t}} - \epsilon_{\mathrm{s}}$ is given as the difference between the triplet and singlet energy. For calculating these energies, we use the analogy between atoms and quantum dots and make use of variational methods similar to the ones in molecular physics. Using the Heitler-London ansatz with ground-state single-dot orbitals, we find [34],

$$J = \frac{\hbar\omega_0}{\sinh\left(2d^2(2b - 1/b)\right)}\left\{\frac{3}{4b}\left(1 + bd^2\right)\right. \tag{12}$$

$$\left. + c\sqrt{b}\left[e^{-bd^2}\, I_0\left(bd^2\right) - e^{d^2(b-1/b)}\, I_0\left(d^2(b - 1/b)\right)\right]\right\},$$

where we have introduced the dimensionless distance $d = a/a_{\mathrm{B}}$ between the dots and the magnetic compression factor $b = B/B_0 = \sqrt{1 + \omega_L^2/\omega_0^2}$ with the Larmor frequency $\omega_L = eB/2mc$. The zeroth order Bessel function is denoted by I_0. In (12), the first term comes from the confinement potential, while the terms proportional to the parameter $c = \sqrt{\pi/2}(e^2/\kappa a_{\mathrm{B}})/\hbar\omega_0$ result from the Coulomb interaction C; the exchange term is recognized by its negative sign. We are mainly interested in the weak coupling limit $|J/\hbar\omega_0| \ll 1$, where the ground-state Heitler-London ansatz is self-consistent. We plot J (12) in Fig. 2

as a function of B and d. We note that $J(B=0) > 0$, which is generally true for a two-particle system with time-reversal invariance. We observe that over a wide range of the parameters c and a, the sign of $J(B)$ changes from positive to negative at a finite value of B (for the parameters chosen in Fig. 2(a) at $B \approx 1.3\,\mathrm{T}$). J is suppressed exponentially either by compression of the electron orbitals through large magnetic fields ($b \gg 1$), or by large distances between the dots ($d \gg 1$), where in both cases the orbital overlap of the two dots is reduced. This exponential suppression, contained in the $1/\sinh$ prefactor in (12), is partly compensated by the exponentially growing exchange term $\propto \exp(2d^2(b-1/b))$. In total, J decays exponentially as $\exp(-2d^2 b)$ for large b or d. Since the sign reversal of J–signalling a singlet-triplet crossing–results from the long-range Coulomb interaction, it is not contained in the standard Hubbard model which takes only short-range interaction into account. In this latter model one finds $J = 4t^2/U > 0$ in the limit $t/U \ll 1$ (see Fig. 2). The Heitler-London result (12) was refined by taking higher levels and double occupancy of the dots into account (implemented in a Hund-Mullikan approach), which leads to qualitatively similar results [34], in particular concerning the singlet-triplet crossing.

Note that the exponential suppression of J is very desirable for minimizing gate errors. In the absence of tunneling between the dots we still might have direct Coulomb interaction left between the electrons. However, this has no effect on the spins (qubit) provided the spin-orbit coupling is sufficiently small, which is the case for s-wave electrons in GaAs structures with unbroken inversion symmetry (this would not be so for hole-doped systems since the hole has a much stronger spin-orbit coupling due to its p-wave character). Finally, the vanishing of J can be exploited for switching by applying a constant homogeneous magnetic field to an array of quantum dots to tune J to zero (or close to some other desirable value). Then, for switching J on and off, only a small gate pulse or a small local magnetic field is needed.

5.2 Vertically Coupled Dots

The case of vertically tunnel-coupled quantum dots was also investigated [47]. A vertical arrangements of the dots has been produced in multilayer self-assembled quantum dots (SAD) [48] as well as in etched mesa heterostructures [49]. We apply the same methods as described in Sect. 5.1 for laterally coupled dots, but now we extend the Hamiltonian (7) from two to three dimensions and take a three-dimensional confinement $V = V_l + V_v$. We implement the vertical confinement V_v as a quartic potential similar to (8), with curvature ω_z at $z = \pm a$ [see Fig. 3(b)], implying an effective Bohr radius $a_\mathrm{B} = \sqrt{\hbar/m\omega_z}$ and a dimensionless distance $d = a/a_\mathrm{B}$. We have modeled a harmonic potential for the lateral confinement, while we have allowed different sizes of the two dots $a_{\mathrm{B}\pm} = \sqrt{\hbar/m\alpha_{0\pm}\omega_z}$. This allows additional switching mechanisms as it is explained in the next paragraph.

Since we are considering a three-dimensional setup, the exchange interaction is not only sensitive to the magnitude of the applied fields, but also to their direction. We now give a brief overview of our results [47] for in-plane ($B_\|$, $E_\|$) and

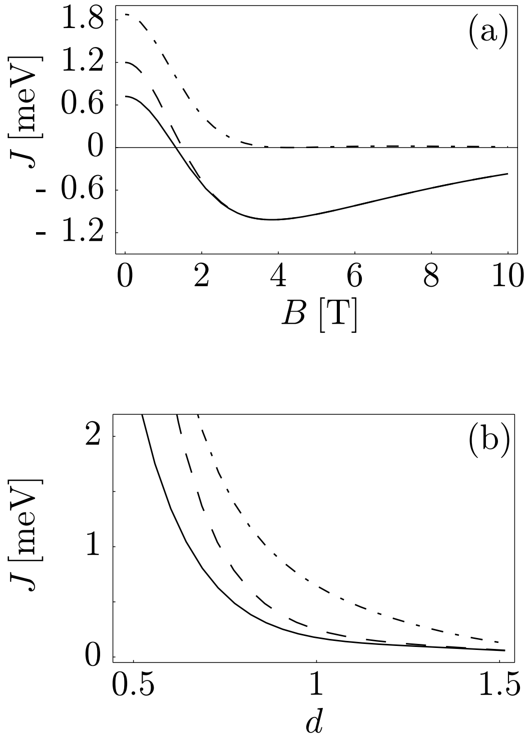


Fig. 2. Exchange coupling J (full line) for GaAs quantum dots with confinement energy $\hbar\omega = 3\,\text{meV}$ and $c = 2.42$. For comparison we plot the usual short-range Hubbard result $J = 4t^2/U$ (dashed-dotted line) and the extended Hubbard result [34] $J = 4t^2/U + V$ (dashed line). In (a), J is plotted as a function of the magnetic field B at fixed inter-dot distance $d = a/a_\text{B} = 0.7$, while in (b) as a function of the inter-dot distance $d = a/a_\text{B}$ at $B = 0$.

perpendicular ($B_\perp$, $E_\perp$) fields; this setup is illustrated in Fig. 3(a): (1) An in-plane magnetic field $B_\|$ suppresses J exponentially; a perpendicular field in laterally coupled dots has the same effect (Sect. 5.1). (2) A perpendicular magnetic fields $B_\perp$ reduces on the one hand the exchange coupling between identically sized dots $\alpha_{0+} = \alpha_{0-}$ only slightly. On the other hand, for different dot sizes $a_{\text{B}+} < a_{\text{B}-}$, the behavior of $J(B_\perp)$ is no longer monotonic: Increasing $B_\perp$ from zero amplifies the exchange coupling J until both electronic orbitals are magnetically compressed to approximately the same size, i.e. $B \approx 2m\alpha_{0+}\omega_z c/e$. From this point, J decreases weakly, as for identically sized dots. (3) A perpendicular electric field $E_\perp$ detunes the single-dot levels, and thus reduces the exchange coupling; the very same finding was made for for laterally coupled dots and an in-plane electric field [34]. (4) An in-plane electric field $E_\|$ and different dot sizes provide another switching mechanism for J. The dots are shifted parallel to the field by $\varDelta x_\pm = E_\|/E_0\alpha_{0\pm}^2$, where $E_0 = \hbar\omega_z/ea_B$. Thus, the larger dot is shifted

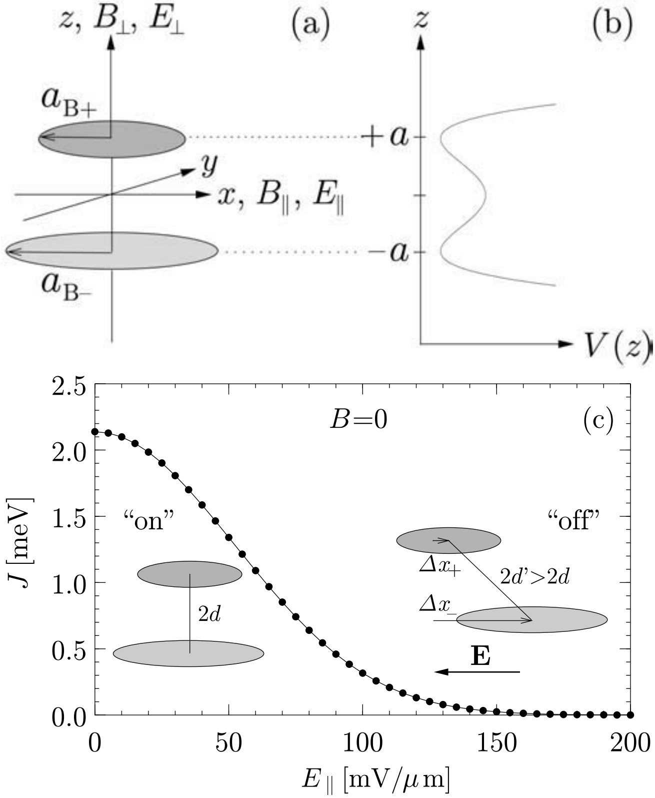


Fig. 3. (a) Two vertically coupled quantum dots with different lateral diameters a_{B+} and a_{B-}. In the text, we discuss magnetic and electric fields applied either in-plane ($B_\parallel$, $E_\parallel$) or perpendicularly ($B_\perp$, $E_\perp$). (b) The quartic double-well potential used for modeling the vertical confinement V_v, see text. (c) Switching of the spin-spin coupling between dots of different size by means of an in-plane electric field $E_\parallel$ at $B = 0$. We have chosen $\hbar\omega_z = 7\,\mathrm{meV}$, $d = 1$, $\alpha_{0+} = 1/2$ and $\alpha_{0-} = 1/4$. For these parameters, $E_0 = \hbar\omega_z/ea_B = 0.56\,\mathrm{mV/nm}$ and $A = (\alpha_{0+}^2 - \alpha_{0-}^2)/2\alpha_{0+}^2\alpha_{0-}^2 = 6$. The exchange coupling J decreases exponentially on the scale $E_0/2A = 47\,\mathrm{mV}/\mu\mathrm{m}$ for the electric field. Thus, the exchange coupling is switched "on" for $E_\parallel = 0$ and "off" for $E_\parallel > 150\,\mathrm{mV}/\mu\mathrm{m}$, see text.

a greater distance $\Delta x_- > \Delta x_+$ and so the mean distance between the electrons grows as $d' = \sqrt{d^2 + A^2(E_\parallel/E_0)^2} > d$, taking $A = (\alpha_{0+}^2 - \alpha_{0-}^2)/2\alpha_{0+}^2\alpha_{0-}^2$. Since the exchange coupling J is exponentially sensitive to the inter-dot distance d', it is suppressed exponentially when an in-plane electric field is applied, $J \approx \exp[-2A^2(E_\parallel/E_0)^2]$, which is illustrated in Fig. 3(c). Thereby we have given an exponential switching mechanism for quantum gate operation relying only on a tunable electrical field, in addition to the magnetically driven switching discussed above.

5.3 Singlet–Triplet Entangling Gate

Encoding a single spin 1/2 state $|\alpha\rangle$ into a singlet (for $\alpha =\downarrow$) or triplet (for $\alpha =\uparrow$) two-spin state one can measure the state of the qubit represented by $|\alpha\rangle$, if a measurement device capable of distinguishing singlet/triplet states is available (see, e.g., Sect. 8.3). Moreover, this operation acts as an "entangler" for electron pairs used in quantum communication (see Sect. 8). We can construct such a two-qubit operation explicitly. While quantum dot 1 is in state $|\alpha\rangle$, we prepare the state of the quantum dot 2 to $|\uparrow\rangle$, perform a $U_{\mathrm{sw}}^{1/2}$ gate and finally apply a local Zeeman term, generating the time evolution $\exp\{\mathrm{i}(\pi/2)S_1^z\}$, thus

$$\left.\begin{array}{l}|\uparrow\uparrow\rangle\\ |\downarrow\uparrow\rangle\end{array}\right\}\xrightarrow{e^{\mathrm{i}\frac{\pi}{2}S_1^z}U_{\mathrm{sw}}^{1/2}}\left\{\begin{array}{l}e^{\mathrm{i}\frac{\pi}{4}}\,|\uparrow\uparrow\rangle\,,\\ -\mathrm{i}\left(|\downarrow\uparrow\rangle-|\uparrow\downarrow\rangle\right)/\sqrt{2}\,.\end{array}\right. \tag{13}$$

In other words, this operation maps the triplet $|\uparrow\uparrow\rangle$ (and $|\downarrow\downarrow\rangle$) into itself, while the state $|\downarrow\uparrow\rangle$ is mapped into the singlet (and $|\uparrow\downarrow\rangle$ into the triplet $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}$), up to phase factors.

6 Single-Spin Operations

In addition to a suitable two-qubit gate (e.g., quantum XOR or "square-root of swap"), arbitrary single-qubit operations are required for doing quantum computing. If the qubits are represented by electron spins, this means that one has to be able to perform arbitrary single-spin rotations. It must therefore be possible to expose a specific qubit to a time-varying Zeeman coupling $(g\mu_B\mathbf{S}\cdot\mathbf{B})(t)$ [34], which can be controlled through both the magnetic field $\mathbf{B}$ and/or the g-factor g. Since only relative phases have a relevance, it is sufficient to rotate all spins of the system at once (e.g. by an external field B), but with a different Larmor frequency. We have proposed a number of possible implementations [8,34,11,15] for spin-rotations:

The location of the electron in equilibrium can be shifted around through electrical gating. Thus, if the electron wave function is pushed into a region with a different magnetic field strength or (effective) g-factor, one produces a relative rotation around the direction of $\mathbf{B}$ by an angle of $\phi = (g'B' - gB)\mu_B\tau/2\hbar$, see Fig. 1. Regions with an increased magnetic field can be provided by a magnetic (dot) material while an effective magnetic field can be produced, e.g., with dynamically polarized nuclear spins (Overhauser effect) [34].

The idea for using g-factor-modulated materials is as follows [11,15]. In bulk semiconductors the free-electron value of the Landé g-factor $g_0 = 2.0023$ is modified by spin-orbit coupling. Similarly, the g-factor can be drastically enhanced by doping the semiconductor with magnetic impurities [4,3]. In confined structures such as quantum wells, wires, and dots, the g-factor is further modified and becomes sensitive to an external bias voltage [50]. We have numerically analyzed a system with a layered structure (AlGaAs-GaAs-InAlGaAs-AlGaAs), in which the effective g-factor of electrons is varied by shifting their equilibrium position from one layer to another by electrical gating [51]. We have found that in this structure the effective g-factor can be changed by about $\Delta g_{\mathrm{eff}} \approx 1$ [15].

ESR techniques can be used as an alternative for switching (as already explained in Sect. 4). Moreover, localized magnetic fields can be generated with the magnetic tip of a scanning force microscope, a magnetic disk writing head, by placing the dots above a grid of current-carrying wires, or by placing a small wire coil above the dot etc.

7 Single-Spin Measurement

7.1 Spin Measurements Through Spontaneous Magnetization

The read-out the spin of an electron on a quantum dots can e.g. be achieved by tunneling of this electron into a supercooled paramagnetic dot [8,11]. In the latter, the spin induces a magnetization nucleation from the paramagnetic metastable phase into a ferromagnetic domain, whose magnetization direction (θ, ϕ) is along the measured spin direction and which can be measured by conventional means. Since this direction is continuous rather than only one of two values, we describe this generalized measurement in the formalism of positive-operator-valued (POV) measurements [52] as projection into the overcomplete set of spin-1/2 coherent states $|\theta, \phi\rangle = \cos(\theta/2)|\uparrow\rangle + e^{\mathrm{i}\phi}\sin(\theta/2)|\downarrow\rangle$. Thus if we interpret a magnetization direction in the upper hemisphere as $|\uparrow\rangle$, we have a 75%-reliable measurement, since $(1/2\pi)\int_{\theta\geq\pi/2} \mathrm{d}\Omega\, |\langle\uparrow|\theta, \phi\rangle|^2 = 3/4$, using the normalization constant 2π for the coherent spin states.

7.2 Spin Measurements via the Charge

The true microscopic nature of the spin qubit involves the advantage of long decoherence times, but the smallness of its magnetic moment makes it also very hard to measure a single spin. However, measuring the charge of single electrons, e.g. with the use of a single-electron transistor, is done routinely nowadays. Thus it is desirable to have a mechanism for detecting the spin of an electron via measuring charge, i.e. voltage or current [8].

A straightforward concept yielding a potentially 100% reliable measurement requires a switchable “spin-filter” tunnel barrier which allows only, say, spin-up but no spin-down electrons to tunnel. When the measurement of a spin in a quantum dot is to be performed, tunneling between this dot and a second dot, connected to an electrometer, is switched on, but only spin-up electrons are allowed to pass (spin-filtering). Thus if the spin had been up, a charge would be detected in the second dot by the electrometer [8], and no charge otherwise. Again, this is a POV type of measurement (see above). It is known how to build electrometers with single-charge detection capabilities; resolutions down to 10^{-8} of one electron charge have been reported [53]. Spin filtering and also spin-state measurements can be achieved by tunneling through a quantum dot [41] as we shall discuss next.

7.3 Quantum Dot as Spin Filter and Read-Out/Memory Device

Here, we consider a quantum dot attached to in- and outgoing current leads $l = 1, 2$–which can be operated as a spin filter, or as a read-out device, or as a spin-memory where a single spin stores the information [41].

For this proposal, it is essential that the spin-degeneracy is lifted with *different* Zeeman splittings in the dot and in the leads, e.g. by using materials with different effective g-factors for leads and dot [41]. This results in Coulomb blockade peaks and spin-polarized currents which are uniquely associated with the spin state on the dot.

The setup is described by a standard tunneling Hamiltonian [54]

$$H = H_0 + H_T, \qquad H_0 = H_L + H_D, \tag{14}$$

$$H_T = \sum_{l,k,p,s} t_{lp} c_{lks}^\dagger d_{ps} + \text{h.c.} \tag{15}$$

where H_0 describes the leads and the dot, H_D includes the charging and interaction energies of the electrons in the dot as well as their Zeeman energy $\pm g\mu_B B/2$ in an external magnetic field $\mathbf{B}$, and the tunneling between leads and the dot is described by H_T, where c_{lks} annihilates electrons with spin s and momentum k in lead l and d_{ps} annihilates electrons in the dot. We consider the Coulomb blockade regime [27] where the charge on the dot is quantized. Then we apply a standard master-equation approach [55,41] with a reduced density matrix of the dot and calculate the transition rates in a "golden-rule" approach up to 2nd order in H_T. The first-order contribution to the current is the sequential tunneling current I_s [27], where the number of electrons on the dot fluctuates and thus the processes of an electron tunneling from the lead onto the dot and vice versa are allowed by energy conservation. The second-order contribution is the cotunneling current I_c [56], involving a virtual intermediate state with a different number of electrons on the dot (see also Sect. 8.3).

Now we assume that the Zeeman splitting in the leads is negligible (i.e. much smaller than the Fermi energy) while on the dot it is given as $\varDelta_z = \mu_B|gB|$. We assume a small bias $\varDelta\mu = \mu_1 - \mu_2 > 0$ between the leads at chemical potential $\mu_{1,2}$ and low temperatures so that $\varDelta\mu$, $kT < \delta$, where δ is the characteristic energy-level distance on the dot. First we consider a quantum dot in the ground state, filled with an odd number of electrons with total spin 1/2, which we assume to be $|\uparrow\rangle$ and to have energy $E_\uparrow = 0$. If an electron tunnels from the lead onto the dot, a spin singlet is formed with energy E_S, while the spin triplets are (usually) excited states with energies $E_{T_\pm}$ and E_{T_0}. At the sequential tunneling resonance, $\mu_1 > E_S > \mu_2$, where the number of electrons on the dot fluctuates between N and $N + 1$, and in the regime $E_{T_+} - E_S$, $\varDelta_z > \varDelta\mu$, kT, energy conservation only allows ground state transitions. Thus, spin-up electrons are not allowed to tunnel from lead 1 via the dot into lead 2, since this would involve virtual states $|T_+\rangle$ and $|\downarrow\rangle$, and so we have $I_s(\uparrow) = 0$ for sequential tunneling. However, spin down electrons may pass through the dot in the process $\downarrow\textcircled{\uparrow}_i \to \textcircled{\uparrow\downarrow}_f$, followed by $\textcircled{\uparrow\downarrow}_i \to \textcircled{\uparrow}\downarrow_f$. Here the state of the quantum dot is drawn inside the circle, while the states in the leads are drawn to the left and right of the circle, respectively. This leads to a *spin-polarized* sequential tunneling current $I_s = I_s(\downarrow)$, which we have calculated as [41]

$$I_s(\downarrow)/I_0 = \theta(\mu_1 - E_S) - \theta(\mu_2 - E_S), \quad k_BT < \Delta\mu, \tag{16}$$

$$I_s(\downarrow)/I_0 = \frac{\Delta\mu}{4k_BT}\cosh^{-2}\left[\frac{E_S - \mu}{2k_BT}\right], \quad k_BT > \Delta\mu, \tag{17}$$

where $\mu = (\mu_1 + \mu_2)/2$ and $I_0 = e\gamma_1\gamma_2/(\gamma_1 + \gamma_2)$. Here $\gamma_l = 2\pi\nu|A_{lnn'}|^2$ is the tunneling rate between lead l and the dot and we have introduced the matrix elements $A_{ln'n} = \sum_{ps} t_{lp}\langle n'|d_{ps}|n\rangle$. Similarly, for N even we find $I_s(\downarrow) = 0$ while for $I_s(\uparrow)$ a similar result holds [41] as in (16) and (17).

Even though I_s is completely spin-polarized, a leakage of current with opposite polarization arises through cotunneling processes [41]; still the leakage is small, and the efficiency for $\Delta_z < |E_{T_+} - E_S|$ for spin filtering in the sequential regime becomes [41]

$$I_s(\downarrow)/I_c(\uparrow) \sim \frac{\Delta_z^2}{(\gamma_1 + \gamma_2)\max\{k_BT, \Delta\mu\}}, \tag{18}$$

and equivalently for $I_s(\uparrow)/I_c(\downarrow)$ at the even-to-odd transition. In the sequential regime we have $\gamma_i < k_BT, \Delta\mu$, thus, for $k_BT, \Delta\mu < \Delta_z$, we see that the spin-filtering is very efficient.

The opposite case where the leads are fully spin polarized with a much smaller Zeeman splitting on the dot [41] can be realized with magnetic semiconductors (with effective g-factors reaching 100 [3]) where spin-injection into GaAs has recently been demonstrated for the first time [3,4]. Another possibility would be to work in the quantum Hall regime where spin-polarized edge states are coupled to a quantum dot [57]. In this setup the device can be used as read-out for the spin state on the dot. Assume now that the spin polarization in both leads is up, and the ground state of the dot contains an odd number of electrons with total spin 1/2. Now the leads can provide and absorb only spin-up electrons. Thus, a sequential tunneling current will only be possible if the dot state is $|\downarrow\rangle$ (to form a singlet with the incoming electron, whereas the triplet is excluded by energy conservation). Hence, the current is much larger for the spin on the dot being in $|\downarrow\rangle$ than it is for $|\uparrow\rangle$. Again, there is a small cotunneling leakage current for the dot state $|\uparrow\rangle$, with a ratio of the two currents given by (18). Thus, we can probe (read out) the spin-state on the quantum dot by measuring the current which passes through the dot. Given that the sequential tunneling current is typically on the order of $0.1 - 1$ nA [27], we can estimate the read-out frequency $I/2\pi e$ to be on the order of $0.1 - 1$ GHz. Combining this with the initialization and read-in techniques from Sect. 4, i.e. ESR pulses to switch the spin state, we have a *spin memory* at the ultimate single-spin limit, whose relaxation time is just the spin relaxation time. This relaxation time can be expected to be on the order of 100's of nanoseconds [2], and can be directly measured via the currents when they switch from high to low due to a spin flip on the dot [41].

7.4 Optical Measurements

Faraday rotation measurements [2] originating from a pair of coupled electrons would allow us to distinguish between spin singlet and triplet [47]: In the singlet state ($S = 0$, no magnetic moment) there is no Faraday rotation, whereas in the triplet state ($S = 1$) the polarization of linearly polarized light is rotated slightly due to the presence of the magnetic moment. A single spin $|\alpha\rangle$ can be measured either directly via Faraday rotation or by first entangling it with another spin $|\uparrow\rangle$ and then applying the singlet/triplet-measurement. This entanglement is achieved by applying the gate defined in Sect. 5.3, resulting in either a triplet or singlet, depending on whether $|\alpha\rangle$ was $|\uparrow\rangle$ or $|\downarrow\rangle$. However, much more work is required to analyze the Faraday rotation (in particular to calculate the oscillator strength for such processes) in order to assess its efficiency for spin measurements.

8 Quantum Communication with Entangled Electrons

Entangled (EPR) pairs of particles or qubits, e.g. spin singlets $|S\rangle$, are a fundamental resource for a number of tasks in quantum communication [26]. By definition, a pure state of two particles (qubits) is entangled, if it cannot be expressed as a tensor product of two single-particle states. Note that also the spin triplet $|T_0\rangle$ is an entangled state, while the other two triplets $|T_\pm\rangle$ are not. The quantum gate mechanism described in Sect. 5.3 is one possibility for producing such entangled states (we call in general such a device an *entangler*, for which a number of realizations are conceivable, see [13,14]). Here we discuss three experimental setups by which the entanglement of electrons can be detected via their charge in transport and noise measurements in mesoscopic nanostructures [11,58,59,46]. This investigation touches on fundamental issues such as the nonlocality of quantum mechanics, especially for massive particles, and genuine two-particle Aharonov-Bohm effects which are fascinating topics in their own right. The main idea here is to exploit the unique relation between the symmetry of the orbital state and the spin state (for two electrons) which makes it possible to detect the spin state again via the charge (orbital) degrees of freedom of the electrons. In quantum optics, violations of Bell inequalities and quantum teleportation with photons have been investigated [60,61], while so far no corresponding experiments for electrons in a solid-state environment are reported.

8.1 Adding Entangled Electrons to the Fermi Sea

When studying the injection of entangled electrons into a Fermi sea, it is important to keep in mind that the injected electrons electrostatically interact with all the other electrons in the leads. Therefore, we need to analyze interaction effects on the entanglement [11,59]. When we add an electron in state q to a Fermi sea (lead), the quasiparticle weight of that state will be renormalized by $0 \le z_q \le 1$

(see below), i.e. some weight $1-z_q$ to find the electron in the original state q will be distributed among all the other electrons [11,59]. This rearrangement of the Fermi system due to the Coulomb interaction happens very quickly, on a time scale given by the inverse plasmon frequency. So, the question now is: how big is this renormalization? More precisely, when a triplet/singlet electron pair (t and s for short) is injected from an entangler into two leads 1 and 2, we obtain the state

$$|\psi^{t/s}_{\mathbf{nn}'}\rangle = \frac{1}{\sqrt{2}}\,(a^\dagger_{\mathbf{n}\uparrow}a^\dagger_{\mathbf{n}'\downarrow} \pm a^\dagger_{\mathbf{n}\downarrow}\,a^\dagger_{\mathbf{n}'\uparrow}\,)\,|\psi_0\rangle, \tag{19}$$

with the filled Fermi sea $|\psi_0\rangle$, $\mathbf{n} = (\mathbf{q}, l)$, $\mathbf{q}$ the momentum of an electron, and l the lead number. The operator $a^\dagger_{\mathbf{n}\sigma}$ creates an electron in state $\mathbf{n}$ with spin σ. The propagation of the triplet or singlet, interacting with all other electrons in the Fermi sea, can be described by the 2-particle Green's function

$$G^{t/s}(\mathbf{12},\mathbf{34};t) = \langle\psi^{t/s}_{\mathbf{12}}, t|\psi^{t/s}_{\mathbf{34}}\rangle. \tag{20}$$

If we prepare a triplet (singlet), $G^{t/s}(\mathbf{12},\mathbf{12};t)$ is the amplitude of finding a triplet (singlet) after time t. Assuming sufficiently separated leads with negligible mutual interaction, we find [11,59] $|G^{t/s}(\mathbf{12},\mathbf{12};t)| = z_F^2$. For a spin-independent Hamiltonian with bare Coulomb interaction only and within the random phase approximation (RPA) [54], the quasiparticle weight for a 2DEG is given by [11,59] $z_F = 1 - r_s\,(1/2 + 1/\pi)$, in leading order of the interaction parameter $r_s = 1/k_F a_B$, where $a_B = \epsilon_0\hbar^2/me^2$ is the Bohr radius and k_F the Fermi wave vector. In a GaAs 2DEG we have $a_B = 10.3$ nm and $r_s = 0.614$, and thus we obtain $z_F = 0.665$. Therefore, we conclude that the entanglement of a pair of electrons injected into a Fermi liquid will be reduced but there is still a finite probability left to preserve the entangled state. This holds provided the spin-scattering effects are small. That this is indeed the case in GaAs 2DEGs is supported by experiments [2] where the electron spin has been transported phase-coherently over distances of up to 100 μm [2].

8.2 Shot Noise of Entangled Electrons

The "bunching" behavior of photons in correlation experiments due to the Bose statistics is well established [62,63]. The opposite behavior is expected theoretically for electrons, being fermions [64,65,66]. Indeed, "antibunching" of electrons was recently found experimentally [67]. However, as we have pointed out [11] the noise of electrons in current-carrying wires is not sensitive to the symmetry of the total wave function but only to the symmetry of the *orbital* part of it, at least if no spin-scattering processes are present. Thus, if we now consider a two-electron state, we expect antibunching for the triplet states, since they have an antisymmetric orbital wave function, whereas the orbital wave function associated with the spin singlet state is symmetric, and so we expect a bunching behavior. This leads to an observable decrease or increase in noise for electrons, depending on their common spin state, as we shall discuss next [59].

Here, we assume that an entangler generates pairs of entangled electrons which are then injected into lead 1 and 2, one electron each, as shown in Fig. 4. A beam splitter is inserted in order to create two-particle interference effects in the sense that there is an equal probability amplitude for incoming electrons (from lead 1 or 2) to leave into lead 3 or 4 (note that the electrons in a Fermi liquid wire hardly interact which each other; the role of the beam splitter is thus to simulate direct and exchange Coulomb processes). The quantity of interest is then the noise, i.e. the current-current correlations, measured in leads 3 and/or 4.

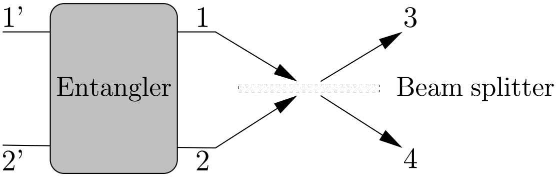


Fig. 4. Proposed scheme for measuring noise of entangled electrons. Uncorrelated electrons are fed into the entangler from the Fermi leads $1'$ and $2'$. Pairs of entangled electrons (singlet $|S\rangle$ or triplet $|T_0\rangle$) are produced in the entangler and then injected into the leads 1 and 2, one electron per lead. The current of these two leads are then mixed with a beam splitter (to induce scattering interference) and the resulting noise is then measured in lead 3 and 4: no noise (antibunching) for triplets, whereas we get enhanced noise (bunching) for singlets (i.e. EPR pairs).

The amplitude of recovering a singlet or triplet state after injecting it into an interacting Fermi sea is reduced by a factor of $z_F^{-2} \approx 2$ (see Sect. 8.1). Except for this renormalization, the entanglement of the singlet or triplet state is not affected by the interacting electrons in the filled Fermi sea. Thus we can now calculate transport quantities using the standard scattering theory for noninteracting quasiparticles in a Fermi liquid. We consider the entangled incident states $|\pm\rangle \equiv |\psi_{\mathbf{12}}^{t/s}\rangle$ with one electron per lead and the quantum numbers $\mathbf{n} = (\varepsilon_n, n)$, where ε_n is the energy of the electron. Considering a multiterminal conductor with density of states ν, we assume that the leads consist of only one quantum channel; the generalization to several channels is straightforward. The (unpolarized) current operator for lead α can be written as [64]

$$I_\alpha(t) = \frac{e}{h\nu}\sum_{\sigma\varepsilon\varepsilon'}\left[a^\dagger_{\alpha\sigma}(\varepsilon)a_{\alpha\sigma}(\varepsilon') - b^\dagger_{\alpha\sigma}(\varepsilon)b_{\alpha\sigma}(\varepsilon')\right]e^{\mathrm{i}(\varepsilon-\varepsilon')t/\hbar}, \tag{21}$$

where $a^\dagger_{\alpha\sigma}(\varepsilon)$ creates an incoming electron with spin σ and energy ε in the lead α. The operators $b_{\alpha\sigma}(\varepsilon)$ for the outgoing electrons are given by $b_{\alpha\sigma}(\varepsilon) = \sum_\beta s_{\alpha\beta}a_{\beta\sigma}(\varepsilon)$ with the scattering matrix $s_{\alpha\beta}$, which is assumed to be spin- and energy-independent. The average currents in the leads, $|\langle I_\alpha\rangle| = e/h\nu$, are not sensitive to the orbital symmetry of the wave function. The spectral densities of the fluctuations $\delta I_\alpha = I_\alpha - \langle I_\alpha \rangle$ between the leads α and β are

$$S_{\alpha\beta}(\omega) = \lim_{T\to\infty} \frac{h\nu}{T} \int_0^T \mathrm{d}t\ e^{\mathrm{i}\omega t}\, \mathrm{Re}\langle \pm|\delta I_\alpha(t)\delta I_\beta(0)|\pm\rangle , \tag{22}$$

which are now evaluated with the scattering matrix for the beamsplitter (Fig. 4) with the reflection and transmission amplitudes r and t, thus $s_{31} = s_{42} = r$, and $s_{41} = s_{32} = t$ and no backscattering, so $s_{12} = s_{34} = s_{\alpha\alpha} = 0$. We obtain for the noise at zero frequency [59]

$$S_{33} = S_{44} = -S_{34} = 2\,\frac{e^2}{h\nu}\,T\,(1-T)\,(1 \mp \delta_{\varepsilon_1\varepsilon_2})\,. \tag{23}$$

Here, the minus (plus) sign refers to the spin triplet (singlet) and $T = |t|^2$ is the transmission coefficient of the beam splitter. If two electrons with the same energies, $\varepsilon_1 = \varepsilon_2$, in the singlet state are injected into the leads 1 and 2, the shot noise is enhanced by a factor of two compared to the value for uncorrelated particles [64,68], $2e^2T(1-T)/h\nu$. This amplification of the noise arises from *bunching* of the electrons due to their symmetric orbital wave function, such that the electrons preferably appear in the same outgoing leads. If the electron pairs are injected as a triplet, an *antibunching* effect appears, completely suppressing the noise, i.e. $S(\omega = 0) = 0$. We stress that the sign of cross-correlations does not carry any signature of statistics, e.g. here the different signs of S_{34} and $S_{33} = S_{44}$ (23) merely reflect current conservation and absence of backscattering. Since the bunching effect appears only for a state with a symmetric orbital wave function, which is not the case for unentangled electron states, measuring noise enhancement in the outgoing arms of the beamsplitter provides unique evidence for entanglement [59].

8.3 Spin-Dependent Current Through a Double Dot–Probing Entanglement

We turn now to a setup by which the entanglement of two electrons in a double-dot can be measured through current and noise [58]. For this we consider a double-dot which is weakly coupled, with tunneling amplitude Γ, to in-and outgoing leads at chemical potentials $\mu_{1,2}$. As shown in Fig. 5, the dots are put in parallel in contrast to the standard series connection. We work in the Coulomb blockade regime [27] where the charge on the dots is quantized and in the cotunneling regime [56,69], with $U > |\mu_1 \pm \mu_2| > J > k_BT, 2\pi\nu\Gamma^2$, where U is the single-dot charging energy, ν the lead density of states, and J the exchange coupling (see Sect. 5). The cotunneling current involves a coherent virtual process where an electron tunnels from a dot to, say, lead 2 and then a second electron tunnels from lead 1 to this dot. Assuming $|\mu_1 - \mu_2| > J$, elastic as well as inelastic cotunneling occurs. Further, Γ is assumed to be sufficiently weak so that the double-dot will return to its equilibrium state before the next electron passes through. Since an electron can either pass through the upper or lower dot, a closed loop is formed by these two paths, and in the presence of a magnetic flux the upper and the lower paths collect a phase difference given by the Aharonov-Bohm phase $\phi = ABe/\hbar$ (with A being the loop area), thus leading to interference effects. If the two electrons on the double-dot are in the *singlet state*, then the tunneling current acquires an additional phase of π (see below and Fig. 5) leading to a sign reversal of the coherent contribution compared to that for triplets. Explicitly, we find for the cotunneling current [58]

$$I = e\pi\nu^2\Gamma^4 \frac{\mu_1 - \mu_2}{\mu_1\mu_2} \, (2 \pm \cos\phi) \, , \tag{24}$$

and for the shot noise power $S(0) = -e|I|$, where the upper sign refers to the triplet states in the double-dot and the lower sign to the singlet state.

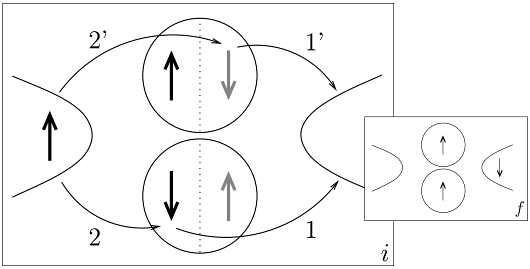


Fig. 5. Two coupled quantum dots with tunnel contacts to in- and outgoing leads to probe the entanglement on the dot (see text). The large box shows an initial state i with one spin-up electron in the left lead and two electrons on the double dot in state $(|{\uparrow\downarrow}\rangle \pm |{\downarrow\uparrow}\rangle)/\sqrt{2}$, where the first term is drawn in black in the left part of the dots and the second term in gray on the right. After the tunneling processes 1, 2 or $1'$, $2'$, the final state f is reached, where a spin-down electron is in the right lead and the state on the dots is $|T_+\rangle = |{\uparrow\uparrow}\rangle$, as shown in the small box.

Equation (24) can be reproduced, up to a prefactor, by the following heuristic argument. Consider the two spins on the double dot to be in the singlet state $|S\rangle = (|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle)/\sqrt{2}$ or in a triplet state, say, $|T_0\rangle = (|{\uparrow\downarrow}\rangle + |{\downarrow\uparrow}\rangle)/\sqrt{2}$. These superpositions are illustrated in Fig. 5 by drawing the first term in black in the left part of the dots and the second term in gray on the right. We consider the contribution I_{T_+} to the current, where we start with one spin-up electron in the left lead and end with a spin-down electron in the right lead and the triplet state $|T_+\rangle$ on the double dot (see inset of Fig. 5). For this process, either a spin-down electron tunnels first (1) from the lower dot into the right lead and then (2) the spin-up electron from the left lead tunnels into the lower dot. Or the upper dot participates via ($1'$) and ($2'$), but now the state $|{\downarrow\uparrow}\rangle$ is involved, thus if the initial state on the double dot is a singlet, the transition amplitudes for upper and lower path acquire opposite signs, whereas there is no sign change if we started out from a triplet (as shown for $|T_0\rangle$ in Fig. 5). Therefore, we can write the transition amplitudes $A_{21} = |A_{21}|e^{i\phi/2} \propto \Gamma^2$ for the lower path and $A_{2'1'} = \pm|A_{21}|e^{-i\phi/2}$ for the upper path, where the upper/lower sign stands for a triplet/singlet initial state on the double-dot. This leads to a total transition amplitude of $A_{fi} = A_{21} + A_{2'1'}$, and a current $I_{T_+} \propto e|A_{fi}|^2 = 2e|A_{21}|^2(1 \pm \cos\phi)$. Note that the transition $|S\rangle \rightarrow |T_+\rangle$ is inelastic whereas $|T_0\rangle \rightarrow |T_+\rangle$ is not. For an initial singlet state on the double-dot, the other inelastic processes $|S\rangle \rightarrow |T_0\rangle$, $|T_-\rangle$ also yield a current proportional to $1 - \cos\phi$, while the current from the elastic process $|S\rangle \rightarrow |S\rangle$ is proportional to $1 + \cos\phi$. Similarly, starting with a triplet, the sign of the $\cos\phi$ term is negative for an inelastic process, while it is positive for an elastic one. Note that there is only one inelastic process $|T\rangle \rightarrow |S\rangle$, whereas there are more elastic processes allowed for $|T\rangle \rightarrow |T\rangle$. The total current is obtained by summing over all terms, yielding $I = \sum_f I_f \propto e\Gamma^4(2 \pm \cos\phi)$, where the upper sign stands for an initial triplet state and the lower sign for a singlet, in agreement with (24). We finally emphasize that for the singlet $|S\rangle$ and for the triplet $|T_0\rangle$ the double-dot state is entangled, i.e. a correlated two-particle state, and thus the proposed setup probes a genuine two-particle interference effect via the Aharonov-Bohm oscillations in the current (noise). Note also that we can continuously transmute the statistics from fermionic to bosonic (like for anyons): the symmetric orbital part of $|S\rangle$ goes into an antisymmetric one at half a flux quantum, and vice versa for $|T_0\rangle$.

We have evaluated the noise also for finite frequencies [58], and found that again $S(\omega) \propto (2 \pm \cos\phi)$, and, moreover, that the odd part of $S(\omega)$ leads to slowly decaying oscillations of the noise in real time, $S(t) \propto \sin(\mu t)/\mu t$, $\mu = (\mu_1 + \mu_2)/2$, which can be ascribed to a charge imbalance on the double dot during an uncertainty time μ^{-1}.

We finally note that the three triplets can be further distinguished by an orientationally inhomogeneous magnetic field which results in a spin-Berry phase [70,58] that leads to left, right or no phase-shift in the Aharonov-Bohm oscillations of the current (noise).

8.4 Double Dot with Superconducting Leads

A further scenario of double-dots has been considered [46], where the dots are aligned in parallel between the leads, as in Sect. 8.3, but now no direct coupling is assumed between them. However, they are coupled with a tunneling amplitude Γ to two superconducting leads. The s-wave superconductor favors an entangled singlet-state on the dots (like in a Cooper pair) and further provides a mechanism for detecting the spin state via the Josephson current. It turns out that in leading order $\propto \Gamma^4$ the spin coupling is again described by a Heisenberg Hamiltonian [46]

$$H_{\text{eff}} \approx J\,(1 + \cos\varphi)\left(\mathbf{S}_a \cdot \mathbf{S}_b - \frac{1}{4}\right)\,, \tag{25}$$

where $J \approx 2\Gamma^2/\epsilon$, and the energy of the dot is ϵ below the lead Fermi energy. Here, φ is the average phase difference across the superconductor–double-dot–superconductor (S-DD-S) junction. We can modify the exchange coupling between the spins by tuning the external control parameters Γ and φ. Thus, we have presented here another implementation of a two-qubit quantum gate (see Sect. 5) or an “entangler” for EPR transport (see Sect. 8.2). Furthermore, the spin state on the dot can be probed if the superconducting leads are joined with one additional (ordinary) Josephson junction with coupling J' and phase difference θ into a SQUID-ring. The supercurrent I_S through this ring is given by [46]

$$I_S/I_J = \begin{cases} \sin(\theta - 2\pi f) + (J'/J)\sin\theta\,, & \text{singlet,} \\ (J'/J)\sin\theta\,, & \text{triplets,} \end{cases} \tag{26}$$

where $I_J = 2eJ/\hbar$. Measurement of the spin- and flux-dependent critical current $I_c = \max_\theta\{|I_S|\}$ probes the spin state of the double dot. This is realized by biasing the system with a dc current I until a finite voltage V appears for $|I| > I_c$ [46].

Acknowledgments

We thank K. Ensslin, E.V. Sukhorukov, and P. Recher for many discussions. This work has been supported by the Swiss NSF, NCCR Nanoscience, DARPA SPINS, QUIST, and ARO.

References

1. G. Prinz: Phys. Today **45**(4), 58 (1995); G. A. Prinz: Science **282**, 1660 (1998)
2. J. M. Kikkawa, I. P. Smorchkova, N. Samarth, D. D. Awschalom: Science **277**, 1284 (1997); J. M. Kikkawa, D. D. Awschalom: Phys. Rev. Lett. **80**, 4313 (1998); D. D. Awschalom, J. M. Kikkawa: Phys. Today **52**(6), 33 (1999)
3. R. Fiederling, M. Keim, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, L. W. Molenkamp: Nature **402**, 787 (1999)
4. Y. Ohno, D. K. Young, B Beschoten, F. Matsukura, H. Ohno, D. D. Awschalom: Nature **402**, 790 (1999)
5. F. G. Monzon, M. L. Roukes: J. Magn. Magn. Mater. **198**, 632 (1999)
6. S. Lüscher, T. Heinzel, K. Ensslin, W. Wegscheider, M. Bichler: Phys. Rev. Lett. **86**, 2118 (2001); cond-mat/0002226
7. For an extended review, see G. Burkard, D. Loss: in *Semiconductor Spintronics and Quantum Computation*, eds. D. D. Awschalom, D. Loss, N. Samarth (Springer, 2002)
8. D. Loss, D. P. DiVincenzo: Phys. Rev. A **57**, 120 (1998); cond-mat/9701055
9. For an earlier review, see G. Burkard, H.-A. Engel, D. Loss: Fortschr. Phys. **48**, 9-11, 965-986 (2000), Special Issue on *Experimental Proposals for Quantum Computation*, eds. S. L. Braunstein, H.-K. Lo (re-published as a book: *Scalable Quantum Computers: Paving the Way to Realization*)
10. A. Steane: Rep. Prog. Phys. **61**, 117 (1998)
11. D. P. DiVincenzo, D. Loss: J. Magn. Magn. Mater. **200**, 202 (1999); cond-mat/9901137
12. C. H. Bennett, D. P. DiVincenzo: Nature **404**, 247 (2000)
13. P. Recher, E. V. Sukhorukov, D. Loss: Phys. Rev. B **63**, 165314 (2001)
14. P. Recher, D. Loss: Phys. Rev. B **65**, 165327 (2002)

15. D. P. DiVincenzo, G. Burkard, D. Loss, E. Sukhorukov: in *Quantum Mesoscopic Phenomena and Mesoscopic Devices in Microelectronics*, eds. I.O. Kulik, R. Ellialtoglu (NATO ASI, Turkey, June 13-25, 1999); see cond-mat/99112445
16. P. W. Shor: in *Proc. 35th Symposium on the Foundations of Computer Science*, (IEEE Computer Society Press), 124 (1994)
17. L. K. Grover: Phys. Rev. Lett. **79**, 325 (1997)
18. D. P. DiVincenzo: Phys. Rev. A **51**, 1015 (1995)
19. J. I. Cirac, P. Zoller: Phys. Rev. Lett. **74**, 4091 (1995); C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, D. J. Wineland: *ibid.* **75**, 4714 (1995)
20. Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, H. J. Kimble: Phys. Rev. Lett. **75**, 4710 (1995)
21. D. Cory, A. Fahmy, T. Havel: Proc. Nat. Acad. Sci. U.S.A. **94**, 1634 (1997); N. A. Gershenfeld, I. L. Chuang: Science **275**, 350 (1997)
22. V. Privman, I. D. Vagner, G. Kventsel: Phys. Lett. A **239** (1998) 141
23. B. Kane: Nature **393**, 133 (1998)
24. Yu. Makhlin, G. Schön, A. Shnirman: Rev. Mod. Phys. **73**, 357 (2001)
25. C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, W. K. Wootters: Phys. Rev. Lett. **70**, 1895 (1993)
26. C. H. Bennett, G. Brassard: in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India (IEEE, NY, 1984), p. 175
27. L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M. Westervelt, N. S. Wingreen: Proceedings of the ASI on *Mesoscopic Electron Transport*, eds. L.L. Sohn, L.P. Kouwenhoven, G. Schön (Kluwer, 1997)
28. S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, L. P. Kouwenhoven: Phys. Rev. Lett. **77**, 3613 (1996)
29. F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. L. Campman, A. C. Gossard: Phys. Rev. Lett. **75**, 705 (1995); C. Livermore, C. H. Crouch, R. M. Westervelt, K. L. Campman, A. C. Gossard: Science **274**, 1332 (1996)
30. T. H. Oosterkamp S. F. Godijn, M. J. Uilenreef, Y. V. Nazarov, N. C. van der Vaart, L. P. Kouwenhoven: Phys. Rev. Lett. **80**, 4951 (1998)
31. R. H. Blick, D. Pfannkuche, R. J. Haug, K. v. Klitzing, K. Eberl: Phys. Rev. Lett. **80**, 4032 (1998); *ibid.* **81**, 689 (1998). T. H. Oosterkamp, T. Fujisawa, W. G. van der Wiel, K. Ishibashi, R. V. Hijman, S. Tarucha, L. P. Kouwenhoven: Nature **395**, 873 (1998); I. J. Maasilta, V.J. Goldman: Phys. Rev. Lett. **84**, 1776 (2000)
32. J. A. Gupta, D. D. Awschalom, X. Peng, A. P. Alivisatos: Phys. Rev. B **59**, R10421 (1999)
33. A. V. Khaetskii, D. Loss, L. Glazman: Phys. Rev. Lett. **88**, 186802 (2002)
34. G. Burkard, D. Loss, D. P. DiVincenzo: Phys. Rev. B **59**, 2070 (1999)
35. A. Imamoḡlu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, A. Small: Phys. Rev. Lett. **83**, 4204 (1999)
36. G. Burkard, D. Loss, D.P. DiVincenzo, J.A. Smolin: Phys. Rev. B **60**, 11404 (1999)
37. P. W. Shor: Phys. Rev. A **52**, R2493 (1995); A. M. Steane: Phys. Rev. Lett. **77**, 793 (1996); D. P. DiVincenzo, P. W. Shor: *ibid.* **77**, 3260 (1996); E. Knill, R. Laflamme: Phys. Rev. A **55**, 900 (1997); D. Gottesman: *ibid.* **54**, 1862 (1996); E. Dennis: quant-ph/9905027
38. L. Kouwenhoven, C. Marcus: private communication
39. M. Dobers, K. v. Klitzing , J. Schneider, G. Weimann, K. Ploog: Phys. Rev. Lett. **61**, 1650 (1988)
40. D. C. Dixon, K. R. Wald, P. L. McEuen, M. R. Melloch: Phys. Rev. B **56**, 4743 (1997)

41. P. Recher, E. V. Sukhorukov, D. Loss: Phys. Rev. Lett. **85**, 1962 (2000); cond-mat/0003089
42. R. Shankar: *Principles of Quantum Mechanics*, Ch. 14, Plenum Press, New York, 1994
43. D. P. DiVincenzo: Phys. Rev. A **51**, 1015 (1995)
44. A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, H. Weinfurter: Phys. Rev. A **52**, 3457 (1995)
45. R. Vrijen, E. Yablonovitch, K. Wang, H. W. Jiang, A. Balandin, V. Roychowdhury, T. Mor, D. DiVincenzo: Phys. Rev. A **62**, 012306 (2000)
46. M.-S. Choi, C. Bruder, D. Loss: Phys. Rev. B **62**, 13569 (2000)
47. G. Burkard, G. Seelig, D. Loss: Phys. Rev. B **62**, 2581 (2000)
48. R. J. Luyken, A. Lorke, M. Haslinger, B. T. Miller, M. Fricke, J. P.Kotthaus, G. Medeiros-Ribiero, P. M. Petroff: Physica E **2**, 704 (1998)
49. D. G. Austing, T. Honda, K. Muraki, Y. Tokura, S. Tarucha: Physica B **249-251**, 206 (1998)
50. E. L. Ivchenko, A. A. Kiselev, M. Willander: Solid State Comm. **102**, 375 (1997)
51. K. Ensslin: private communication
52. A. Peres: *Quantum Theory: Concepts and Methods* (Kluwer, Dondrecht, 1993)
53. M. Devoret, D. Estève, Ch. Urbina: Nature (London) **360**, 547 (1992)
54. G. D. Mahan: *Many Particle Physics*, 2nd Ed. (Plenum, New York, 1993)
55. L. P. Kouwenhoven, G. Schön, L. L. Sohn: *Mesoscopic Electron Transport*, NATO ASI Series E: Applied Sciences-Vol. 345 Kluwer Academic Publishers, 1997
56. D. V. Averin, Yu. V. Nazarov: in *Single Charge Tunneling*, eds. H. Grabert, M. H. Devoret, NATO ASI Series B: Physics Vol. 294, Plenum Press, New York, 1992
57. M. Ciorga, A. S. Sachrajda, P.Hawrylak, C. Gould, P. Zawadzki, S.Jullian, Y. Feng, Z. Wasilewski: Phys. Rev. B **61**, R16315 (2000)
58. D. Loss, E. V. Sukhorukov: Phys. Rev. Lett. **84**, 1035 (2000)
59. G. Burkard, D. Loss, E. V. Sukhorukov: Phys. Rev. B **61**, R16303 (2000); cond-mat/9906071
60. A. Aspect, J. Dalibard, G. Roger: Phys. Rev. Lett. **49**, 1804 (1982); W. Tittel, J. Brendel, H. Zbinden, N. Gisin: Phys. Rev. Lett. **81**, 3563 (1998)
61. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, A. Zeilinger: Nature **390**, 575 (1997); D. Boschi, S. Branca, F. De Martini, L. Hardy, S. Popescu: Phys. Rev. Lett. **80**, 1121 (1998)
62. R. Loudon: Phys. Rev. A **58**, 4904 (1998)
63. R. Hanbury Brown, R. Q. Twiss: Nature (London) **177**, 27 (1956)
64. M. Büttiker: Phys. Rev. Lett. **65**, 2901 (1990); Phys. Rev. B **46**, 12485 (1992)
65. T. Martin, R. Landauer: Phys. Rev. B **45**, 1742 (1992).
66. E. V. Sukhorukov, D. Loss: Phys. Rev. B **59**, 13054 (1999).
67. R. C. Liu, B. Odom, Y. Yamamoto, S. Tarucha: Nature **391**, 263 (1998); M. Henny, S. Oberholzer, C. Strunk, T. Heinzel, K. Ensslin, M. Holland, C. Schönenberger: Science **284**, 296 (1999); W. D. Oliver, J. Kim, R. C. Liu, Y. Yamamoto: *ibid.*, 299 (1999)
68. V. A. Khlus: Zh. Eksp. Teor. Fiz. **93**, 2179 (1987)
69. J. König, H. Schoeller, G. Schön: Phys. Rev. Lett. **78**, 4482 (1997)
70. D. Loss, P. Goldbart: Phys. Rev. B **45**, 13544 (1992)