

**Giuliano Benenti**   **Giulio Casati**  
**Davide Rossini**   **Giuliano Strini**

S	A	T	O	R
A	R	E	P	O
T	E	N	E	T
O	P	E	R	A
R	O	T	A	S

**Principles of Quantum Computation  
and Information**  
A Comprehensive Textbook

**World Scientific**

# Principles of Quantum Computation and Information

A Comprehensive Textbook

**This page intentionally left blank**

# Principles of Quantum Computation and Information

A Comprehensive Textbook

**Giuliano Benenti and Giulio Casati**

*Università degli Studi dell'Insubria, Italy*

**Davide Rossini**

*Università di Pisa, Italy*

**Giuliano Strini**

*Università di Milano, Italy*



NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

*Published by*

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

*USA office:* 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

*UK office:* 57 Shelton Street, Covent Garden, London WC2H 9HE

**Library of Congress Cataloging-in-Publication Data**

Names: Benenti, Giuliano, 1969– author. | Casati, Giulio, 1942– author. | Rossini, Davide, 1979– author. | Strini, Giuliano, 1937– author.

Title: Principles of quantum computation and information / Giuliano Benenti and Giulio Casati (Università degli Studi dell'Insubria, Italy), Davide Rossini (Scuola Normale Superiore, Pisa, Italy), Giuliano Strini (Università di Milano, Italy).

Description: Second edition. | Singapore ; Hackensack, NJ : World Scientific Publishing Co. Pte. Ltd., [2019] | Includes bibliographical references and index.

Identifiers: LCCN 2018051847| ISBN 9789813237223 (hardcover ; alk. paper) | ISBN 9789813279995 (pbk. ; alk. paper)

Subjects: LCSH: Quantum computers.

Classification: LCC QA76.889 .B46 2019 | DDC 006.3/843--dc23

LC record available at <https://lccn.loc.gov/2018051847>

**British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library.

Copyright © 2019 by World Scientific Publishing Co. Pte. Ltd.

*All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the publisher.*

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

For any available supplementary material, please visit  
<https://www.worldscientific.com/worldscibooks/10.1142/10909#t=suppl>

*To Silvia, Arianna and Isabella*  
g.b.

*To my wife for her love and encouragement*  
g.c.

*To Carlo*  
d.r.

*To my family and friends*  
g.s.

## ***About the Cover***

This acrostic is the famous *sator* formula. It can be translated, among possible different interpretations, as:

*'Arepo the sower holds the wheels at work'*

The text may be read in four different ways:

- (i) horizontally, from left to right (downward) and from right to left (upward);
- (ii) vertically, downward (left to right) and upward (right to left).

The resulting phrase is always the same.

It has been suggested that it might be a form of secret message.

This acrostic was unearthed during archeological excavation work at Pompeii, which was buried, as well known, by the eruption of Vesuvius in 79 A.D. The formula can be found throughout the Roman Empire, probably also spread by legionnaires. Moreover, it has been found in Mesopotamia, Egypt, Cappadocia, Britain and Hungary.

The *sator* acrostic may have a mystical significance and might have been used as a means for persecuted Christians to recognize each other (it can be rearranged into the form of a cross, with the opening words of the Lord's prayer, *A Paternoster O*, both vertically and horizontally, intersecting at the letter N, the Latin letters A and O corresponding to the Greek letters alpha and omega, beginning and end of all things).

# Preface

## *Purpose of the book*

This book can be used as a broad range textbook for a course in quantum information and computation, both for upper-level undergraduate students and for graduate students. It may also be useful as general education for readers who want to know the fundamental principles of quantum information and computation and who have the basic background in physics and mathematics acquired in their undergraduate courses in physics, mathematics, or computer science, as well as for researchers interested in some of the latest spin-off of the field, including the use of quantum information in the theories of condensed matter and many-body systems.

Note that, to address not only students in physics, but also in mathematics, computer science, chemistry, or engineering, no prior knowledge of either quantum mechanics or classical computation is required to follow this book. Indeed, the first two chapters are a simple introduction to classical computation and quantum mechanics. Our aim is that these chapters should provide the necessary background for an understanding of the subsequent chapters. Chapters 3-5 develop the basic principles and discuss the main results of quantum computation and communication. The first five chapters would thus be suitable for a one-semester introductory course in quantum information and computation, for both undergraduate and graduate students.

Chapters 6-11 deal with various, more advanced, important aspects of quantum computation and information. The areas include entanglement measures, quantum discord, the Kraus and the Bloch-Fano representations, the GKLS master equation, non-Markovian quantum dynamics, quantum to classical transition, quantum measurements, quantum trajectories, quantum data compression, accessible information, quantum channels, quantum error correction, stabilizer coding, decoherence-free subspaces, dynamical decoupling, the Zeno effect, fault-tolerant quantum computation, the basic principles of cavity quantum electrodynamics, trapped ions and solid-state qubits, quantum simulators, area-law scaling of entanglement, DMRG and basic tensor-network structures.

### *General approach*

Quantum computation and information is a rapidly developing field. It is therefore not easy to grasp the fundamental concepts and central results without having to face many technical details. Our purpose in this book is to provide the reader interested in the field with a useful guide. Mathematical rigour is therefore not our primary concern. Instead, we have tried to present a simple and systematic treatment, such that the reader might understand the presented material without the need for consulting other texts.

### *Note to the reader*

Some of the material presented is not necessary for understanding the rest of the book and may be omitted on a first reading. We have adopted two methods of highlighting such parts:

- 1) The sections or subsections with an asterisk before the title contain more advanced or complementary material. Such parts may be omitted without risk of encountering problems in reading the rest of the book.
- 2) Comments, notes or examples are printed in a small typeface.

To gain complete familiarity with the subject, it is important to practice problem solving. The book contains a large number of exercises (with solutions), which are an essential complement to the main text. In order to develop a solid understanding of the arguments dealt with here, it is indispensable that the student try to solve a large part of them.

### *Further reading*

We shall conclude each chapter with a short guide to the bibliography. Our aim is to give general references that might be used by the reader as an entry point for a more in-depth analysis of the topics discussed in this book. We shall therefore often refer to review papers instead of the original articles.

### *About the second edition*

In the eleven years between the first and this (second) edition, there has been a tumultuous development in the field, both experimental and theoretical. On the experimental side, we have witnessed the emergence of quantum cryptography on the market, the first steps towards a global, satellite-based, quantum communication network, and the progress of quantum simulators. On the theoretical side, the cross-fertilization between research devoted to quantum information science and to quantum many-body systems has led to new ideas, methods and insights in both fields. These considerations have pushed us to write a second edition. While we have tried to retain as much as possible the spirit of the first, new material has been added on entanglement quantifiers, quantum discord, non-Markovian dynamics, quantum simulators, and in particular, on quantum information in condensed matter and strongly correlated quantum many-body systems.

*Acknowledgments*

We are indebted to several colleagues and students for valuable suggestions. In particular, we gratefully acknowledge Mari-Carmen Bañuls and Karol Życzkowski for their critical reading of parts of the manuscript. Obviously no responsibility should be attributed to any of the above regarding possible flaws that might remain, for which the authors alone are to blame.

**This page intentionally left blank**

# Contents

<i>Preface</i>	vii
<i>Introduction</i>	1
1. Introduction to classical computation	9
1.1 The Turing machine . . . . .	9
1.1.1 Addition on a Turing machine . . . . .	11
1.1.2 The Church–Turing thesis . . . . .	13
1.1.3 The universal Turing machine . . . . .	14
1.1.4 The probabilistic Turing machine . . . . .	15
1.1.5 * The halting problem . . . . .	15
1.2 The circuit model of computation . . . . .	16
1.2.1 Binary arithmetics . . . . .	17
1.2.2 Elementary logic gates . . . . .	18
1.2.3 Universal classical computation . . . . .	21
1.3 Computational complexity . . . . .	22
1.3.1 Tractable vs. intractable problems . . . . .	23
1.3.2 Complexity classes . . . . .	30
1.3.3 * The Chernoff bound . . . . .	34
1.4 * Computing dynamical systems . . . . .	34
1.4.1 * Deterministic chaos . . . . .	35
1.4.2 * Algorithmic complexity . . . . .	37
1.5 Energy and information . . . . .	39
1.5.1 Maxwell’s demon . . . . .	39
1.5.2 Landauer’s principle . . . . .	40
1.5.3 Extracting work from information . . . . .	42
1.6 Reversible computation . . . . .	43
1.6.1 Toffoli and Fredkin gates . . . . .	45
1.6.2 * The billiard-ball computer . . . . .	46
1.7 * Energy dissipation in computation . . . . .	48
1.7.1 * Experimental realization of a Maxwell’s demon . . . . .	48

1.7.2	* Experimental verification of Landauer's principle . . . . .	49
1.7.3	* Energy dissipation in real classical computer . . . . .	50
1.7.4	* Experimental realization of reversible computers . . . . .	51
1.7.5	* Neuromorphic computing . . . . .	53
1.8	A guide to the bibliography . . . . .	54
2.	Introduction to quantum mechanics	55
2.1	The Stern–Gerlach experiment . . . . .	56
2.2	Young's double-slit experiment . . . . .	59
2.3	The postulates of quantum mechanics . . . . .	63
2.3.1	Dynamical evolution . . . . .	63
2.3.2	Outcomes of a measurement . . . . .	64
2.3.3	The post-measurement state . . . . .	67
2.3.4	Heisenberg's uncertainty principle . . . . .	69
2.4	The EPR paradox . . . . .	72
2.5	Bell's inequalities . . . . .	77
2.6	The density matrix . . . . .	81
2.6.1	Composite systems . . . . .	86
2.7	The Schmidt decomposition . . . . .	89
2.8	Purification . . . . .	91
2.9	Generalized measurements . . . . .	93
2.9.1	POVM measurements . . . . .	94
2.10	A guide to the bibliography . . . . .	96
3.	Quantum computation	97
3.1	The qubit . . . . .	98
3.1.1	Pure qubit states: The Bloch sphere . . . . .	99
3.1.2	Mixed qubit states: The Bloch ball . . . . .	100
3.2	Measuring the state of a qubit . . . . .	103
3.2.1	Pure qubit states . . . . .	103
3.2.2	Mixed qubit states . . . . .	104
3.3	The circuit model of quantum computation . . . . .	105
3.4	Single-qubit gates . . . . .	108
3.4.1	Rotations of the Bloch sphere . . . . .	109
3.5	Controlled gates and entanglement generation . . . . .	111
3.5.1	The Bell basis . . . . .	115
3.6	Hamiltonian model for one- and two-qubit gates . . . . .	116
3.7	Universal quantum gates . . . . .	117
3.7.1	* Preparation of the initial state . . . . .	124
3.8	Unitary errors . . . . .	127
3.9	Function evaluation . . . . .	128
3.10	* The quantum adder . . . . .	132

3.11	Adiabatic theorem . . . . .	134
3.11.1	Adiabatic condition . . . . .	136
3.11.2	Berry phase . . . . .	137
3.12	* Non-Abelian geometric phase . . . . .	141
3.13	Adiabatic quantum computation . . . . .	145
3.14	* Maximum speed of quantum gates . . . . .	149
3.14.1	* Speed limit of an autonomous time evolution . . . . .	150
3.14.2	* Speed limit of single-qubit gates . . . . .	151
3.15	* Holonomic quantum computation . . . . .	152
3.16	A guide to the bibliography . . . . .	155
4.	Quantum algorithms	157
4.1	Deutsch's algorithm . . . . .	157
4.1.1	The Deutsch–Jozsa problem . . . . .	158
4.1.2	* An extension of Deutsch's algorithm . . . . .	159
4.2	Quantum search . . . . .	161
4.2.1	Searching one item out of four . . . . .	161
4.2.2	Searching one item out of $N$ . . . . .	163
4.2.3	Geometric visualization . . . . .	164
4.2.4	Searching by adiabatic quantum evolution . . . . .	166
4.3	The quantum Fourier transform . . . . .	168
4.4	Quantum phase estimation . . . . .	171
4.5	* Finding eigenvalues and eigenvectors . . . . .	173
4.6	Period finding and Shor's algorithm . . . . .	175
4.7	Quantum computation of dynamical systems . . . . .	179
4.7.1	Quantum simulation of the Schrödinger equation . . . . .	179
4.7.2	* The quantum baker's map . . . . .	183
4.7.3	* The quantum sawtooth map . . . . .	185
4.7.4	Information extraction for dynamical quantum systems . . . . .	188
4.8	Universal quantum simulation . . . . .	190
4.9	A guide to the bibliography . . . . .	192
5.	Quantum communication	195
5.1	Classical cryptography . . . . .	195
5.1.1	The Vernam cypher . . . . .	196
5.1.2	The public-key cryptosystem . . . . .	197
5.1.3	The RSA protocol . . . . .	198
5.2	The no-cloning theorem . . . . .	199
5.2.1	Faster-than-light transmission of information? . . . . .	201
5.2.2	* The no-signalling condition . . . . .	203
5.2.3	* Universal quantum cloning . . . . .	204
5.2.4	* The universal-NOT gate . . . . .	206

5.3	Quantum cryptography . . . . .	207
5.3.1	The BB84 protocol . . . . .	207
5.3.2	The E91 protocol . . . . .	210
5.4	Dense coding . . . . .	212
5.5	Quantum teleportation . . . . .	215
5.5.1	* Conclusive teleportation . . . . .	219
5.6	Quantum mechanics with continuous variables . . . . .	220
5.6.1	* General framework for Gaussian states . . . . .	233
5.7	Quantum cryptography with continuous variables . . . . .	237
5.8	A guide to the bibliography . . . . .	240
6.	Entanglement and non-classical correlations	241
6.1	Definition of entanglement . . . . .	241
6.1.1	Basic properties . . . . .	242
6.2	Bipartite separability criteria . . . . .	243
6.2.1	The Peres separability criterion . . . . .	244
6.2.2	Positive maps . . . . .	245
6.2.3	Entanglement witnesses . . . . .	246
6.2.4	Positive maps and witnesses . . . . .	248
6.3	The Shannon entropy . . . . .	248
6.3.1	Mutual information . . . . .	250
6.4	The von Neumann entropy . . . . .	252
6.4.1	Example 1: source of orthogonal pure states . . . . .	255
6.4.2	Example 2: source of non-orthogonal pure states . . . . .	255
6.5	Entanglement concentration . . . . .	256
6.5.1	* Entanglement of a random state . . . . .	260
6.6	Requirements for bipartite entanglement measures . . . . .	263
6.7	Other entanglement measures . . . . .	264
6.7.1	* Concurrence . . . . .	264
6.7.2	* Negativity . . . . .	265
6.8	* Multipartite entanglement . . . . .	266
6.8.1	* Monogamy of entanglement and tangle measures . . . . .	268
6.9	Quantum discord . . . . .	269
6.9.1	Definition . . . . .	270
6.9.2	Basic properties . . . . .	273
6.9.3	Examples . . . . .	274
6.9.4	* Other measures of quantum correlations . . . . .	275
6.10	* Quantum discord in continuous systems . . . . .	277
6.10.1	* Entropy of a Gaussian state . . . . .	277
6.10.2	* Discord of a Gaussian state . . . . .	278
6.11	* Entropies in physics . . . . .	279
6.11.1	* Thermodynamic entropy . . . . .	280

6.11.2 * Statistical entropy . . . . .	282
6.11.3 * Dynamical Kolmogorov–Sinai entropy . . . . .	284
6.12 A guide to the bibliography . . . . .	285
7. Decoherence . . . . .	287
7.1 The Kraus representation . . . . .	287
7.2 Decoherence models for a single qubit . . . . .	293
7.2.1 The quantum black box . . . . .	294
7.2.2 Measuring a quantum operation acting on a qubit . . . . .	295
7.2.3 Quantum circuits simulating noise channels . . . . .	296
7.2.4 The bit-flip channel . . . . .	298
7.2.5 The phase-flip channel . . . . .	299
7.2.6 The bit-phase-flip channel . . . . .	300
7.2.7 The depolarizing channel . . . . .	301
7.2.8 Amplitude damping . . . . .	302
7.2.9 Phase damping . . . . .	303
7.2.10 De-entanglement . . . . .	305
7.3 * The Bloch-Fano representation . . . . .	307
7.3.1 * Bloch-Fano representation of a state . . . . .	307
7.3.2 * Bloch-Fano representation of a quantum operation . . . . .	308
7.4 The master equation . . . . .	310
7.4.1 * Derivation of the master equation . . . . .	311
7.4.2 The master equation and quantum operations . . . . .	319
7.4.3 The master equation for a single qubit . . . . .	322
7.5 * Non-Markovian quantum dynamics . . . . .	324
7.6 Quantum to classical transition . . . . .	329
7.6.1 Schrödinger’s cat . . . . .	329
7.6.2 Decoherence and destruction of cat states . . . . .	330
7.7 Decoherence and quantum measurements . . . . .	335
7.7.1 * Weak measurements . . . . .	337
7.7.2 * Decoherence and quantum trajectories . . . . .	340
7.8 A guide to the bibliography . . . . .	344
8. Quantum information theory . . . . .	345
8.1 Classical data compression . . . . .	346
8.1.1 Shannon’s noiseless coding theorem . . . . .	346
8.1.2 Examples of data compression . . . . .	348
8.1.3 Capacity of classical channels . . . . .	349
8.2 Quantum data compression . . . . .	351
8.2.1 Schumacher’s quantum noiseless coding theorem . . . . .	351
8.2.2 Compression of an $n$ -qubit message . . . . .	352
8.2.3 Example 1: two-qubit messages . . . . .	354

8.2.4	Example 2: three-qubit messages . . . . .	355
8.3	Accessible information . . . . .	357
8.3.1	The Holevo bound . . . . .	358
8.3.2	Example 1: two non-orthogonal pure states . . . . .	359
8.3.3	* Example 2: three non-orthogonal pure states . . . . .	362
8.4	Capacities of quantum channels . . . . .	363
8.4.1	Classical capacity . . . . .	364
8.4.2	Quantum capacity . . . . .	365
8.5	* Quantum memory channels . . . . .	371
8.6	A guide to the bibliography . . . . .	378
9.	Quantum error correction	379
9.1	The three-qubit bit-flip code . . . . .	381
9.2	The three-qubit phase-flip code . . . . .	384
9.3	The nine-qubit Shor code . . . . .	385
9.4	General properties of quantum error correction . . . . .	389
9.4.1	The quantum Hamming bound . . . . .	391
9.5	Stabilizer coding . . . . .	392
9.5.1	The nine-qubit Shor code revisited . . . . .	392
9.5.2	* General formalism for stabilizer codes . . . . .	394
9.5.3	* Logical operators for stabilizer codes . . . . .	395
9.6	* The five-qubit code . . . . .	396
9.7	Decoherence-free subspaces . . . . .	399
9.7.1	* Conditions for decoherence-free dynamics . . . . .	401
9.7.2	* The spin-boson model . . . . .	402
9.8	* Dynamical decoupling . . . . .	404
9.8.1	* Explicit form of control Hamiltonian . . . . .	406
9.9	* The Zeno effect . . . . .	407
9.10	Fault-tolerant quantum computation . . . . .	411
9.10.1	Avoidance of error propagation . . . . .	411
9.10.2	Fault-tolerant quantum gates . . . . .	413
9.10.3	The noise threshold for quantum computation . . . . .	414
9.11	A guide to the bibliography . . . . .	415
10.	Principles of experimental implementations of quantum protocols	417
10.1	Cavity quantum electrodynamics . . . . .	418
10.1.1	Interaction of a two-level atom with a classical field . . . . .	420
10.1.2	The Jaynes–Cummings model . . . . .	421
10.1.3	Rabi oscillations . . . . .	422
10.1.4	Entanglement generation . . . . .	423
10.2	The ion-trap quantum computer . . . . .	424
10.2.1	The Paul trap . . . . .	425

10.2.2	Laser pulses . . . . .	427
10.3	Solid-state qubits . . . . .	433
10.3.1	Spins in semiconductors . . . . .	433
10.3.2	Quantum dots . . . . .	434
10.3.3	Superconducting qubit circuits . . . . .	437
10.4	Quantum communication with photons . . . . .	443
10.4.1	Linear optics . . . . .	444
10.4.2	Non-linear optics and probabilistic gates . . . . .	446
10.4.3	Experimental quantum-key distribution . . . . .	449
10.5	Problems and prospects . . . . .	454
10.6	A guide to the bibliography . . . . .	455
11.	Quantum information in many-body systems	457
11.1	Quantum simulators . . . . .	458
11.1.1	Ultracold atoms . . . . .	458
11.1.2	Arrays of coupled QED cavities . . . . .	461
11.2	Emergence of quantum correlations . . . . .	466
11.2.1	The Hubbard model . . . . .	467
11.3	The spin-1/2 quantum Ising chain . . . . .	469
11.3.1	Jordan–Wigner transformation . . . . .	470
11.3.2	Diagonalization of the Ising chain . . . . .	472
11.3.3	Two-spin concurrence . . . . .	478
11.3.4	Entanglement block entropy . . . . .	479
11.3.5	The Ising model revisited: Kitaev chain . . . . .	481
11.4	Area-law scaling of the entanglement . . . . .	484
11.5	Matrix product states . . . . .	487
11.5.1	Examples of MPS wave functions . . . . .	490
11.6	Graphical representation of matrix product states . . . . .	492
11.6.1	Expectation values of observables . . . . .	494
11.6.2	* Scaling of correlation functions with the distance . . . . .	496
11.6.3	Gauge freedom . . . . .	498
11.6.4	Schmidt decomposition of a MPS . . . . .	500
11.7	Ground-state search in the Hilbert space corner . . . . .	503
11.7.1	Density-matrix renormalization group . . . . .	506
11.7.2	* DMRG as a variational optimization over the MPS class . . . . .	510
11.8	Time evolution of matrix product states . . . . .	512
11.8.1	Finite-temperature calculations . . . . .	515
11.8.2	Mixed-state time evolution . . . . .	517
11.9	* General tensor-network structures . . . . .	519
11.9.1	* Projected entangled pair states . . . . .	519
11.9.2	* Hierarchical tensor networks . . . . .	524
11.10	A guide to the bibliography . . . . .	526

Conclusions and prospects	529
Appendix A Elements of linear algebra	535
A.1 Finite-dimensional vector spaces	535
A.1.1 Basic properties of vector spaces	535
A.1.2 Inner product and norm of a vector	536
A.1.3 Linear independence and the notion of basis	538
A.1.4 Linear operators	540
A.1.5 Tensor product	545
A.1.6 Matrix decompositions	547
A.1.7 Symplectic decompositions	555
A.2 Infinite-dimensional vector spaces	557
A.2.1 Discrete and continuous bases	557
A.2.2 The Dirac delta function	558
A.2.3 Orthonormality and completeness relations	559
A.2.4 Position and momentum representations	560
A.2.5 Position and momentum operators	563
Appendix B Solutions to the exercises	565
B.1 Chapter 1	565
B.2 Chapter 2	566
B.3 Chapter 3	573
B.4 Chapter 4	584
B.5 Chapter 5	585
B.6 Chapter 6	594
B.7 Chapter 7	600
B.8 Chapter 8	615
B.9 Chapter 9	618
B.10 Chapter 10	625
B.11 Chapter 11	644
B.12 Appendix A	648
<i>Bibliography</i>	651
<i>Index</i>	677

# Introduction

Since its dawn at the beginning of the twentieth century, quantum mechanics has had an outstanding technological and societal impact. To appreciate this fact, it is sufficient to consider the invention of the transistor, perhaps the most remarkable among the countless other applications of quantum mechanics. On the other hand, the enormous impact of computers on everyday life is on the eye of the beholder. Their importance is such that it is appropriate to say that we are now living in the *information age*. Needless to say, this information revolution became possible thanks to the invention of the transistor, that is, thanks to the synergy between computer science and quantum physics.

Today this synergy paves the way for a cross-fertilization of ideas, which are offering completely new opportunities and promising exciting advances in both fundamental science and technological application. We are referring to the fact that **quantum mechanics can be used to process and transmit information**.

Miniaturization provides us with an intuitive way of understanding why, in the near future, quantum laws will become important for computation. The electronics industry for computers grows hand-in-hand with the decrease in size of integrated circuits. This miniaturization is necessary to increase computational power, that is, the number of floating-point operations per second (flops) a computer can perform. In the 1950's, electronic computers based on vacuum-tube technology were capable of performing approximately  $10^3$  floating-point operations per second, while nowadays (2018) there exists a supercomputer (Sunway TaihuLight in China) whose power is as great as 93 petaflops (1 petaflop is equal to  $10^{15}$  flops), and exascale computing is under development (that is, computing systems capable of at least 1 exaflops, equal to  $10^{18}$  flops). As we have already remarked, this enormous growth of computational power has been made possible owing to progress in miniaturization, which may be quantified empirically in Moore's law. This law is the result of a remarkable observation made by Gordon Moore, co-founder of Intel, in 1965: the number of transistors that may be placed on a single integrated-circuit chip doubles approximately every 18–24 months. This exponential growth has not yet saturated and Moore's law is still valid. At the time of writing this book, the limit

is approximately  $2\text{--}3 \times 10^{10}$  transistors per chip and the typical size of circuit components is as small as 12 nanometres (with 5–7 nm components under development). Extrapolating Moore’s law, one would estimate that around the year 2020 we shall reach the atomic size for storing a single bit of information. At that point, quantum effects will become unavoidably dominant.

It is clear that, notwithstanding quantum effects, other factors could bring Moore’s law to an end. In the first place, there are economic considerations. Indeed, the cost of building fabrication facilities to manufacture chips has also increased exponentially with time. Nevertheless, it is of the utmost importance to understand the ultimate limitations set by quantum mechanics. As a matter of fact, even though we might overcome economic barriers by means of technological breakthroughs, quantum physics sets fundamental limitations on the size of the circuit components. The first question under debate is whether it would be more convenient to push the silicon-based transistor to its physical limits or instead to develop alternative devices, such as quantum dots, single-electron transistors or molecular switches. A common feature of all these devices is that they are on the nanometre length scale, and therefore quantum effects play a crucial role.

So far, we have talked about quantum switches that could substitute silicon-based transistors and possibly be connected together to execute classical algorithms based on Boolean logic. In this perspective, quantum effects are simply unavoidable corrections that must be taken into account, owing to the nanometre size of the switches. A quantum computer would rather represent a radically different challenge: the aim is to build a machine *based on quantum logic*, that is, which processes the information and performs logic operations by exploiting the laws of quantum mechanics.

The unit of quantum information is known as the *qubit* (the quantum counterpart of the classical *bit*) and a quantum computer may be viewed as a many-qubit system. Physically, a qubit is a simple two-level system, like the two spin states of a spin- $\frac{1}{2}$  particle, the vertical and horizontal polarization states of a single photon, or the ground and excited states of an atom, just to quote a few examples. The state-of-the-art in the implementation of quantum bits closely resembles the situation in the construction of the transistors at the beginning of the 1960’s: at that time, there were tens of different technologies (and almost all required the manual assembly of single transistors), before the nowadays universally adopted photolithographic technology was invented and developed.

A quantum computer is a system of many qubits, whose evolution can be controlled, and a quantum computation is a sequence of unitary transformations that act on the many-qubit state describing the quantum computer. The power of such devices ultimately resides in typical quantum phenomena, such as the *superposition* of quantum states, *entanglement* and non-classical correlations. There is an inherent quantum parallelism associated with the superposition principle. In simple terms, a quantum computer can process a large number of classical inputs in a single

run. On the other hand, this implies a large number of possible outputs. It is the task of quantum algorithms, which are based on quantum logic, to exploit the inherent quantum parallelism of quantum mechanics to highlight the desired output. In short, to be useful, quantum computers require the development of appropriate quantum software, that is, of efficient quantum algorithms.

In the 1980's Feynman suggested that a quantum computer based on quantum logic would be ideal for simulating quantum-mechanical systems and his ideas have spawned an active area of research in physics. It is also remarkable that quantum mechanics can help in the solution of basic problems of computer science. In 1994, Peter Shor proposed a quantum algorithm that efficiently solves the prime-factorization problem: given a composite integer, find its prime factors. This is an important problem in computer science and it is conjectured, though not proven, that for a classical computer it is computationally difficult to find the prime factors. Shor's algorithm efficiently solves the integer factorization problem and therefore it provides an exponential improvement in speed with respect to any known classical algorithm. It is worth mentioning here that there are cryptographic systems, such as RSA, that are used extensively today and that are based on the conjecture that no efficient algorithms exist for solving the prime factorization problem. Hence, Shor's algorithm, if implemented on a large-scale quantum computer, would break the RSA cryptosystem. Lov Grover demonstrated that quantum mechanics can also be useful for solving the problem of searching for a marked item in an unstructured database. In this case, the gain with respect to classical computation is quadratic.

Another interesting aspect of the quantum computer is that, in principle, it avoids dissipation. Present day classical computers, which are based on irreversible logic operations (gates), are *intrinsically dissipative*. The minimum energy requirements for irreversible computation are set by Landauer's principle: each time a single bit of information is erased, the amount of energy dissipated into the environment is at least  $k_B T \ln 2$ , where  $k_B$  is Boltzmann's constant and  $T$  the temperature of the environment surrounding the computer. Each irreversible classical gate must dissipate at least this amount of energy (in practice, present-day computers dissipate more by orders of magnitude<sup>1</sup>). In contrast, quantum evolution is unitary and thus quantum logic gates must be reversible. Therefore, at least in principle, there is no energy dissipation during a quantum computer run. Obviously this is by no means the full story: to execute a single unitary operation in a real system, a complex hardware with a huge amount of energy dissipation is at present necessary. Furthermore, a fault-tolerant quantum computer would require error correction, with the unavoidably associated energy cost.

---

<sup>1</sup>A rough estimate is obtained by using the best performance per watt (as a measure of power efficiency) of  $1.7 \times 10^{10}$  flops/W from the November 2017 Green500 list. Such a list ranks the 500 most powerful supercomputers in terms of energy efficiency. Assuming  $10^5$  logical operations per flop and  $k_B T \approx 4 \times 10^{-21}$  J at  $T = 300$  K, we obtain an energy dissipation of about  $10^5 k_B T$  per logical operation. It might be interesting to note here that the human brain, which dissipates about 30 W and performs of the order of  $10^{19}$ – $10^{20}$  operations per second, is much more efficient. Indeed for every operation it dissipates about  $10^2$ – $10^3$   $k_B T$ .

It is well known that a small set of elementary logic gates allows the implementation of any complex computation on a classical computer. This is very important: it means that, when one addresses a different problem, it is not necessary to modify the computer hardware as well. Fortunately, the same property holds for a quantum computer. It turns out that, in the quantum circuit model, each unitary transformation acting on a many-qubit system can be decomposed into gates acting on a single qubit and a single gate acting on two qubits, for instance the CNOT gate.

A large number of different proposals to build real quantum computers have been put forward. They range from cold ion traps to superconducting tunnel-junction circuits and spin in semiconductors, to name but a few. Even though in some cases elementary quantum gates have been realized and quantum algorithms with a small number of qubits demonstrated, it is too early to say what type of implementation will be the most suitable to build a scalable piece of quantum hardware. Although for some computational problems the quantum computer is more powerful than the classical computer, still we need at least 50–1000 qubits and from thousands to millions of quantum gates to perform tasks inaccessible to the classical computer (the exact numbers depend, of course, on the specific quantum algorithm).

The technological challenge of realizing a quantum computer is very demanding: we need to be able to control the evolution of a large number of qubits for the time necessary to perform many quantum gates. Decoherence may be considered the ultimate obstacle to the practical realization of a quantum computer. Here the term decoherence denotes the decay of the quantum information stored in a quantum computer, due to the inevitable interaction of the quantum computer with the environment. Such interaction affects the performance of a quantum computer, introducing errors into the computation. Another source of errors that must be taken into account is the presence of imperfections in the quantum-computer hardware. Even though quantum error-correcting codes exist, a necessary requirement for a successful correction procedure is that one can implement many quantum gates inside the decoherence time scale. Here “many” means  $10^4$ – $10^6$ , the exact value depending on the kind of error. It is very hard to fulfill this requirement in complex many-qubit quantum systems.

The following question then arises: is it possible to build a useful quantum computer that could outperform existing classical computers in important computational tasks? And, if so, when? Besides the problem of decoherence, we should also remark on the difficulty of finding new and efficient quantum algorithms. We know that the integer-factoring problem can be solved efficiently on a quantum computer, but we do not know the answer to the following fundamental question: what class of problems could be simulated efficiently on a quantum computer? Quantum computers open up fascinating prospects, but it does not seem likely that they will become a reality with practical applications in a few years. How long might it take to develop the required technology? Even though unexpected technological breakthroughs are, in principle, always possible, one should remember the

enormous effort that was necessary in order to develop the technology of classical computers.

Nevertheless, even the first, modest, demonstrative experiments are remarkable, as they allow for testing the theoretical principles of quantum mechanics. Since quantum mechanics is a particularly counter-intuitive theory, we should at the very least expect that experiments and theoretical studies on quantum computation will provide us with a better understanding of quantum mechanics. Moreover, such research stimulates the control of individual quantum systems (atoms, electrons, photons *etc.*). We stress that this is not a mere laboratory curiosity, but has interesting technological applications. For instance, it is now possible to realize single-ion clocks that are more precise than standard atomic clocks<sup>2</sup>. Another useful byproduct is the generation of random bit sequences on the basis of quantum effects, exploiting the fundamental randomness of quantum mechanics. In a sense we may say that quantum computation rationalizes the efforts of the various experiments that manipulate individual quantum systems. It might be useful to recall here that in 2012 Serge Haroche and David Wineland were awarded the Nobel prize in physics “for ground-breaking experimental methods that enable measuring and manipulation of individual quantum systems”.

A different research line is the implementation of *quantum simulators*, grounded on the seminal Feynman’s idea that a quantum system could be efficiently simulated by using another, controllable, quantum system. Rather than a general-purpose quantum computer, a quantum simulator is a simpler, analog device. The simulation is not performed by decomposing the quantum system’s evolution into elementary quantum gates, but it is done directly by observing the dynamics of the quantum many-body system (the simulator). This basic idea has been developed in several different realizations, including cold atoms in optical lattices, superconducting circuits and integrated photonic platforms. In view of the technological breakthroughs in such context, nowadays one of the most widely pursued approaches is the *quantum annealing*, otherwise named adiabatic quantum computation. Its strategy assumes to encode the solution of a given problem into the ground state of a suitable Hamiltonian. The protocol proceeds in such a way to determine such state by performing an adiabatic connection (if possible) with another Hamiltonian, typically describing a much simpler physical system. A number of different situations that might enable a considerable speedup induced by quantum mechanics have been extensively analyzed in the context of Hamiltonian complexity theory. This stimulated many private companies, such as D-Wave, Google, IBM, Intel, and Microsoft, to initiate a research line, with the ultimate purpose to build up a functioning quantum annealer which outperforms any conceivable classical hardware, thus putting in practice the so-called *quantum advantage*.

---

<sup>2</sup>However, better performances are obtained by the latest atomic clocks with millions of strontium atoms in optical lattices, with a measurement precision of  $5 \times 10^{-19}$  (at JILA, Boulder, Colorado, 2017).

From a broader perspective, quantum information finds applications also in condensed matter physics. Specifically it has been discovered that, for a large variety of strongly correlated quantum many-body systems, including prototype spin systems and Hubbard-like models, the amount of bipartite quantum correlations that may establish in the low-energy physics of such models is severely bounded by an *area-law* behaviour. This concept stands at the basis of the *density-matrix renormalization group* (DMRG), a virtually exact numerical algorithm that has been devised in the Nineties for finding the ground state of one-dimensional quantum lattice systems. Its working mechanism is to build up a portion of the system and then recursively enlarge it, until the desired system size is reached. At every step, the basis of the corresponding Hamiltonian is truncated by suitably exploiting the above mentioned area law, so that the size of the Hilbert space is kept manageable as the physical system grows. Nowadays several extensions of this basic idea have been put in practice, giving rise to the theory of quantum tensor networks. Besides condensed matter, all these concepts may find applications in several other fields, including quantum chemistry, high-energy physics, and cosmology.

Another important research direction concerns the (secure) transmission of information. In this case, quantum mechanics allows us to perform not only faster operations but also operations *inaccessible* to classical means. Entanglement is at the heart of many quantum-information protocols. It is the most spectacular and counter-intuitive manifestation of quantum mechanics, observed in composite quantum systems: it signifies the existence of non-local correlations between measurements performed on well-separated particles. After two classical systems have interacted, they are in well-defined individual states. In contrast, after two quantum particles have interacted, in general, they can no longer be described independently of each other. There will be purely quantum correlations between two such particles, independently of their spatial separation. This is the content of the celebrated EPR paradox, a *Gedanken* experiment proposed by Einstein, Podolsky and Rosen in 1935. These authors showed that quantum theory leads to a contradiction, provided that we accept the two, seemingly natural, principles of realism and locality. The reality principle states that, if we can predict with certainty the value of a physical quantity, then this value has physical reality, independently of our observation. The locality principle states that, if two systems are causally disconnected, the results of any measurement performed on one system cannot influence the result of a measurement performed on the second system. In other words, information cannot travel faster than the speed of light.

In 1964 Bell proved that this point of view (known as called local realism) leads to predictions, Bell's inequalities, that are in contrast with quantum theory. Aspect's experiments (1982), performed with pairs of entangled photons, exhibited an unambiguous violation of a Bell's inequality by tens of standard deviations and an impressive agreement with quantum mechanics. These experiments also showed that it is possible to perform laboratory investigations on the more fundamental,

non-intuitive aspects of quantum theory. More recently, other experiments have come closer to the requirements of the ideal EPR scheme. More generally, thanks to the development and increasing precision of experimental techniques, *Gedanken* experiments of the past become present-day real experiments.

The profound significance of Bell's inequalities and Aspect's experiments lies far beyond that of a mere consistency test of quantum mechanics. These results show that entanglement is a fundamentally new resource, beyond the realm of classical physics, and that it is possible to experimentally manipulate entangled states.

Quantum entanglement is central to many quantum-communication protocols. Of particular importance are *quantum dense coding*, which permits transmission of two bits of classical information through the manipulation of only one of two entangled qubits, and *quantum teleportation*, which allows the transfer of the state of one quantum system to another over an arbitrary distance. In 2017, it was possible to observe the violation of a Bell inequality for two entangled photons, distributed from a satellite to two ground stations (Delingha and Lijiang in China), at a distance of 1203 Km from each other. Ground-to-satellite teleportation experiments based on entangled photons were also realized in 2017, over distances of up to 1400 Km. These experiments pave the way to the realization of a global quantum communication network. Furthermore, from a fundamental viewpoint it is possible to test the nonlocality of quantum mechanics over unprecedented distances. Moreover, it will become possible to investigate the effects of gravitational fields on entanglement, and this might provide useful insights on the link between gravitation and quantum mechanics.

Quantum mechanics also provides a unique contribution to cryptography: it enables two communicating parties to detect whether the transmitted message has been intercepted by an eavesdropper. This is not possible in the realm of classical physics as it is always possible, in principle, to copy classical information without changing the original message. In contrast, in quantum mechanics the measurement process, in general, disturbs the system for fundamental reasons. Put plainly, this is a consequence of the Heisenberg uncertainty principle. Experimental advances in the field of *quantum cryptography* are impressive and quantum-cryptographic protocols have been demonstrated, using optical fibres, over distances up to a few hundred kilometres. However, on the basis of present technology it appears difficult to improve significantly this figure since photon-absorption losses grows exponentially with the length of the fibre. It should be noted that the aforementioned quantum communication with satellites has been precisely designed to realize the quantum key distribution over large distances. In 2017, a kilohertz key rate from a satellite to the ground over a distance of up to 1200 kilometres was achieved. Finally, we note that there are currently companies offering commercial quantum key distribution systems. Hence, quantum cryptography is the first quantum-information protocol to find commercial applications.

To conclude this introduction, let us quote Schrödinger [*Brit. J. Phil. Sci.*, **3**, 233 (1952)]: “*We never experiment with just one electron or atom or (small) molecule. In thought-experiments we sometimes assume that we do; this invariably entails ridiculous consequences . . . we are not experimenting with single particles, any more than we can raise Ichthyosauria in the zoo.*” It is absolutely remarkable that only fifty years later experiments on single electrons, atoms and molecules are routinely performed in laboratories all over the world.

## Chapter 1

# Introduction to classical computation

This chapter introduces the basic concepts of computer science that are necessary for an understanding of quantum computation and information. We discuss the Turing machine, the fundamental model of computation since it formalizes the intuitive notion of an algorithm: if there exists an algorithm to solve a given problem, then this algorithm can be run on a Turing machine. We then introduce the circuit model of computation, which is equivalent to the Turing machine model but is nearer to real computers. In this model, the information is carried by wires and a small set of elementary logical operations (gates) allows implementation of any complex computation. It is important to find the minimum resources (computer memory, time and energy) required to solve a given problem with the best possible algorithm. This is the task of computational complexity, for which we provide a quick glance at the key concepts. We also examine the energy resources necessary to perform computations. Here we discuss the relation between energy and information, which was explained by Landauer and Bennett in their solution of Maxwell's demon paradox. In particular, Landauer's principle sets the minimum energy requirements for irreversible computation. On the other hand, it turns out that it is, in principle, possible to perform any complex computation by means of reversible gates, without energy dissipation. A concrete model of reversible computation, the so-called billiard-ball computer, is briefly discussed. Finally, we address the problem of energy dissipation in a real computation and show how it may be managed using reversible computation techniques.

### 1.1 The Turing machine

An *algorithm* is a set of instructions for solving a given problem. Examples of algorithms are those learnt at primary schools for adding and multiplying two integer numbers. Such algorithms always give the correct result when applied to any pair of integer numbers.

The *Turing machine*, introduced by the mathematician Alan Turing in the 1930's, provides a precise mathematical formulation of the intuitive concept of algorithm. This machine contains the essential elements (memory, control unit and

read/write unit) on which any modern computer is based. Turing's work was stimulated by an intense debate at that time regarding the following question: for which class or classes of problems is it possible to find an algorithm? This debate was motivated by a profound question raised by David Hilbert at the beginning of the twentieth century. Hilbert asked whether or not an algorithm might exist that could, in principle, be used to solve all mathematical problems. Hilbert thought (erroneously, as we shall see in this section) that the answer to his question was positive.

A closely related problem is the following: given a logical system defined by an ensemble of axioms and rules, can all possible propositions be proved, at least in principle, to be either a consequence of the axioms and rules or not? (Shortly: to be or not to be a theorem.) At the beginning of twentieth century it was widely believed that the answer to this question was positive. (Of course, the question does not address the problem that, in practice, it may be extremely difficult to prove whether a proposition is or not a theorem.) Contrary to this belief, in the 1930's Kurt Gödel proved a theorem stating that there exist propositions of any given logical system (barring the most trivial ones, like Pressburger arithmetic) that are *undecidable*, meaning that they can neither be proved nor disproved using axioms and rules inside the same logical system.<sup>1</sup> This does not exclude that we can enlarge the system, introducing between the axioms the undecidable proposition or its opposite as a new axiom and thus decide whether a given proposition is or not a theorem. However, it will also be possible to find undecidable propositions in this new system. Thus, it turns out that logical systems are intrinsically *incomplete*. Notice that Gödel's theorem also sets limits on the possibilities of a computer: it cannot answer all questions on arithmetics.

The main elements of a Turing machine are illustrated in Fig. 1.1. The general idea is that the machine performs a computation as a "human computer" would. Such a human computer is capable of storing only a limited amount of information in his brain, but has at his disposal an (ideally) unlimited amount of paper for reading and writing operations. Likewise, the Turing machine contains the following three main elements:

- (1) A *tape*, which is infinite and divided into cells. Each cell holds only one letter  $a_i$  from a finite alphabet  $\{a_1, a_2, \dots, a_k\}$  or is blank. Except for a finite number of cells, all the other cells are blank.
- (2) A *control unit*, which has a finite number of states,  $\{s_1, s_2, \dots, s_l, H\}$ , where  $H$  is a special state, known as the halting state: if the state of the control unit becomes  $H$ , then the computation terminates.
- (3) A *read/write head*, which addresses a single cell of the tape. It reads and (over)writes or erases a letter in this cell, after which, the head moves one cell to the left or to the right or the computation halts.

---

<sup>1</sup>In Pressburger arithmetic, which contains only the addition and equality but not multiplication, any proposition is decidable.

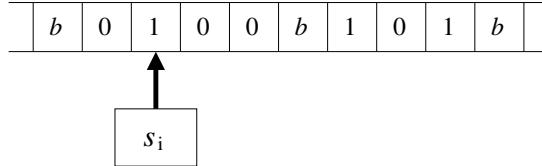


Fig. 1.1 Schematic drawing of a Turing machine. The symbol  $b$  denotes a blank cell.

The working of a Turing machine is governed by a *program*, which is simply a finite set of instructions. Each instruction governs one step of the Turing machine and induces the following sequence of operations:

- (i) the transition of the control unit from the state  $s$  to the state  $\bar{s}$ ,
- (ii) the transition of the cell addressed by the read/write head from the letter  $a$  to the letter  $\bar{a}$ ,
- (iii) the displacement of the read/write head one cell left or right or the computation halts.

Therefore, an instruction in the Turing machine is defined by three functions  $f_S$ ,  $f_A$  and  $f_D$ , defined as follows:

$$\bar{s} = f_S(s, a), \quad \bar{a} = f_A(s, a), \quad d = f_D(s, a), \quad (1.1)$$

where  $d$  indicates the displacement of the head to the left ( $d = l$ ) or to the right ( $d = r$ ). In short, the functions  $f_S$ ,  $f_A$  and  $f_D$  define the mapping

$$(s, a) \rightarrow (\bar{s}, \bar{a}, d). \quad (1.2)$$

### 1.1.1 Addition on a Turing machine

Let us now describe a concrete example: a Turing machine performing the addition of two integers. For the sake of simplicity, we write the integer numbers in the unary representation: an integer  $N$  is written as a sequence of  $N$  1's, that is,  $1 = 1$ ,  $2 = 11$ ,  $3 = 111$ ,  $4 = 1111$  and so on. As an example, we compute the sum  $2 + 3$ . Our Turing machine needs five internal states  $\{s_1, s_2, s_3, s_4, H\}$  and a unary alphabet, namely, the single letter 1. We denote by  $b$  the blank cells of the tape. The initial condition of the machine is shown in Fig. 1.2: the initial state is  $s_1$ , the head points to a well-defined cell and the numbers to be added,  $2 = 11$  and  $3 = 111$ , are written on the tape, separated by a blank.

The program for computing the sum of two integer numbers is shown in Table 1.1. The program has a total number of six lines. The internal state  $s$  of the machine and the letter  $a$  being read on the tape determine which program line is executed. The last three columns of Table 1.1 denote the new state  $\bar{s}$ , the letter  $\bar{a}$  overwritten on the tape and the left/right direction ( $d = l$  or  $d = r$ ) of the read/write head motion. Note that in the fourth line of the program we write  $d = -$  since the machine halts and the head moves no further. It is easy to check that, if

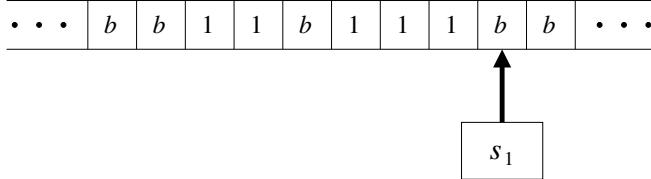


Fig. 1.2 Initial conditions of a Turing machine for computing the sum  $2 + 3$ .

Table 1.1 The algorithm for computing the sum of two integers on a Turing machine.

$s$	$a$	$\bar{s}$	$\bar{a}$	$d$
$s_1$	$b$	$s_2$	$b$	$l$
$s_2$	$b$	$s_3$	$b$	$l$
$s_2$	$1$	$s_2$	$1$	$l$
$s_3$	$b$	$H$	$b$	$-$
$s_3$	$1$	$s_4$	$b$	$r$
$s_4$	$b$	$s_2$	$1$	$l$

we start from the initial conditions of Fig. 1.2 and run the program of Table 1.1, the machine halts in the configuration depicted in Fig. 1.3 and we can read on the tape the result of the sum,  $2 + 3 = 5$ . It is also easy to convince ourselves that the same program can compute the sum of two arbitrary integers  $m$  and  $n$ , provided that the initial conditions are set as in Fig. 1.4.

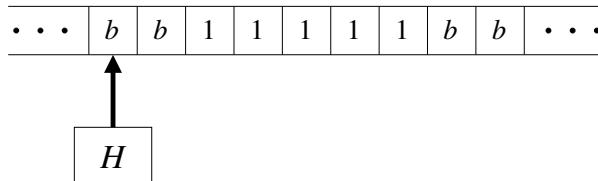


Fig. 1.3 A Turing machine after computation of the sum  $2 + 3$ . The machine started from the initial conditions of Fig. 1.2 and implemented the program of Table 1.1.

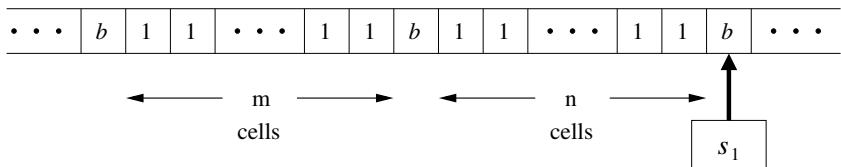


Fig. 1.4 The initial conditions of a Turing machine for computing the sum  $m + n$  of two generic integers.

### 1.1.2 The Church–Turing thesis

It turns out that Turing machines are capable of solving very complex problems. As far as we know, they can be used to simulate any operation carried out on a modern computer. It is useful to note that the only difference is the execution time. If there exists an algorithm to compute a function, then the computation can be performed by a Turing machine. This idea was formalized independently by Church and Turing:

**The Church–Turing thesis:** *The class of all functions computable by a Turing machine is equivalent to the class of all functions computable by means of an algorithm.*

This statement provides a rigorous mathematical definition of the intuitive concept of “function computable by an algorithm”: a function is computable if and only if it can be computed by a Turing machine. The thesis, formulated in 1936, has never been disproved since we do not know of any algorithm that computes a function not computable by a Turing machine. Indeed, much evidence has been gathered in favour of the Church–Turing thesis.

It is worthwhile to note that in all the previous discussion it is essential to assume a finite (but unbounded) number of steps in the computation. This condition excludes the so-called “accelerated” Turing machines, capable of running an infinite number of steps in a finite time, or other exotic machines dubbed *hypercomputers* (the accelerated Turing machine has the first cycle of the clock of, say, 1 second, the second of  $1/2$  second, the third of  $1/4$  second, and so on).

One may think that quantum mechanics gives access to hypercomputation because the amplitudes of the states of a system are given by complex numbers, instead of the integers (0 and 1) of standard classical computers. As one may think that analog computers (working with real numbers) are more powerful than the digital ones (working with integers), it is possible to suspect that quantum computers may be more powerful than classical computers. However, this argument does not take into account that analog computers suffer from problems of noise and unpredictability. To note realistic physical limits to hypercomputation, it is useful to consider the proposal of Kieu of using the Hilbert space of a quantum field (harmonic oscillators) to solve the Hilbert’s tenth problem (existence of an algorithm to decide the solvability of all Diophantine equations). Note that such problem is equivalent to the Turing halting problem, which is known to be mathematically undecidable (see Sec. 1.1.5 below). This proposal is not realistic as for sufficiently high excitation of the quantum field, there is a finite probability to create undesired new particles, destroying the algorithm. Other subtle flaws in the use of the adiabatic theorem (this theorem is discussed in Chap. 3) in the Kieu’s algorithm are reported in the literature (see the guide to the bibliography at the end of this chapter). Similar problems may be found in other hypercomputation proposals.

Intuition tells us that an accelerated Turing machine is impossible. However, for such a fundamental problem the intuition is not enough: it is necessary to give precise physical limits to such a hypothesized impossibility. It is impossible to give clear restrictions to the accelerated Turing machine in the framework of the non-relativistic classical mechanics. Indeed, starting from customary machines it is in principle possible to build smaller and faster machines, and from these, build still smaller and faster ones, *ad infinitum*. So there

is no fundamental limit to the miniaturization of mechanical computer elements in classical mechanics.

But quantum mechanics, special relativity, the standard model of field theory and the general relativity give precise limits to the building of accelerated Turing machines and to other systems dubbed supersystems. It is possible to give a (partial) list of the physical limits as presently understood:

- (i) Clock frequency: at present the practical clock frequency limit for classical and quantum computers are roughly related to the atomic Bohr frequency. So, the clock timing cannot be shorter than few tens femtoseconds (one femtosecond is  $10^{-15}$  seconds) (see the discussion on the speed of gates in Chap. 3).
- (ii) Special relativity gives precise limits to the speed of signals: it is not possible to reach speeds greater than the speed of light. This limit is very strong: in the laboratory transistors have been realized, working at the terahertz level (1 terahertz is equal to  $10^{12}$  hertz), and at this frequency the wavelength of the radiation in vacuum is 0.03 cm, posing synchronization problems of data transferred between the different units of a computer.
- (iii) With respect to general relativity, we note that a rough time scale at which quantum gravitational effects are likely to become important is set by the Planck time  $t_P \equiv \sqrt{\hbar G/c^5} \approx 5.39 \times 10^{-44}$  s, where  $\hbar = h/2\pi \approx 1.05 \times 10^{-34}$  m<sup>2</sup>Kg/s is the reduced Planck constant,  $G \approx 6.67 \times 10^{-11}$  m<sup>3</sup>Kg<sup>-1</sup>s<sup>-2</sup> is the gravitational constant, and  $c \approx 2.998 \times 10^8$  m/s is the speed of light in vacuum. In quantum general relativity (quantum gravity), it is believed that time intervals smaller than the Planck time are meaningless.

So hypercomputers are extremely improbable on the basis of quantum mechanics and relativity theory, rather than of non-relativistic classical mechanics. The above discussion makes it clear that the limits of computation are physical, not logical.

Finally, it might be interesting to mention that an important problem not solved is if the human brain works by means of algorithms. The problem is strictly related to the free will problem that after millennia of debate is largely unsolved. As this is a book on quantum computation, it is natural to ask if the quantum computation paradigm may help to solve this problem. At present it is better to affirm that this problem is unanswered.<sup>2</sup>

### **1.1.3 The universal Turing machine**

The universal Turing machine  $U$  is a single machine that encompasses all Turing machines; that is, it is capable of computing any algorithm. A Turing machine  $T$  running a given program, on the basis of the input  $x$  written on tape, produces some output  $T(x)$  and then halts. The universal Turing machine can simulate any Turing machine  $T$ , provided that on the tape of the universal Turing machine we specify the description of the machine  $T$ . It can be shown that an integer number  $n_T$  may be uniquely associated with the corresponding machine  $T$  (it is sufficient to apply any recipe, such as the Gödel numbering, that converts reversibly a list of instructions into a unique natural number). This number is known as the

---

<sup>2</sup>For a proposal of an experiment to test Bell's inequalities in which humans are used to decide the settings, see Hardy (2017) (the meaning of Bell's inequalities will be discussed in Sec. 2.5).

Turing number associated with the machine. Therefore, if we give the description  $n_T$  of  $T$  and  $x$  on input, the universal Turing machine  $U$  produces the output  $U(n_T, x) = T(x)$ . It is important to stress that in the universal Turing machine the (finite) set of internal states  $\{s_i\}$  and the program are fixed once and for all. Thus, we can run any computation by simply changing the initial state of the tape.

#### 1.1.4 The probabilistic Turing machine

A probabilistic Turing machine is characterized by the fact that the mapping  $(s, a) \rightarrow (\bar{s}, \bar{a}, d)$  is probabilistic. This means that there exist *coin-tossing states* in which the machine tosses a coin to decide the output. The coin lands heads up with probability  $p$  and tails up with probability  $1 - p$ . In the first case the new internal state of the machine is given by  $\bar{s} = \bar{s}_h$  while in the latter we have  $\bar{s} = \bar{s}_t$ . A probabilistic Turing machine may be more powerful than a deterministic Turing machine, in that it can solve many computational problems faster. If a probabilistic Turing machine gives the result of some application, failing with an error of less than, say,  $\frac{1}{3}$ , then, by running the program a sufficient number of times, it is possible to lower the error probability below any preset threshold. An example illustrating the usefulness of probabilistic algorithms will be discussed in Sec. 1.3. However, we should note that the probabilistic Turing machine does not enlarge the class of functions computable by a deterministic Turing machine. Indeed, a deterministic Turing machine can always simulate a probabilistic Turing machine by exploring, one after another, all possible paths corresponding to different values of the coin tosses.

#### 1.1.5 \* The halting problem

Let us now consider the following problem: will some given Turing machine  $T$  eventually halt for some given input  $x$ ? The question is quite natural: the machine could either end up in the internal state  $H$  and stop after some finite time or loop infinitely without ever reaching the state  $H$ . Turing demonstrated that there exists no algorithm capable of solving this problem, known as the halting problem. An instance of this problem is the following: will a given machine  $T$  attain the halt state  $H$  after input of its own Turing number  $n_T$ ? In other words, is there an algorithm (or Turing machine)  $A$  whose output  $A(n_T)$  tells us whether or not some Turing machine  $T$  eventually halts on input of  $n_T$ ?

Let us assume that such an algorithm exists. In other words, if machine  $T$  halts for input  $n_T$ , then for the same input  $A$  writes “yes” and halts, otherwise  $A$  writes “no” and halts. In the following, we shall prove that such a machine  $A$  cannot exist. Let us consider another machine  $B$ , defined as follows: if  $A$  writes “yes” for some input  $n_T$  then  $B$  does not halt, if instead  $A$  writes “no” then  $B$  does halt. If  $A$  exists, then  $B$  exists as well. Therefore, for any  $n_T$ ,  $B(n_T)$  halts if and only if  $T(n_T)$  does not halt. We now consider the case in which the input of the machine

$B$  is its own Turing number  $n_B$ . Therefore,  $B(n_B)$  halts if and only if  $B(n_B)$  does not halt. This is a contradiction and therefore the machine  $A$  cannot exist.

To understand in an intuitive (but not rigorous) way the logical basis of this proof by *reductio ad absurdum*, consider the following paradoxical statement: “This sentence is false.” While it does not violate any grammatical law, that is, its construction is perfectly legitimate, there is no possible answer to the question “Is the statement true or not?” The above problem is equivalent to asking a computer to provide just such answer. However, it should also be remarked that the *practical* stumbling block to such an algorithm is clearly the difficulty in demonstrating the infinite-loop condition without waiting an infinite amount of time.

It might be interesting to note that with a similar technique it can be proved that a perfect antivirus does not exist. Let us assume that a decision procedure  $D$  exists, which decides if  $V$  is a virus, that is, a program that can infect other programs by modifying them. Hence, the virus  $V$  is detected by  $D$ . We can now consider a new program  $P$ , which invokes the decision procedure  $D$ . If  $D$  decides  $P$  is not a virus, then  $P$  will infect another program. If, on the other hand,  $D$  decides  $P$  is a virus, then  $P$  will not infect other programs. This is a contradiction and therefore the desired decision procedure  $D$  does not exist.

Finally, we note that the halting theorem is at the basis of many impossibility theorems, obtained reducing a given theorem to the halting problem. For instance, the impossibility to solve the tenth Hilbert problem on the solution of Diophantine equations and the gap problem in adiabatic evolutions (see Chap. 3).

## 1.2 The circuit model of computation

In terms of computational power, the circuit model of computation is equivalent to the Turing machine model discussed in the previous section but is nearer to a real computer. Let us first introduce the *bit*, the elementary unit of classical information. A bit is defined as a two-valued or binary variable, whose values are typically written as the binary digits 0 and 1. A circuit is made of *wires* and *gates*; each wire carries one bit of information since it may take the value 0 or 1. As we shall see below, the gates perform logic operations on these bits. The classical computer is a *digital* device since the information enters the computer as a binary sequence of 0's and 1's and the output of any computation is again a binary sequence of 0's and 1's. For instance, an integer number  $N < 2^n$  is stored as follows:

$$N = \sum_{k=0}^{n-1} a_k 2^k, \quad (1.3)$$

where the value of each binary digit  $a_k$  may be equal to 0 or to 1. We may write equivalently

$$N = a_{n-1} a_{n-2} \dots a_1 a_0. \quad (1.4)$$

For instance, we have  $3 = 11$ ,  $4 = 100$ ,  $5 = 101$  and  $49 = 110001$ . We may also write binary fractions, using the notation  $\frac{1}{2} = 0.1$ ,  $\frac{1}{4} = 0.01$ ,  $\frac{1}{8} = 0.001$  and so on. Let

we write the binary codes of a few non-integer numbers:  $5.5 = 101.1$ ,  $5.25 = 101.01$  and  $5.125 = 101.001$ . It is clear that any real number may be approximated to any desired accuracy by a binary fraction.

The advantage of the binary notation is that binary numbers are well suited to being stored in electrical devices since only two possible values need be set: computers use high and low voltage values or switches with only two positions (*on* and *off*) to load one bit of information. For instance, in Fig. 1.5 we show the sequence of voltages required to load the integer number  $N = 49$ .

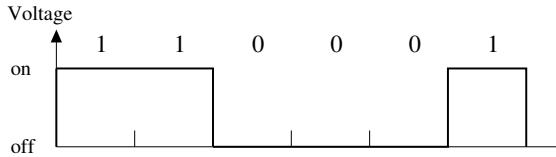


Fig. 1.5 The sequence of voltages representing the integer  $N = 49$ .

### 1.2.1 Binary arithmetics

The arithmetical rules also turn out to be much simpler in the binary representation. As an example, in Table 1.2 we show the binary addition table, where  $s = a \oplus b$  is the addition, modulo two, of the two bits  $a$  and  $b$  while  $c$  is the carry over.

Table 1.2 The binary addition table.

$a$	$b$	$s$	$c$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

The following examples should help clarify the procedure for computing the sum and product of two numbers written in their binary representations. Decimal addition and multiplication are also shown to provide a more familiar comparison.

	BINARY	DECIMAL
ADDITION	$  \begin{array}{r}  1 & 1 & 1 & 0 & 1 \\  1 & 0 & 1 & 0 & 1 \\  \hline  1 & 1 & 0 & 0 & 1 & 0  \end{array}  $	$  \begin{array}{r}  2 & 9 \\  2 & 1 \\  \hline  5 & 0  \end{array}  $
MULTIPLICATION	$  \begin{array}{r}  1 & 1 & 1 & 0 & 1 \\  1 & 0 & 1 & 0 & 1 \\  \hline  1 & 1 & 1 & 0 & 1 \\  1 & 1 & 1 & 0 & 1 \\  \hline  1 & 0 & 0 & 1 & 1 & 0 & 1  \end{array}  $	$  \begin{array}{r}  2 & 9 \\  2 & 1 \\  \hline  2 & 9 \\  5 & 8 \\  \hline  6 & 0 & 9  \end{array}  $

### 1.2.2 Elementary logic gates

In any computation, we must provide an  $n$ -bit input to recover an  $l$ -bit output. Namely, we must compute a logical function of the form

$$f : \{0, 1\}^n \rightarrow \{0, 1\}^l. \quad (1.5)$$

As we shall show later in this section, the evaluation of any such function may be decomposed into a sequence of elementary logical operations. First of all, we introduce a few logic gates that are useful for computation.

Figure 1.6 (left) shows a trivial one-bit gate, the identity gate: the value of the output bit is simply equal to the value of the input bit. The simplest non-trivial gate is the NOT gate, which acts on a single bit and flips its value: if the input bit is 0, the output bit is set to 1, and vice versa. In binary arithmetics,

$$\bar{a} = 1 - a, \quad (1.6)$$

where  $\bar{a}$  denotes NOT  $a$ . The circuit representation and the truth table for the NOT gate are shown in Fig. 1.6 (right).



Fig. 1.6 The truth table and circuit representation for the identity (left) and for the NOT gate (right).

We next introduce a list of two-bit logic gates useful for computation. These gates have two input bits and one output bit and are therefore binary functions  $f : \{0, 1\}^2 \rightarrow \{0, 1\}$ .

- (i) The AND ( $\wedge$ ) gate (see the left side of Fig. 1.7): produces output 1 if and only if both input bits are set to 1. In binary arithmetics,

$$a \wedge b = ab. \quad (1.7)$$

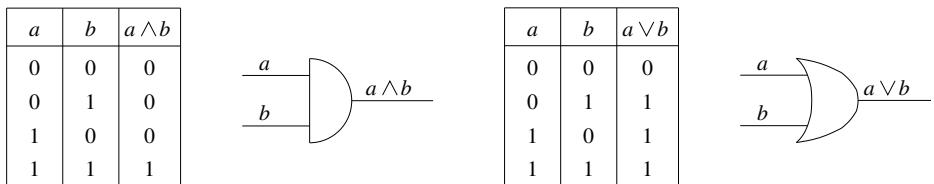


Fig. 1.7 The truth table and circuit representation for the AND (left) and the OR gate (right).

- (ii) The OR ( $\vee$ ) gate (see the right side Fig. 1.7): produces output 1 if and only if at least one of the input bits is set to 1. In binary arithmetics,

$$a \vee b = a + b - ab. \quad (1.8)$$

- (iii) The XOR ( $\oplus$ ) gate (see Fig. 1.8): produces output 1 if only one of the input bits is set to 1, otherwise the output is 0. The XOR (also known as the exclusive OR) gate outputs the sum, modulo 2, of the inputs:

$$a \oplus b = a + b \pmod{2}. \quad (1.9)$$

$a$	$b$	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

Fig. 1.8 The truth table and circuit representation for the XOR gate.

- (iv) The NAND ( $\uparrow$ ) gate (see the left side of Fig. 1.9): produces output zero if and only if both inputs are set to one. It is obtained by the application of a NOT gate to the output of an AND gate:

$$a \uparrow b = \overline{a \wedge b} = \overline{ab} = 1 - ab. \quad (1.10)$$

$a$	$b$	$a \uparrow b$
0	0	1
0	1	1
1	0	1
1	1	0

$a$	$b$	$a \downarrow b$
0	0	1
0	1	0
1	0	0
1	1	0

Fig. 1.9 The truth table and circuit representation for the NAND (left) and the NOR (right) gate.

- (v) The NOR ( $\downarrow$ ) gate (see the right side Fig. 1.9): produces output 1 if and only if both inputs are set to zero. It is obtained by the application of a NOT gate to the output of an OR gate:

$$a \downarrow b = \overline{a \vee b} = \overline{a + b - ab} = 1 - a - b + ab. \quad (1.11)$$

Other important gates are the FANOUT (also known as COPY) gate, which takes one bit into two bits:

$$\text{COPY} : a \rightarrow (a, a), \quad (1.12)$$

and the CROSSOVER (or SWAP), which interchanges the values of two bits:

$$\text{Crossover} : (a, b) \rightarrow (b, a). \quad (1.13)$$

The circuit representations for these two gates are shown in Fig. 1.10.

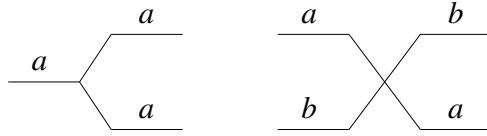


Fig. 1.10 Circuit representations for the FANOUT (COPY) gate (left) and the CROSSOVER (SWAP) gate (right).

The elementary gates described above may be put together to implement any complex computation. As an example, we shall construct a circuit to perform the summation  $s$  of two integer numbers  $a$  and  $b$  using a sequence of AND, OR, XOR and FANOUT gates. We can compute the sum  $s = a + b$  bit-by-bit. Given the binary representations  $a = (a_n, a_{n-1}, \dots, a_1, a_0)$  and  $b = (b_n, b_{n-1}, \dots, b_1, b_0)$ , the  $i$ -th bit of the sum is

$$s_i = a_i \oplus b_i \oplus c_i \pmod{2}, \quad (1.14)$$

where  $c_i$  is the carry over from the sum  $a_{i-1} \oplus b_{i-1} \oplus c_{i-1}$ . We denote  $c_{i+1}$  the carry over of the sum (1.14): it is set to 1 if two or more of the input bits  $a_i$ ,  $b_i$  and  $c_i$  are 1 and 0 otherwise. It is easy to check that the circuit of Fig. 1.11 takes as input  $a_i$ ,  $b_i$  and  $c_i$  and outputs  $s_i$  and  $c_{i+1}$ .

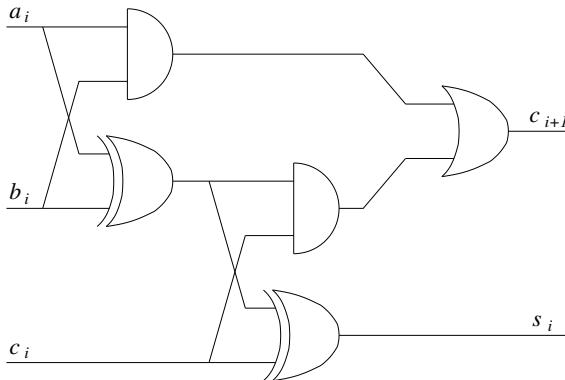


Fig. 1.11 A circuit for computing the sum  $s_i = a_i \oplus b_i \oplus c_i$  and the carry  $c_{i+1}$ . The bifurcating wires are achieved by means of FANOUT gates.

We remark that the elementary gates introduced above are not all independent. For instance, AND, OR and NOT are related by De Morgan's identities:

$$\overline{a \wedge b} = \overline{a} \vee \overline{b}, \quad \overline{a \vee b} = \overline{a} \wedge \overline{b}. \quad (1.15)$$

It is also easy to check that the XOR gate can be constructed by means of the AND, OR and NOT gates as follows:

$$a \text{ XOR } b = (a \text{ OR } b) \text{ AND } ((\text{NOT } a) \text{ OR } (\text{NOT } b)). \quad (1.16)$$

### 1.2.3 Universal classical computation

**Universal gates:** Any function

$$f : \{0,1\}^n \rightarrow \{0,1\}^m \quad (1.17)$$

can be constructed from the elementary gates AND, OR, NOT and FANOUT. Therefore, we say that these constitute a universal set of gates for classical computation.

**Proof.** The  $m$ -bit function (1.17) is equivalent to  $m$  one-bit (Boolean) functions

$$f_i : \{0,1\}^n \rightarrow \{0,1\}, \quad (i = 1, 2, \dots, m), \quad (1.18)$$

where  $f = (f_1, f_2, \dots, f_m)$ . One way to compute these Boolean functions  $f_i(a)$ ,  $a = (a_{n-1}, a_{n-2}, \dots, a_1, a_0)$ , is to consider its *minterms*  $f_i^{(l)}(a)$ , defined, for each  $a^{(l)}$  such that  $f_i(a^{(l)}) = 1$ , as

$$f_i^{(l)}(a) = \begin{cases} 1 & \text{if } a = a^{(l)}, \\ 0 & \text{otherwise.} \end{cases} \quad (1.19)$$

Then the function  $f_i(a)$  reads as follows:

$$f_i(a) = f_i^{(1)}(a) \vee f_i^{(2)}(a) \vee \dots \vee f_i^{(k)}(a), \quad (1.20)$$

where  $f_i(a)$  is the logical OR of all  $k$  minterms (with  $0 \leq k \leq 2^n - 1$ ). It is therefore sufficient to compute the minterms and to perform the OR gates in order to obtain  $f_i(a)$ . We note that the decomposition (1.20) also requires the implementation of FANOUT gates. We need  $k$  copies of the input  $a$ , since each minterm must act on it.

The evaluation of  $f_i^{(l)}$  may be performed as follows. If, for instance,  $a^{(l)} = 110100\dots001$ , we have

$$f_i^{(l)}(a) = a_{n-1} \wedge a_{n-2} \wedge \bar{a}_{n-3} \wedge a_{n-4} \wedge \bar{a}_{n-5} \wedge \bar{a}_{n-6} \wedge \dots \wedge \bar{a}_2 \wedge \bar{a}_1 \wedge a_0. \quad (1.21)$$

Thus,  $f_i^{(l)}(a) = 1$  if and only if  $a = a^{(l)}$ . This completes our proof: we have constructed a generic function  $f(a)$  from the elementary logic gates AND, OR, NOT and FANOUT.  $\square$

As an illustration of the above procedure, we consider the Boolean function  $f(a)$ , where  $a = (a_2, a_1, a_0)$ , defined as follows:  $f(a) = 1$  if  $a = a^{(1)} = 1$  ( $a_2 = 0, a_1 = 0, a_0 = 1$ ) or if  $a = a^{(2)} = 3$  ( $a_2 = 0, a_1 = 1, a_0 = 1$ ) or if  $a = a^{(3)} = 6$  ( $a_2 = 1, a_1 = 1, a_0 = 0$ ) and  $f(a) = 0$  otherwise. The minterms of  $f(a)$  are  $f^{(1)}(a)$ ,  $f^{(2)}(a)$  and  $f^{(3)}(a)$ , which are equal to one if and only if  $a = a^{(1)}$ ,  $a = a^{(2)}$  and  $a = a^{(3)}$ , respectively. We have  $f(a) = f^{(1)}(a) \vee f^{(2)}(a) \vee f^{(3)}(a)$ , where  $f^{(1)}(a) = \bar{a}_2 \wedge \bar{a}_1 \wedge a_0$ ,  $f^{(2)}(a) = \bar{a}_2 \wedge a_1 \wedge a_0$  and  $f^{(3)}(a) = a_2 \wedge a_1 \wedge \bar{a}_0$ .

Actually, it is even possible to reduce the number of elementary operations. It turns out, for example, that NAND and FANOUT are a smaller universal set. Indeed, we have already seen that OR can be obtained from NOT and AND by means of De Morgan's identities. It is also easy to obtain NOT from NAND and FANOUT:

$$a \uparrow a = \overline{a \wedge a} = 1 - a^2 = 1 - a = \bar{a}. \quad (1.22)$$

### Exercise 1.1 Construct AND and OR from NAND and FANOUT.

In computers the NAND gate is usually implemented via transistors, as shown in Fig. 1.12. A bit is set to 1 if the voltage is positive and to 0 if the voltage is zero. It is easy to verify that the output is  $a$  NAND  $b$ . Indeed, the current flows through the transistors if and only if both inputs have positive voltage ( $a = b = 1$ ). In this case, the output has zero voltage. If at least one of the inputs has zero voltage, there is no current flow and therefore the output has positive voltage.

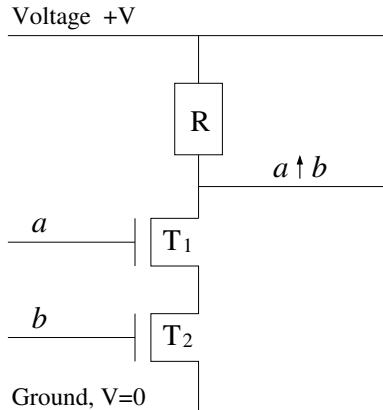


Fig. 1.12 The electrical circuit for a NAND gate.  $R$  denotes a resistor,  $T_1$  and  $T_2$  two transistors.

### 1.3 Computational complexity

To solve any problem, a certain amount of resources is necessary. For instance, to run an algorithm on a computer we need *space* (that is, memory), *time* and *energy*. Computational complexity is the study of the resources required to solve computational problems. Sometimes it is obvious if one problem is easier to solve than another. For instance, we know that it is easier to add two numbers than to multiply them. In other cases, it may be very difficult to evaluate the complexity of a problem. This is a particularly important objective, which affects many fields, from computer science to mathematics, physics, biology, medicine, economics and even the social sciences. A crucial task of computational complexity is to find the *minimum* resources needed to solve a given problem with the best possible algorithm. Hereafter we only deal with problems admitting discrete variables.

Let us start from a very simple example. As noted above, intuitively it is easier to add two numbers than to multiply them. This statement is based on two algorithms learnt at primary school: the addition of two integers requires a number of steps that grows linearly with the total number  $n$  of bits in the input numbers, that is, the time necessary to execute the algorithm is  $t_a = \alpha n$ . The number of steps required to compute the multiplication of these two numbers is instead proportional to the square of  $n$ :  $t_m = \beta n^2$ . Therefore, one might be tempted to conclude that

multiplication is more complex than addition. However, such a conclusion would be based on particular algorithms for computing addition and multiplication: those learnt at primary school. Could different algorithms lead us to different conclusions? It is clear that the addition of two numbers cannot be performed in a number of steps smaller than  $n$ : we must at least read the two  $n$ -digit input numbers. Therefore, we may conclude that the complexity of addition is  $O(n)$  (given two functions  $f(n)$  and  $g(n)$ , we say that  $f = O(g)$  if, for  $n \rightarrow \infty$ ,  $c_1 \leq |f(n)/g(n)| \leq c_2$ , with  $0 < c_1 \leq c_2 < \infty$ ). On the other hand, in 1971 Schönhage and Strassen discovered an algorithm, based on the fast Fourier transform, that requires  $O(n \log n \log \log n)$  steps to carry out the multiplication of two  $n$ -digit numbers on a Turing machine (see Schönhage and Strassen, 1971). This bound has been overcome by Fürer, who found a novel method with a slightly lower asymptotic complexity (see Fürer, 2007). However the advantage of the latter algorithm for realistic values of  $n$  is tiny. Is there a better algorithm to compute multiplication? If not, we should conclude that the complexity of multiplication is those of Fürer's algorithm and that addition is easier than multiplication. However, we cannot exclude the existence of better algorithms for computing multiplication.

### 1.3.1 Tractable vs. intractable problems

To be more quantitative, let us introduce a distinction between problems that can be solved using *polynomial* versus *exponential* resources. More precisely, if  $n$  denotes the *input size*, that is, the number of bits required to specify the input, we may divide solvable problems into two categories:

- (1) Problems that can be solved using resources that are bounded by a polynomial in  $n$ . We say that these problems can be solved *efficiently*, or that they are *easy*, *tractable* or *feasible*. Addition and multiplication belong to this class.
- (2) Problems requiring resources that are *superpolynomial* (*i.e.*, which grow faster than any polynomial in  $n$ ). These problems are considered as *difficult*, *intractable* or *unfeasible*. For instance, it is believed (though not proven) that the problem of finding the prime factors of an integer number is in this class. That is, the best known algorithm for solving this problem is superpolynomial in the input size  $n$ , but we cannot exclude that a polynomial algorithm might exist.

#### *Comments*

- (i) An example may help us clarify the difficulty of superpolynomial problems: the best known algorithm for the factorization of an integer  $N$ , the number field sieve, requires  $\exp(O(n^{1/3}(\log n)^{2/3}))$  operations, where  $n = \log N$  is the input size. The largest factored semiprime number (*i.e.*, a number which is the product of two prime numbers) is 232 decimal digits long. This goal was achieved at the end of 2009, by a collaboration of several research institutions with many hundreds of machines in almost two years, taking the equivalent of about 1500

years of computing on a single-core 2.2 GHz AMD Opteron processor with 2 GB RAM (see Kleinjung *et al.*, 2010). Therefore, we may conclude that the problem is in practice impossible to solve with existing algorithms and any conceivable technological progress.

- (ii) It is clear that also a polynomial algorithm scaling as  $n^\alpha$ , with  $\alpha \gg 1$ , say  $\alpha = 1000$ , can hardly be regarded as easy. However, it is in practice very unusual to encounter useful algorithms with  $\alpha \gg 1$ . In addition, there is a more fundamental reason to base the theory of computational complexity on the distinction between polynomial and super-polynomial algorithms. Indeed, according to the strong Church–Turing thesis, this classification is *robust* when the model of computation is changed.

**The strong Church–Turing thesis:** *A probabilistic Turing machine can simulate any model of computation with at most a polynomial increase in the number of elementary operations required.*

This thesis states that, if a problem cannot be solved with polynomial resources on a probabilistic Turing machine, it has no efficient solution on any other machine. Any model of computation is at best polynomially equivalent to the probabilistic Turing machine model.

In this connection it is interesting to note that quantum computers challenge the strong Church–Turing thesis. Indeed, as will be shown in Chap. 3, there exists an algorithm, discovered by Peter Shor, that solves the integer factorization problem on a quantum computer. As discussed above, we do not know of any algorithm that solves this problem polynomially on a classical computer. And indeed, if such an algorithm does not exist, then we should conclude that the quantum model of computation is more powerful than the probabilistic Turing machine model and the strong Church–Turing thesis should be rejected.

Before giving a more formal definition of the most important complexity classes together with a map of complexity (Sec. 1.3.2), we will discuss few paradigmatic problems which significantly contributed, during the last centuries, to the advance of computational complexity theory.

### 1.3.1.1 \* Optimization problems

Optimization problems consist in searching the optimum solution out of all the feasible ones, which are typically in an exponential number with the problem size  $n$ . This task can be formally expressed in terms of finding an integer, a permutation, or a graph from a finite (large) set of choices.

#### Minimum spanning tree

Suppose to have a company with offices in  $n$  cities, and you want to connect them with phone lines. The phone company charges different amounts of money to connect different pairs of cities. Your task is to build up a set of lines connecting all the

offices with a minimum total cost. Mathematically we can formalize this problem by defining a weighted graph  $\mathcal{G} = (V, E)$ , where each vertex  $v \in V$  is represented by a city, and an edge  $e \in E$  is a phone line with a given “weight”; that is, there is a function  $w : E \rightarrow \mathbb{N}$  that assigns an integer  $w(e)$  to each edge  $e$ . The task is to find a subgraph  $\mathcal{T} \subseteq \mathcal{G}$  connecting all the vertices of  $\mathcal{G}$  with minimum total weight; that is, the minimum spanning tree (MST). Of course  $\mathcal{T}$  should not contain closed loops, since it is always possible to remove an edge from a loop by keeping all the vertices connected and reducing the cost (a graph without loops is named a tree).

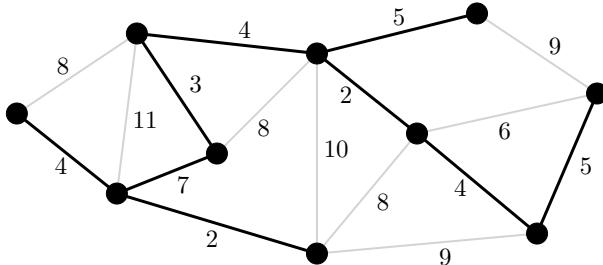


Fig. 1.13 A weighted graph and its minimum spanning tree (dark edges).

Unfortunately the number of all possible trees with  $n$  vertices is extremely large. From Cayley’s formula, we know that there are  $n^{n-2}$  of such trees, so that exhaustive enumeration is prohibitive already for  $n \sim 10$ . However it turns out that the MST problem is tractable, in the sense depicted above. The following lemma is crucial. Let  $U \subset V$  be a given subset of the vertices of  $\mathcal{G}$ , and let  $e$  be the edge with the smallest weight of all the edges connecting  $U$  and  $V - U$ . Then  $e$  is part of the MST. The proof follows by contradiction. Suppose indeed that  $e$  connects the vertexes  $u \in U$  and  $v \in V - U$ , and that  $\mathcal{T}$  is a MST not containing  $e$ . By hypothesis  $\mathcal{T}$  contains a path from  $u$  to  $v$ , which, together with  $e$ , forms a closed loop. This path has to include another edge  $f$  (or edges) connecting  $u$  and  $v$ . Then  $\mathcal{T} + e - f$  is a spanning tree with less total weight than  $\mathcal{T}$ . Using this lemma it is straightforward to solve MST edge by edge with a fast, i.e. polynomial, algorithm.

### Travelling salesman problem

Let us now investigate a slightly different problem which admits a similar formalization. Suppose to plan an itinerary for a travelling salesman who shall visit once all the  $n$  cities, minimizing the path he has to travel. The cities represent the vertexes of a weighted graph  $\mathcal{G}$ , where each edge connecting the  $i$ -th and the  $j$ -th city is weighted by their distance  $d_{i,j}$  (every pair of distinct vertexes is connected by an edge, so that the resulting graph is fully connected). If we define a *walk* as an alternating sequence of vertexes and edges, and a *loop* as a closed walk, a generic itinerary of the salesman shall be represented by a loop on  $\mathcal{G}$  containing all its vertexes and in which all of them are distinct. This kind of itinerary is also named

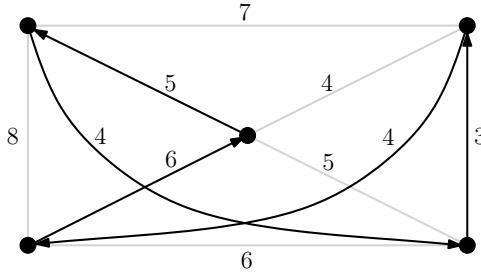


Fig. 1.14 A fully connected weighted graph and the TSP solution (dark edges); that is, the shortest itinerary to be travelled by a salesman.

a *Hamiltonian cycle*. Solving the travelling salesman problem (TSP) consists in finding the Hamiltonian cycle in a fully connected graph with the minimum weight.

Even for this problem, the number of possible solutions (itineraries) is very large, and is given by the  $(n - 1)!$  cyclic permutations  $\pi : [1 \dots n] \rightarrow [1 \dots n]$  of  $n$  cities. Exhaustive enumeration thus scales even faster than exponentially. Unfortunately no polynomial algorithm for TSP has been found yet. According to our previous classification the problem is intractable, even if there are efficient algorithms to find approximate “good” solutions but do not guarantee to yield the “best” solution. At present there is however no proof excluding the existence of a polynomial algorithm for TSP.

### 1.3.1.2 \* Decision problems

Another class of problems poses a simpler question, which can be answered without comparing all the feasible solutions. In the specific, we focus on *decision problems*, where the solution is given in terms of “yes” or “no” to a specific question. Even in this case exhaustive searching would enable to answer the question, but this approach is typically intractable and in many cases a polynomial wayout can be found.

#### Eulerian cycle

A famous example, which was discussed and first solved by Euler, is dated back to the 18th century and is named as the seven-bridge problem of Königsberg. At that time, seven bridges crossed the Pregel river and its two arms, (see the left panel of Fig. 1.15). The problem asks whether it was possible to walk through the four land masses of the city crossing each bridge exactly once, and returning back home. Except at the endpoints of the walk, whenever one enters a land mass by a bridge, one has also to leave the same land mass by a bridge. In other words, the number of times one enters a non-terminal land mass equals the number of times one leaves it. If every bridge has to be traversed exactly once, the number of bridges touching each non-terminal land mass should be even. However, all the four land masses of Königsberg are touched by an odd number of bridges. Since at most two land

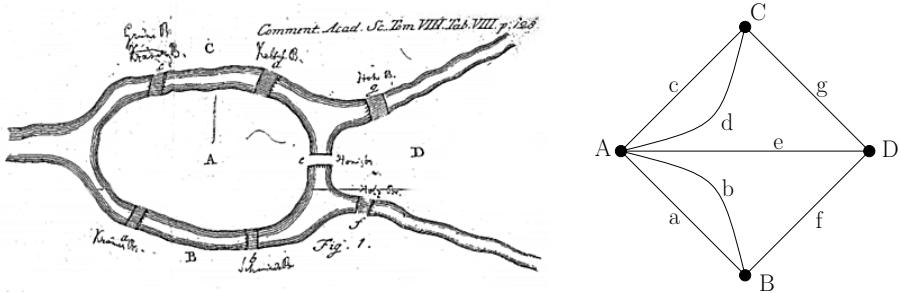


Fig. 1.15 The seven bridges of Königsberg's bridge puzzle (left) and its formal representation in terms of a graph (right). In the graph, the island and the riverbanks are represented by vertices, while the bridges are the edges. The figure on the left is taken from Euler (1736).

masses can serve as the endpoints of a walk, one cannot give a positive answer to the original problem of finding a walk traversing each bridge once.

This puzzle can be formalized by an unweighted graph  $\mathcal{G} = (V, E)$ , in which the vertices  $v \in V$  represent the land masses, while the bridges are the edges  $e \in E$  (see the right panel of Fig. 1.15 formalizing the seven-bridge problem of Königsberg). The goal is finding the existence of a loop in  $\mathcal{G}$  that contains all the edges  $E$ , and where all the edges are distinct. This kind of loop is called Eulerian cycle, and the question is named as the Eulerian-cycle problem (EC). Following Euler's argument, it is easy to see that EC can be solved in a polynomial time. Indeed, in an Eulerian cycle, one has to go through each vertex in such a way that the incoming edge is different from the outgoing one; that is, with the exception of the endpoints of the loop, the number of edges adjacent to the vertices has to be even. This condition is also sufficient to have an Eulerian cycle.

### Hamiltonian cycle

A slightly different problem, called Icosian puzzle, was posed by Hamilton in the 19th century. In its general formulation, his question asks for the existence of a loop in an unweighted graph  $\mathcal{G} = (V, E)$ , that contains all its vertices (instead of its edges) once. This kind of loop is also named Hamiltonian cycle, and the question is named as the Hamiltonian-cycle problem (HC).

Despite the similarity with EC, this puzzle is believed to be intractable. More specifically, we know that HC is a special case of TSP. Indeed the former can be obtained by setting the distance between two cities to one if they are adjacent (i.e., connected by an edge) and two otherwise, and verifying whether or not the total travelled distance is equal to  $n$ . As we shall see later in Sec. 1.3.2, HC is said to be *reducible* to TSP. Furthermore it can be shown that the two problems belong to the same complexity class.

### Satisfiability

Another fundamental decision problem refers to the satisfiability of a given Boolean formula. Let us call a bit  $a$  (or its negation  $\bar{a}$ ) a literal. Literals can be combined in *clauses*, using AND and OR gates. Different clauses can be combined to yield Boolean functions. These can be always written in *conjunctive normal form* (CNF), i.e. as a set of clauses combined with the AND operator, and where the literals in each clause are combined with the OR operator. The satisfiability problem (SAT) inquires for the existence of an assignment of the  $N$  variables  $\mathbf{a} \equiv (a_1, \dots, a_N)$  satisfying a given Boolean function  $f(\mathbf{a})$  with  $M$  clauses. In practice, a tentative assignment of the variables gives a positive outcome if and only if in each of the  $M$  clauses there is at least one literal equal to 1. Fixing the number of literals in each clause to a given amount  $k$ , defines the so-called  $k$ -SAT problem. The 1-SAT and the 2-SAT are known to be tractable. For general SAT and  $k$ -SAT with  $k \geq 3$  no polynomial algorithm is known.

**Exercise 1.2** Solve the SAT problem for the following CNF functions:

$$f_1(\mathbf{a}) = (a_1 \vee \bar{a}_2 \vee a_3) \wedge (a_2 \vee \bar{a}_3) \wedge (\bar{a}_1 \vee a_2) \wedge (\bar{a}_1 \vee \bar{a}_3), \quad (1.23)$$

$$f_2(\mathbf{a}) = (\bar{a}_1 \vee a_2) \wedge (\bar{a}_2 \vee a_3) \wedge (\bar{a}_1 \vee a_3) \wedge a_1. \quad (1.24)$$

**Exercise 1.3** Three people, hereafter named Alice Bob and Charlie, are accused of stealing money from a bank. During the interrogation at the police station, they make the following claims:

- (1) Alice says: “Bob is guilty and Charlie is innocent”.
- (2) Bob says: “If Alice is guilty, then Charlie is guilty too”.
- (3) Charlie says: “I am innocent. One of the others, or maybe both are guilty”.

Are the three statements contradictory? Assuming that all of them are guilty, who lied during the interrogation? Assuming that nobody lied, who is guilty? To answer the questions, translate the above sentences into a Boolean form and solve the corresponding satisfiability problem.

At this point it is worth stressing that the computational complexity is based on the worst-case analysis. Sometimes it might happen that an algorithm requires exponential time only for pathological instances. For example, let us concentrate on 3-SAT. Solving it for a generic Boolean function  $f$  generally requires a polynomial running time, unless the ratio  $M/N$  is carefully tuned. If  $M/N$  is small, the problem is underconstrained and has many possible solutions; a clever algorithm will find one of them quickly. If  $M/N$  is large, the problem is overconstrained and has a number of contradictory constraints which again can be easily discovered. The real hard instances correspond to ratios  $\alpha = M/N$  close to some critical value  $\alpha_c$ . It turns out that the transition from underconstrained to overconstrained formulas resemble the properties of phase transitions in a physical system, and have been analyzed with methods from statistical mechanics.

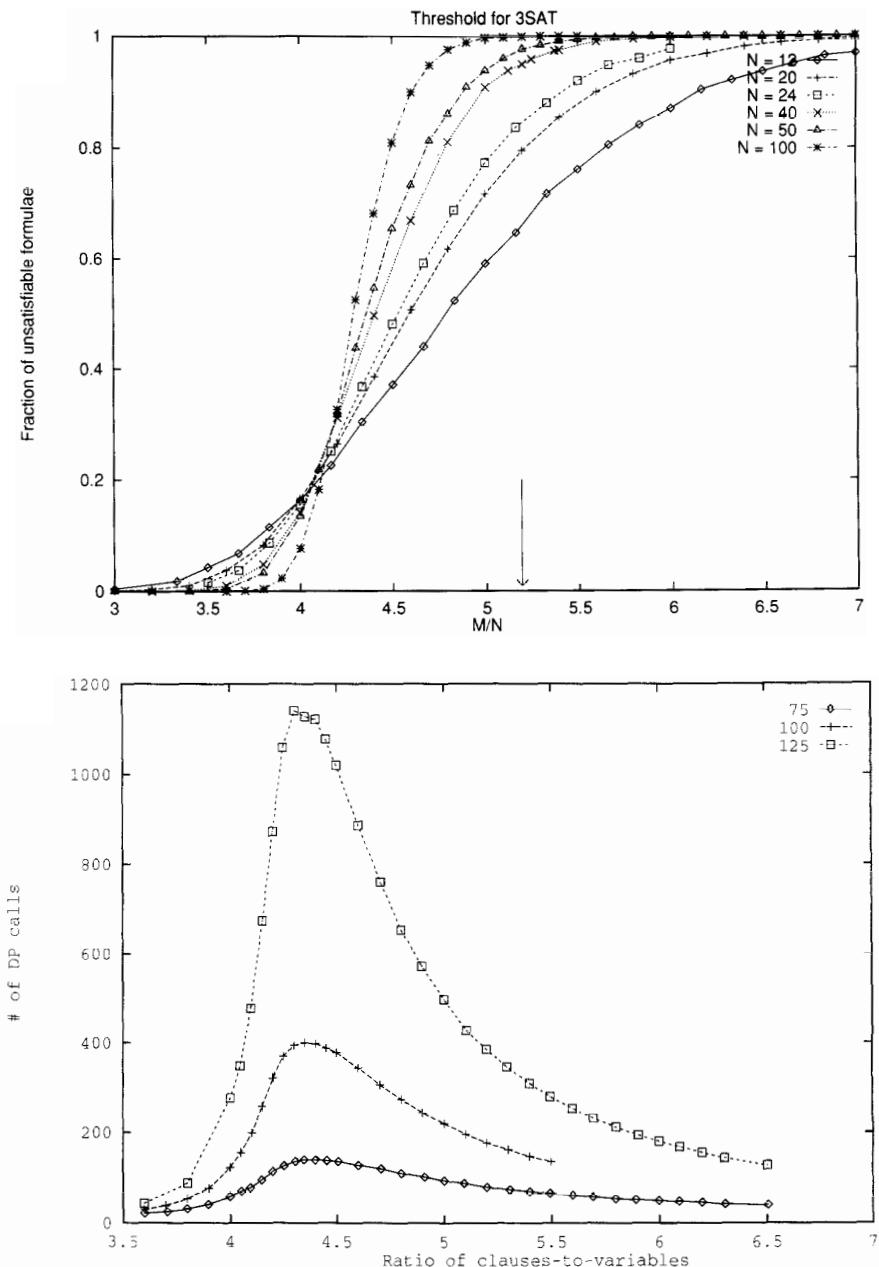


Fig. 1.16 Critical behaviour in the computational cost of 3-SAT. Ten thousands of random Boolean formulas have been drawn using a set of  $N$  variables (different curves) and keeping the clause-to-variable ratio  $\alpha$  fixed. The upper panel shows that the fraction of unsatisfied formulas grows monotonically with  $\alpha$  and sharpens up when increasing  $N$ , as is typical of threshold phenomena. In the lower panel the median cost of determining satisfiability is plotted as a function of  $\alpha$ . The computational cost is quantified by the number of calls to the extended Davis-Putnam recursive algorithm. The figures are reprinted with permission from Selman and Kirkpatrick (1996). ©(1996) Elsevier Science B.V.

Let us have a closer look at the 3-SAT numerical data shown in the upper panel of Fig. 1.16, after averaging over thousands of randomly selected CNF Boolean functions (see Selman and Kirkpatrick, 1996). Here  $\alpha$  represents the control parameter ( $x$  axis), while the order parameter of the transition is the probability  $\mathcal{P}$  of a given formula being unsatisfiable ( $y$  axis). We immediately realize that the transition between all satisfiable and all unsatisfiable formulas sharpens up with increasing system size  $N$ , which is a general feature of critical phenomena. We can gain more insight by employing a finite-size scaling. After introducing the scale-invariant parameter

$$y = N^{1/\nu}(\alpha - \alpha_c)/\alpha_c \quad \text{with } \nu \text{ and } \alpha_c \text{ constants,} \quad (1.25)$$

it can be seen that  $\mathcal{P}$  exhibits a universal behaviour according to

$$\mathcal{P}(y) = e^{-2^{-y}}. \quad (1.26)$$

The data are best fitted by choosing  $\nu = 1.5$  and  $\alpha_c \approx 4.17$ , which therefore represents the critical point for the satisfiability transition. The lower panel of Fig. 1.16 displays the computational cost as a function of  $\alpha$  using a specific recursive search algorithm commonly employed in solving combinatorial problems. A sudden growth of time resources is signaled for  $\alpha \sim \alpha_c$ . Even in this case the curves sharpen up for higher values of  $N$ , so that the time resources increase rapidly with the system size. In particular the maximum cost at  $\alpha_c$  increases exponentially as  $e^{N/c}$  with  $c \approx 11.1$ , thus spotlighting the hardness of the problem.

Similar finite-size scaling analysis have been performed for  $k$ -SAT with a generic  $k > 2$  dimension of the clauses, pointing out an analogous qualitative behaviour (see Monasson *et al.*, 1999).

### 1.3.2 Complexity classes

After having discussed the solvability of some prototype computational problems in the previous section, we now come back to our working plan and provide a rigorous definition of the most important complexity classes. We will mostly concentrate on time resources (CPU time) needed to run algorithms which solve the assigned problem in the worst possible instance; that is, the supremum over all possible input states. Of course the worst-case time is an upper bound for the actual running time.

In the theory of computational complexity, it is customary to say that a problem belongs to class **P** if it can be solved in *polynomial time*, namely, in a number of steps that is polynomial in the input size. According to what we learnt in Sec. 1.3.1, examples of **P** problems are MST, EC, 1-SAT and 2-SAT. Another landmark class is **NP**, standing for *non-deterministic polynomial time*. This denotes problems which can be solved in polynomial time by a non-deterministic algorithm. The latter is defined as the same as an ordinary algorithm, except that it may use an additional instruction that splits the computation into two parallel processes. By iterating this kind of procedure, the algorithm may branch as a tree into a sequence of parallel computations. It is easy to show that **NP** problems are equivalently defined as those admitting an output whose validity can be verified in polynomial time. It is thus evident that, for a **NP** problem, the only source of intractability is

the typical exponential size of the search space. All the problems discussed before belong to **NP**, however there are puzzles which do not enter this classification. An example of problem that is outside the **NP** class is a simpler version of the halting problem, which asks whether a deterministic Turing machine  $T$  will halt in at most  $n$  steps, for some given input  $x$  (written in binary notation). Such a task can be only solved in exponential time, since one simulation requires  $O(n)$  time, while the input  $x = n$  is encoded using  $O(\log n)$  bits. It is clear that  $\mathbf{P} \subseteq \mathbf{NP}$ . Nonetheless all **NP** problems have a chance to be in **P**. Nowadays, it is still a fundamental open problem of mathematics and computer science whether there exist problems in **NP** that are not in **P**. It is conjectured, though not proven, that  $\mathbf{P} \neq \mathbf{NP}$ . If this were the case, there would be problems hard to solve but whose solution could be easily checked. As previously hinted, TSP, HC and  $k$ -SAT with  $k > 2$  are believed to be in **NP** but not in **P**.

Another important concept is that of the reducibility of two problems. Namely, problem A is said to be *polynomially reducible* to problem B ( $A \leq_p B$ ) if there exists a polynomial algorithm for A, provided there is a polynomial algorithm for B. This is equivalent to say that there exists an algorithm solving A as a suitable instance of B. It is evident that  $A \leq_p B$  implies that A cannot be harder than B. For example, as discussed in Sec. 1.3.1, we know that EC  $\leq_p$  TSP. We say that a **NP** problem is also **NP**-complete (**NPC**) if any problem in **NP** is *polynomially reducible* to it. Therefore, if an algorithm capable of efficiently solving an **NPC** problem is discovered, then we should conclude that  $\mathbf{P} = \mathbf{NP}$ . All the problems discussed in Sec. 1.3.1 which are believed to be intractable are also **NPC** (in the example discussed above it has been also proven that TSP  $\leq_p$  EC). There are however problems that are conjectured to be neither **P** nor **NPC**. One of them is the integer-factoring problem. This belongs to the **NP** class, since it is easy to check whether a given number  $m$  is a prime factor of an integer  $N$ , but is believed to be not in **P** (at least on a classical computer). Moreover we are not aware of a polynomial reduction which classifies it as **NPC**. It has been proven that, if  $\mathbf{P} \neq \mathbf{NP}$ , then there exist **NP** problems that are neither in **P** nor in **NPC**. The possible maps of **NP** problems are drawn in Fig. 1.17.

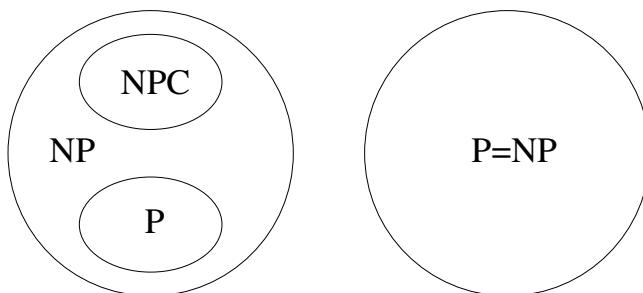


Fig. 1.17 Possible maps of **NP** problems. It is conjectured, though not proven, that the left map is the correct one.

So far, we have only discussed time resources. However, to run a computation, space and energy resources are also important. The discussion of energy resources will be postponed to Secs. 1.5–1.7.

*Space (i.e., memory) resources.* Space and time resources are linked. Indeed, if at any step of, say, a Turing machine, we use a new memory cell, then space and time resources scale equivalently. However, there is a fundamental difference: space resources can be reused. We define **PSPACE** as the class of problems that can be solved by means of space resources that are polynomial in the input size, independently of the computation time. It is evident that  $\mathbf{P} \subseteq \mathbf{PSPACE}$ , since in a polynomial time a Turing machine can explore only a polynomial number of memory cells. It is also conjectured that  $\mathbf{P} \neq \mathbf{PSPACE}$ . Indeed, it seems reasonable that, if we have unlimited time resources and polynomial space resources, we can solve a larger class of problems than if we have polynomial time (and space) resources. However, there is no proof that there exist problems in **PSPACE** not belonging to **P**. It is easy to show that **NP** is a subset of **PSPACE**, that is, any problem in **NP** can be solved by means of polynomial space resources. Indeed, we can always try to find the solution of an **NP** problem by exhaustive search; since each possible solution can be verified in polynomial time and space for **NP** problems, we may reuse the same (polynomial) space resources to test all possible solutions. In summary, we know that  $\mathbf{P} \subseteq \mathbf{NP} \subseteq \mathbf{PSPACE}$ , but we do not know if these inclusions are strict.

Finally, let us consider the case in which a probabilistic computer, such as a probabilistic Turing machine, is used to solve a decision problem. The problem is said to be in **BPP** class (*bounded-error probabilistic polynomial time*) if there exists a polynomial-time algorithm such that the probability of getting the right answer is larger than  $\frac{1}{2} + \delta$  for every possible input and  $\delta > 0$ . The Chernoff bound, discussed in the next subsection, shows that the probability of getting the right answer can be quickly amplified by running the algorithm several times and then applying majority voting. Indeed, in order to reduce the error probability below  $\epsilon$  in a **BPP** problem, it is sufficient to repeat the algorithm a number of times logarithmic in  $1/\epsilon$ .

The following simple example demonstrates that sometimes it is convenient to relax the requirement that a solution is always correct and allow some very small error probability. Let us consider a database of  $N$  bits  $j_1, \dots, j_N$ . Suppose we know in advance that either they are all equal ( $j_1 = \dots = j_N = 0$  or  $j_1 = \dots = j_N = 1$ ) or half of them are 0 and half 1. We call the first possibility “constant” and the second “balanced”. Our problem is to distinguish between these two possibilities. In the case in which the bits are all equal, we must observe  $N/2 + 1$  of them to be sure of our answer to the problem. Indeed, if we observe  $N/2$  bits (for instance, from  $j_1$  to  $j_{N/2}$ ) and they are all equal to, say, 0, we cannot exclude with certainty the balanced case: we could have  $j_1 = \dots = j_{N/2} = 0$  and  $j_{N/2+1} = \dots = j_N = 1$ . To solve our problem probabilistically, we toss a random number  $i$  between 1 and  $N$

and we observe  $j_i$ . This is repeated  $k$  times. If we find two different bits, then we can conclude with certainty that we are in the balanced case. If all bits are constant, we say that we are in the constant case. Of course, there is a chance that we give the wrong answer to the problem. However, the probability of obtaining the same response every time when we are in the balanced case is  $1/2^{k-1}$ . Therefore, we can reduce the probability of error below some level  $\epsilon$  if  $k$  is such that  $1/2^{k-1} < \epsilon$ . This is obtained in  $k = O(\log(1/\epsilon))$  bit observations, independently of  $N$ .

The example shows that **BPP** better than **P** should be regarded as the class of problems that can be solved efficiently on a classical computer. It is evident that **P**  $\subseteq$  **BPP**, while the relation between **NP** and **BPP** is presently unknown. In particular, there are problems known to be in **BPP** but not in **P**. One of them is the polynomial identity testing; that is, determining whether a polynomial  $P(x_1, \dots, x_n)$  is identically equal to zero or not. While we are not aware of any deterministic polynomial algorithm to solve it, it has been shown independently by Schwartz (1980) and Zippel (1979) that it suffices to choose each variable's value uniformly at random from a finite subset of at least  $d+1$  values to achieve bounded error probability,  $d$  being the total degree of the polynomial  $P$ . It is however a matter of fact that the number of such problems is decreasing, as it occurred for the so-called primality testing. This is a paradigmatic puzzle in arithmetics, which consists in determining whether a number is prime or composite. For a long time it was known to be in **BPP**, but not known to be in **P**. It was only recently that Agrawal *et al.* (2004) discovered a new algorithm running in  $\tilde{O}(\log^{12} n)$  operations, that is able to solve the problem (the symbol  $\tilde{O}(f(n))$  is a shorthand for  $O(f(n) \log^k f(n))$  for some integer  $k$ ). A few years later, this bound was further reduced to  $\tilde{O}(\log^6 n)$ . Nowadays computer scientists are inclined towards the conjecture that **P** = **BPP**.

We close this section by introducing the class **BQP**, standing for *bounded-error quantum probabilistic polynomial*. We say that a decision problem belongs to class **BQP** if there exists a polynomial-time quantum algorithm giving the right answer with probability larger than  $\frac{1}{2} + \delta$  (with  $\delta > 0$ ). Since the integer-factoring problem may be reduced to a decision problem, Shor's algorithm belongs to this class. Indeed, it solves the factoring problem in  $O(n^2 \log n \log \log n \log(1/\epsilon))$  operations, where  $\epsilon$  is the probability of error. Note that  $\epsilon$  does not depend on the input size  $n$  and therefore we can take it as small as we like and still have an efficient algorithm. We stress that there is no known classical algorithm, deterministic or probabilistic, that solves this problem in a number of operations polynomial in the input size. We know that **P**  $\subseteq$  **BPP**  $\subseteq$  **BQP**  $\subseteq$  **PSPACE** and it is conjectured that **BPP**  $\neq$  **BQP**, namely, that a quantum computer is more powerful than a classical computer. The relation between **BQP** and **NP** is presently unknown. It is worth stressing that, even if integer factoring is **NP**, unfortunately it is not **NPC**, hence Shor's algorithm cannot be regarded as a polynomial-time quantum algorithm to which all **NP** problems can be reduced.

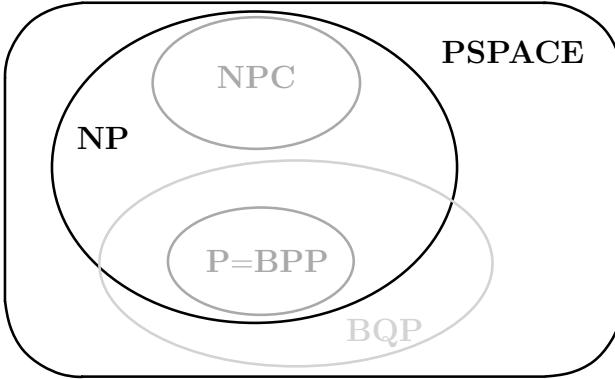


Fig. 1.18 Tentative map of complexity.

### 1.3.3 \* The Chernoff bound

When solving a decision problem, a probabilistic algorithm produces a non-deterministic binary output  $f$ . Let us assume, without any loss of generality, that the correct answer is  $f = 1$  and the wrong answer is  $f = 0$ . Let us repeat the algorithm  $k$  times and then apply majority voting. At each step  $i$  ( $i = 1, 2, \dots, k$ ) we obtain  $f_i = 1$  with probability  $p_1 > 1/2 + \delta$  and  $f_i = 0$  with probability  $p_0 < 1/2 - \delta$ . Majority voting fails when  $s_k \equiv \sum_i f_i \leq k/2$ . Note that the average value of  $s_k$  is larger than  $k(1/2 + \delta) > k/2$ . The most probable sequences  $\{f_i\}$  that lead to a failure of majority voting are those in which  $s_k$  is nearest to its average value, that is,  $s_k = k/2$ . Such sequences occur with probability

$$p(\{f_1, f_2, \dots, f_k\}; s_k = \frac{k}{2}) < (\frac{1}{2} - \delta)^{\frac{k}{2}} (\frac{1}{2} + \delta)^{\frac{k}{2}} = \frac{(1 - 4\delta^2)^{\frac{k}{2}}}{2^k}. \quad (1.27)$$

Since there are  $2^k$  possible sequences  $\{f_i\}$ , we may conclude that majority voting fails with probability

$$p(s_k \leq \frac{k}{2}) < 2^k \frac{(1 - 4\delta^2)^{\frac{k}{2}}}{2^k} = (1 - 4\delta^2)^{\frac{k}{2}}. \quad (1.28)$$

Finally, since  $1 - x \leq \exp(-x)$ , we obtain the *Chernoff bound*:

$$p(s_k \leq \frac{k}{2}) < \exp(-2\delta^2 k). \quad (1.29)$$

Therefore, the error probability drops below  $\epsilon$  after a number of runs

$$k > \frac{1}{2\delta^2} \ln\left(\frac{1}{\epsilon}\right). \quad (1.30)$$

### 1.4 \* Computing dynamical systems

One of the main applications of computers is the simulation of dynamical models describing the evolution of complex systems. We refer here not only to problems of

interest for physics and mathematics, but also to a much wider class of problems in different fields such as chemistry, biology, economics, medicine, engineering, social sciences, meteorology, population dynamics and so on. From the viewpoint of computational complexity, the following question naturally arises: can such complex problems be solved efficiently? More precisely, given a generic dynamical system, is it possible to find its solution at time  $t$  efficiently? That is, since the number of bits required to specify the time  $t$  is  $\log t$ , can we solve the problem in a number of operations polynomial in  $\log t$ ? We shall see in this section that this is not the case for a generic dynamical system, whose evolution is typically described by non-linear equations.

#### 1.4.1 \* Deterministic chaos

Deterministic chaos has been one of the most significant discoveries of the last century. Let us briefly explain the meaning of the wording “deterministic chaos”. A system is said to be *deterministic* when its future, as well as its past, are determined by its present state. For instance, Newton’s laws of motion unambiguously determine the future (and the past) of a system, once its state at some time  $t_0$  is assigned. On the other hand, the motion of the system can be so complex as to be indistinguishable in practice from purely *chaotic* motion. This property allows us to reconcile the determinism of physical laws and the apparent chaoticity of natural phenomena, such as turbulence, which we observe in everyday life. Hence, the term “deterministic chaos” is not self-contradictory, since a phenomenon can be both deterministic and chaotic: deterministic since it is governed by laws that fully determine its future state from initial conditions; chaotic since its motion is so complex as to be completely *unpredictable* in practice. Let us try to clarify this statement. We first consider the harmonic oscillator, namely, the simplest example of a classical solvable or so-called integrable system. Its equation of motion,  $d^2x/dt^2 + \omega^2 x = 0$ , can be solved analytically. The solution is  $x(t) = x_0 \cos(\omega t + \phi_0)$ , with  $x_0$  and  $\phi_0$  the initial conditions. Given a time  $t$ , a computer can output  $x(t)$  from the above solution with  $O(\log t)$  operations. In contrast, for chaotic motion, as we shall see below, the number of operations required is  $O(t)$ . This means that, while for an integrable system the motion is predictable and computable, for a chaotic system it is not possible to predict the future “before it arrives”. That is, it is not possible to describe the orbit of a chaotic system by means of an algorithm that scales better than  $O(t)$ : the system itself is “its own best computer”.

In order to clarify this concept, let us consider a conservative system described by the Hamiltonian  $H(q, p)$ , where  $q = (q_1, \dots, q_n)$  and  $p = (p_1, \dots, p_n)$  denote canonical variables. Since the total energy  $E$  is a constant of motion, the system’s orbit moves on the constant-energy surface, defined by the equation  $H(q, p) = E$ . We now make a *partition* of this surface, that is, we divide it into a finite set of non-overlapping cells and we identify each cell by means of an integer. If we have perfect knowledge of the system’s orbit, we can assign, at time intervals  $\tau$ , the number of

the cell in which the system resides. In this way, we obtain a sequence of integers, which provides a coarse-grained description of the orbit. For a chaotic system, no regularity appears. The knowledge of the cells occupied by the system up to time  $t$  is not sufficient to determine the cell number at time  $t + 1$  (a unit of discrete time  $t$  corresponds to the time interval  $\tau$ ). Therefore, for chaotic orbits, knowledge of the coarse-grained past is not sufficient to determine the coarse-grained future. In contrast, this is possible in non-chaotic systems, since the coarse-grained orbit exhibits regularities. Note that no restrictions on the size of the partition have been made. That is, a sequence of finite precision measurements is unable to predict the future of a chaotic system, independently of their (finite) precision.

Let us illustrate the concept of deterministic chaos by means of an example. We consider the *logistic map*, one of the best known models for studying the transition to chaos. It is defined by the first-order difference equation

$$x_{n+1} = \alpha x_n(1 - x_n), \quad (1.31)$$

where  $0 \leq \alpha \leq 4$ , so that the unit interval  $[0, 1]$  is mapped into itself. The behaviour of the logistic map is very complicated and exhibits regions of regular or chaotic motion when the parameter  $\alpha$  is varied. In particular, the map is fully chaotic for  $\alpha = 4$ . Now, let  $x_n = \sin^2(\pi y_n)$  and substitute into (1.31) for  $\alpha = 4$ . After some straightforward algebra we find  $\sin^2(\pi y_{n+1}) = \sin^2(2\pi y_n)$ . Hence, the logistic map, for  $\alpha = 4$ , is equivalent to

$$y_{n+1} = 2y_n \pmod{1}. \quad (1.32)$$

This equation maps the unit interval  $[0, 1]$  into itself. It has the following simple analytic solution:

$$y_n = 2^n y_0 \pmod{1}. \quad (1.33)$$

A more transparent form of the solution may be obtained by writing the binary representation of the initial condition  $y_0$ :

$$y_0 = 0.1101001100011010\dots. \quad (1.34)$$

It is easy to check that each iteration of map (1.32) moves the decimal point, in the binary representation, one digit to the right and then drops the integer part to the left of the decimal point. Thus, one bit of information is erased at each map step.

It is now easy to show that the solution of the deterministic equation (1.32) is completely unpredictable. In our example the unit interval  $[0, 1]$  plays the role of the energy surface (the orbit resides in this interval). We partition this “energy surface” into two cells, a left cell ( $0 \leq y < 1/2$ ) and a right cell ( $1/2 \leq y < 1$ ). From the binary representation (1.34), we recognize that  $y_n$  resides in the left or right cell, depending on the value 0 or 1 of the first digit (after the decimal point) of its binary representation. Since the decimal point moves one digit to the right at each map step, the coarse-grained orbit corresponds to the binary representation (1.34) of the initial state  $y_0$ : 0 means the left cell and 1 the right cell. It is clear that, if

we know the first  $t$  numbers of the coarse-grained orbit, we cannot determine the number  $t + 1$ . Indeed, if we know the first  $t$  digits of the binary expansion of  $y_0$ , we cannot determine the subsequent digits. As time goes on, the solution will depend on ever diminishing details of the initial condition. In other words, when we fix  $y_0$  we supply the system with infinite complexity which arises owing to the chaotic nature of the motion.

How random is the solution of Eq. (1.32)? Let us assume that someone who knows the precise solution of (1.32) tells us the sequence of digits in  $y_0$ . Can we deduce whether this person is really telling us the solution of Eq. (1.32) or merely a sequence of random digits that he has obtained, for instance, by flipping a fair coin? (Say that 0 corresponds to heads, 1 to tails.) The answer is no. Indeed, we can easily convince ourselves that the set of all possible initial conditions  $y_0$  is in one-to-one correspondence with the set of all possible coin-tossing sequences.<sup>3</sup> Since a coin-tossing sequence is random, the binary representation of  $y_0$  is also random. Therefore, the orbit itself is also random.

#### 1.4.2 \* Algorithmic complexity

At this point we need to clarify exactly what we mean when we say that a digit string is *random*. Each binary digit carries one bit of information. Therefore, an  $n$ -bit binary sequence can carry  $n$  bits of information. However, if there are correlations between digits, the information contained in this  $n$ -bit string can be expressed by a shorter sequence. Following Kolmogorov, we define the *complexity*  $K_M(x)$  of a given  $n$ -bit sequence  $x$  as the bit length of the shortest computer program (algorithm) capable of computing this sequence using machine  $M$ . Note that complexity can be made machine independent. Indeed, Kolmogorov has proved the existence of a universal machine  $U$  such that

$$K_U(x) \leq K_M(x) + C_M , \quad (1.35)$$

where  $C_M$  is a positive constant, which depends on  $M$  but not on  $x$ .

Let us consider the sequence 01010101... This string can be computed by the program “PRINT 01,  $n$  times”. The length of this program is  $\log_2 n + A$ , where  $\log_2 n$  is the number of bits required to specify  $n$  and  $A$  is a constant that depends on the machine. Therefore, the complexity of this sequence is  $O(\log_2 n)$ , independently of the machine used. On the other hand, the complexity of an  $n$ -bit string  $x = (x_1, x_2, \dots, x_n)$  can never exceed  $O(n)$ . Indeed, this sequence can always be produced by the copy program “PRINT  $(x_1, x_2, \dots, x_n)$ ”, which is  $O(n)$  bits long. Following Kolmogorov, we say that an  $n$ -bit sequence is random if it cannot be calculated by a computer program whose length is smaller than  $O(n)$  bits. A chaotic orbit is random in the sense that it cannot be compressed into a shorter sequence; it is therefore unpredictable.

---

<sup>3</sup>Note that for chaotic dynamics, the set of coarse-grained orbits is complete, that is, it actually contains all possible sequences.

For infinite sequences, we can define the complexity as

$$K_\infty = \lim_{n \rightarrow \infty} [K^{(n)}/n], \quad (1.36)$$

where  $K^{(n)}$  is the complexity of the first  $n$  bits of the sequence. Note that Kolmogorov's result on the existence of a universal machine tells us that  $K_\infty$  is machine independent, this follows trivially from Eq. (1.35). It is possible to show that, in general, limit (1.36) exists. Martin-Löf proved that almost all sequences having positive complexity ( $K_\infty > 0$ ) would pass all computable tests for randomness. This justifies the statement that positive complexity sequences are random. Moreover, Martin-Löf proved that almost all sequences have positive complexity and are therefore random. It follows that almost all orbits that are solutions of map (1.32) are random; their information content is thus both infinite and incompressible.

A further consequence of algorithmic complexity theory is that almost all real numbers cannot be computed by finite algorithms. Of course, exceptions are the integer and rational numbers. We also note that irrational numbers such as  $\pi$  or  $e$  are not random, since efficient algorithms to compute them to any desired accuracy exist. These algorithms imply  $K_\infty = 0$ .

We can now clarify the connection between chaotic dynamics and positive algorithmic complexity. Chaos is defined in terms of sensitivity to initial conditions. If  $\delta x_0$  is an infinitesimal change of the initial condition for a given dynamical system and  $\delta x_t$  is the corresponding change at time  $t > 0$ , then in general

$$|\delta x_t| \approx e^{\lambda t} |\delta x_0|, \quad (1.37)$$

where  $\lambda$  is the largest so-called Lyapunov exponent. We say that the dynamics is chaotic when  $\lambda > 0$ , that is, when any amount of error in determining the initial conditions diverges exponentially, with rate  $\lambda$ . It is clear from the solution (1.33) of map (1.32) that

$$|\delta y_t| = 2^t |\delta y_0| = e^{(\ln 2)t} |\delta y_0|, \quad (1.38)$$

and therefore  $\lambda = \ln 2$ . The exponential sensitivity to initial conditions means that one digit of orbital accuracy is lost per suitably chosen unit of time. To recover this digit of accuracy, we must add one digit of accuracy to the initial condition  $y_0$ . Therefore, to be able to follow our orbit up to time  $t$  accurately, we must input  $O(t)$  bits of information. Thus, a chaotic orbit has positive complexity, namely, it is random.

For non-chaotic systems (in particular, for integrable systems) errors only grow linearly with time and therefore knowledge of the coarse-grained past is sufficient to predict the future.

In conclusion, the solutions of chaotic systems cannot be computed efficiently, since the computational resources required to determine the orbit up to time  $t$  grows like the time  $t$  itself. Only for non-chaotic systems is it possible, at least in principle, to compute the solution efficiently. That is, with computational resources that grow as the input size  $\log t$ .

## 1.5 Energy and information

### 1.5.1 Maxwell's demon

In this section, we discuss the connection between energy and information, two concepts that, at first sight, might seem hardly related. We may say that the discussion on the relation between energy and information goes back to *Maxwell's demon* paradox, introduced by James Clerk Maxwell in 1867. He imagined that a demon was capable of monitoring the positions and velocities of the individual molecules of a gas contained in a box and initially in thermal equilibrium at temperature  $T$  (see Fig. 1.19). At equilibrium, the individual molecular velocities are

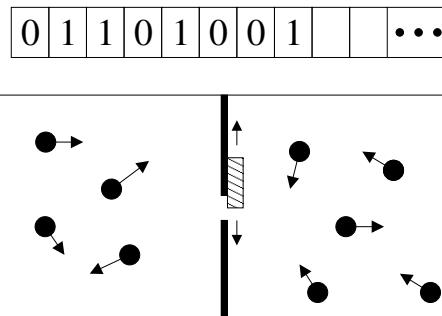


Fig. 1.19 Schematic drawing of Maxwell's demon paradox. The demon's memory is represented by a binary string and stores the results of his measurements of the positions and velocities of the molecules.

distributed according to the Maxwell distribution and with random directions. The box is divided in two halves, which communicate through a little door. The demon opens and closes the door so as to allow the faster molecules to move from the right half of the box to the left half and the slower molecules to move in the opposite way. By doing this many times the demon separates the faster molecules, which end up on the left side of the box, from the slower molecules, which finish on the right side. As a consequence, the temperature  $T_l$  of the gas in the left chamber becomes higher than the temperature  $T_r$  of the gas in the right chamber. Since we have obtained two gases at different temperatures, we can now use these two “thermal baths” to produce work. Therefore, the demon has been able to convert heat from a source at uniform temperature  $T$  into work, in apparent violation of the second law of thermodynamics. Actually, as we shall see later, there is no such violation, since the transformation of heat into work is not the only result of the process considered as a whole.

Maxwell's demon paradox may be equivalently stated in terms of *entropy*. Entropy is defined as

$$S = k_B \ln \Omega, \quad (1.39)$$

where  $k_B \approx 1.38 \times 10^{-23}$  Joule/K is Boltzmann's constant and  $\Omega$  is the number

of microscopic states of the system (that is, positions and velocities of the single molecules) that give the same macroscopic state (identified by a few parameters, such as the volume  $V$  and temperature  $T$ ). It is clear that the demon introduces order into the system: the faster molecules are constrained to stay on the left side of the container, the slower molecules on the right. Therefore, the number of microscopic states accessible to the system diminishes and the total entropy of the gas is reduced, thus apparently violating the second law of thermodynamics. Moreover, the demon could divide the box into many cells and separate with great accuracy the molecules according to their velocity. As a consequence, the entropy of the gas would become smaller and smaller as the number of cells increased. Indeed, our knowledge of the microscopic state of the system would increase, implying a decrease in the entropy. However, such violation of the second law is only apparent: indeed, a careful analysis shows that the total entropy (of the gas, demon and environment) does not decrease.

### 1.5.2 Landauer's principle

Maxwell's demon spawned much discussion and different solutions were proposed to solve the paradox. At the beginning, it was widely believed that the resolution of the paradox lays in the energy cost of the *measurements* performed by the demon. For example, in order to locate the molecules the demon needs to illuminate them and this has an energy cost. However, Rolf Landauer and Charles Bennett were able to show that the measurement process can, in principle, be performed without energy expenditure. They eventually succeeded in finding the solution of the paradox: the results of the measurements must be stored in the demon's memory. Since his memory is finite, the demon will eventually need to *erase* his memory to free up memory cells for new measurements. There is energy dissipation associated with this erasure. This is the content of Landauer's principle, stated in 1961.

**Landauer's principle:** *Each time a single bit of information is erased, the amount of energy dissipated into the environment is at least  $k_B T \ln 2$ , where  $k_B$  is Boltzmann's constant and  $T$  the temperature of the surrounding environment. Equivalently, we may say that the entropy of the environment increases by at least  $k_B \ln 2$ .*

Thus, the decrease in the entropy of the gas is compensated by an increase in the demon's entropy. To erase the information gathered by the demon in the measurement process, we must dissipate energy into the environment. Therefore, the energy cost, according to Landauer's principle, is not due to the measurement process itself, but to information erasure.

The following example is useful to illustrate Landauer's principle. Suppose that information is embodied in the state of a physical system, for instance, we might store a bit of information via a single molecule in a box. We say that the bit is set to 0 if the molecule is on the left side of the box, to 1 if it is on the right. Even though we have a single-molecule system, we may apply the laws of thermodynamics. As

is well known,

$$dE = \delta L + \delta Q, \quad (1.40)$$

where  $dE$  is the variation of the internal energy of the gas,  $\delta L$  the work done on the gas and  $\delta Q$  the heat absorbed by the gas. If we consider a quasi-static transformation (that is, a transformation so slow that we may consider the system to be always in an equilibrium state), we may write

$$dS = \frac{\delta Q}{T}, \quad (1.41)$$

where  $dS$  is the variation of the entropy of the gas. Let us assume that our box is in contact with a thermal bath at temperature  $T$  and that we compress the gas by means of a frictionless piston (see Fig. 1.20). If the displacement of the piston is  $dx$ , the work done on the gas is given by

$$\delta L = -Fdx = -pAdx = -pdV, \quad (1.42)$$

where  $F$  is the force of the gas on the piston,  $p$  its pressure,  $A$  the surface of the piston and  $V$  the volume of the gas. Of course, since we only have a single molecule, concepts such as pressure and force must be understood in a time-averaged sense, that is, we need to average over many collisions of the molecule against the piston. Let us consider a transformation that halves the volume of the gas. Taking into account the equation of state for ideal gases,

$$pV = Nk_B T, \quad (1.43)$$

where  $N$  is the number of particles in the gas (here  $N = 1$ ), we can compute the work done on the gas as follows:

$$L = - \int_V^{V/2} pdV' = - \int_V^{V/2} \frac{k_B T}{V'} dV' = k_B T \ln 2. \quad (1.44)$$

Note that here we have used the fact that the transformation is isothermal since the system is in contact with a thermal bath at temperature  $T$ .

As we have assumed that the gas is ideal, its internal energy does not change because the temperature is constant. Therefore, from the first law of thermodynamics, Eq. (1.40), the work done on the gas is transformed into heat dissipated into the environment:  $\Delta Q = -L$ . Note that  $\Delta Q < 0$  since heat is dissipated, not absorbed. The change in the entropy of the gas is given by Eq. (1.41):

$$\Delta S = \frac{\Delta Q}{T} = \frac{-L}{T} = -k_B \ln 2. \quad (1.45)$$

We have  $\Delta S < 0$  since, after compression, the volume available to the molecule is halved and therefore the number of available microstates is reduced correspondingly. The entropy of the system diminishes, while the entropy of the environment  $\Delta S_{\text{env}}$  increases: since the total entropy of the universe can never decrease, we have  $\Delta S + \Delta S_{\text{env}} \geq 0$ . Thus,  $\Delta S_{\text{env}} \geq k_B \ln 2$ , in agreement with Landauer's principle.

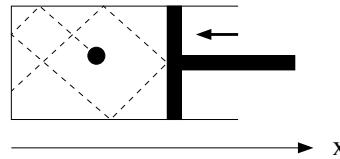


Fig. 1.20 Compression of a single-molecule gas by means of a piston.

Let us now assume that a binary message is stored by means of a sequence of single-molecule boxes. Each box carries a single bit of information, set in the state 0 or 1, depending on the left/right position of the molecule (see Fig. 1.21). We shall now show the validity of the following statement: the information contained in the message is proportional to the energy necessary to erase the message, that is, to move all the molecules to the left (or right) side of the boxes. First of all, we must define the information content of a message. It is defined as the information we should gain if we knew the values of the bits that constitute the message. Therefore, information can be seen as a measure of our *ignorance* about the message. If we already knew the values of the bits, we should obtain no further information from the message. In this case the information contained in the message is zero and, according to the statement made above, no energy expenditure is necessary to erase the message. Let us show that indeed no work is required to set the state of each bit to 0. Indeed, if the molecule is already on the left side, no further action is required. On the other hand, if the molecule is on the right side, we can move it to the left without energy expenditure: it is sufficient to enclose it inside a smaller, inner box and to shift this box to the left, as shown in Fig. 1.22. No work is required to perform this operation, since the molecule bounces as many times against the left wall of the inner box as against the right. Only in the case in which we do not know in advance the position of the molecule must we halve the volume of the gas and, as we have seen previously, this requires work  $L = k_B T \ln 2$  to be done on the gas. In this latter case the information content of the message is different from zero and we must expend energy to erase it.

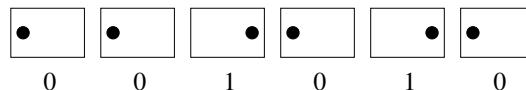


Fig. 1.21 A sequence of single-molecule boxes associated with a binary string.

### 1.5.3 Extracting work from information

For a better understanding of the relation between information and energy, it is instructive to consider the following example, devised by Bennett, which shows that information may be used as fuel to move a machine. A trolley is in contact with a thermal bath at temperature  $T$  and a ribbon, made of a string of single-molecule

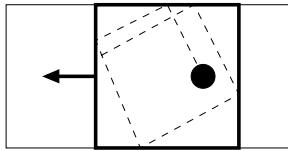


Fig. 1.22 A procedure enabling the transfer of a molecule to the left side of a box without energy cost.

boxes, enters the trolley (see Fig. 1.23). If we know in advance the left/right position of every molecule, such information can be used to extract work to move the trolley. Indeed, it is sufficient to insert a piston in the middle of each box. As shown in Fig. 1.24, the piston is movable to the right if the molecule is on the left side of the box and to the left if the molecule is on the right side. Since the whole system is at temperature  $T$ , we extract work  $L = k_B T \ln 2$ . If we have an  $N$ -bit ribbon, the total work is  $Nk_B T \ln 2$  and it may be used to displace the trolley. We stress that, when the ribbon comes out of the trolley, the molecules can be anywhere inside the volume  $V$ . The information content of the string of boxes has been completely lost and used as a fuel to move the trolley. On the other hand, if we do not know in advance the left/right position of the molecules, then we cannot extract useful work: indeed, if we insert a piston, half of the time the gas produces work, but the other half of the time work is done on the gas. On average, the extracted work is thus zero.

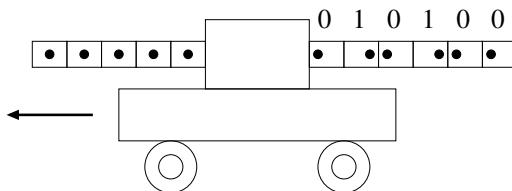


Fig. 1.23 Use of information to produce work.



Fig. 1.24 Extraction of work from a single-molecule gas, which at the beginning is on the left or right side of a container.

## 1.6 Reversible computation

In this section we discuss the energy requirements for computation. Most of the logic gates introduced in Sec. 1.2 are *irreversible*. This means that, given the

output, we cannot recover the input. For instance, if the OR gate has output 1, the input bits could have been set to  $(0, 1)$ ,  $(1, 0)$  or  $(1, 1)$ . The Boolean functions  $f : \{0, 1\}^2 \rightarrow \{0, 1\}$  erase a bit of information and therefore, according to Landauer's principle, the amount of energy dissipated into the environment must be at least  $k_B T \ln 2$ . This is analogous to the example of a molecule in a box discussed in the previous section: instead of halving the volume accessible to a gas, we pass from a two-bit input to a single-bit output.

We note that the value  $k_B T \ln 2$  represents only a lower bound for the energy dissipation of a two-bit irreversible gate. Present-day, real computers dissipate more by orders of magnitude. However, the energy dissipation per gate has been reduced enormously over the years thanks to technological progress. On the other hand, if we increase the power of a computer (that is, the number of operations per second), we also increase the energy dissipated, unless we are able to reduce the energy cost of each elementary logic gate. It is therefore important to keep in mind that Landauer's principle sets a lower bound to any future reduction of the energy dissipated by irreversible computation.

Since the energy cost of computation is related to irreversibility, the following question arises: is it possible to build a reversible computation without energy consumption? We can see in advance that it should be possible, since the fundamental laws of physics (Newton's equations in classical mechanics) are reversible. Therefore, there must be some underlying reversible physical process that allows us to implement irreversible gates. Indeed, any irreversible function  $f : \{0, 1\}^m \rightarrow \{0, 1\}^n$  can be embedded into a reversible function. It is sufficient to define the function

$$\tilde{f} : \{0, 1\}^{m+n} \rightarrow \{0, 1\}^{m+n}, \quad (1.46)$$

such that

$$\tilde{f}(x, y) = (x, [y + f(x)] \text{ (mod } 2^n)), \quad (1.47)$$

where  $x$  represents  $m$  bits, while  $y$  and  $f(x)$  represent  $n$  bits. Since  $\tilde{f}$  takes distinct inputs into distinct outputs, it is an invertible  $(m + n)$ -bit function.

According to the above argument, it is possible to find universal reversible gates for computation. We shall indeed show that the universal gates NAND and FANOUT can be constructed from reversible gates. In order to avoid information loss, a reversible function must take  $n$  input bits into  $n$  output bits. Reversible functions are permutations of the  $2^n$  possible inputs and therefore their number is  $(2^n)!$ . For  $n = 1$  there are 2 reversible single-bit gates: the identity and the NOT gate. A very important two-bit gate is the controlled-NOT (CNOT) or reversible XOR, shown in Fig. 1.25. The first bit ( $a$ ) acts as a control and its value is unchanged on output:  $a' = a$ . The second (target) bit is flipped if and only if the first bit is set to one and therefore  $b' = a \oplus b$ . Therefore, on output the second bit provides the XOR of the inputs  $a$  and  $b$ . Furthermore, the CNOT gate is reversible since from the output  $(a', b')$  we may infer the input  $(a, b)$ . Note that, if we set the target bit to 0, the CNOT gate becomes the FANOUT gate:  $(a, 0) \rightarrow (a, a)$ . It

is easy to check that CNOT is self-inverse. Indeed, the application of two CNOT gates, one after the other, leads to

$$(a, b) \rightarrow (a, a \oplus b) \rightarrow (a, a \oplus (a \oplus b)) = (a, b). \quad (1.48)$$

Therefore,  $(\text{CNOT})^2 = I$ , that is,  $\text{CNOT}^{-1} = \text{CNOT}$ .

An interesting consequence of the fact that the FANOUT gate can be obtained from the CNOT is that a measurement process is, in principle, possible without energy expenditure. Indeed, a measurement establishes a correlation between the state of a system and the state of a memory register. Therefore, it is equivalent to a copying operation (FANOUT), which can be performed reversibly. This argument shows that the solution of Maxwell's demon paradox does not indeed lie in the measurements performed by the demon.

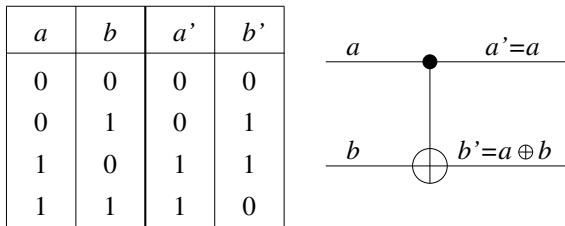


Fig. 1.25 The truth table and circuit representation for the CNOT gate.

### 1.6.1 Toffoli and Fredkin gates

It is possible to show that two-bit reversible gates are not enough for universal computation (see, *e.g.*, Preskill, 1998a). Instead, a universal gate is the controlled-controlled-NOT (CCNOT) or Toffoli gate, which is a three-bit gate. Its truth table and circuit representation are shown in Fig. 1.26. This gate acts as follows: the two control bits are unchanged ( $a' = a$  and  $b' = b$ ) while the target bit is flipped if and only if the two control bits are set to 1, that is,  $c' = c \oplus ab$ . In order to prove that the Toffoli gate is universal, we shall use it to construct both NAND and FANOUT gates. Indeed, if we set  $a = 1$ , the Toffoli gate acts on the other two bits as a CNOT, and we have seen that the FANOUT gate can be constructed from the CNOT. To construct the NAND gate, we set  $c = 1$ , so that  $c' = 0$  if and only if  $a = 1$  and  $b = 1$ , that is,  $c' = 1 \oplus ab = a \text{ NAND } b$

**Exercise 1.4** Construct the NOT, AND, OR gates from the Toffoli gate.

Another universal reversible gate is the Fredkin or controlled-EXCHANGE gate, whose truth table and circuit representation are shown in Fig. 1.27. This gate swaps the input bits  $b$  and  $c$  if and only if the control bit  $a$  is set to 1. As is easy to check, both the Toffoli and Fredkin gates are self-inverse.

**Exercise 1.5** Show that the Fredkin gate is universal.

$a$	$b$	$c$	$a'$	$b'$	$c'$
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

Fig. 1.26 The truth table and circuit representation for the Toffoli gate.

$a$	$b$	$c$	$a'$	$b'$	$c'$
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	1	0
1	1	0	1	0	1
1	1	1	1	1	1

Fig. 1.27 The truth table and circuit representation for the Fredkin gate.

We have seen that irreversible gates, such as AND and OR, can be embedded into reversible gates. However, the price to pay is the introduction of additional bits and on output this produces “garbage” bits, which are not reused during the computation. These extra bits are needed to store the information that would allow us to reverse the operations. For instance, if we set  $c = 1$  at the input of the Toffoli gate, we obtain  $c' = a \text{NAND } b$  plus two garbage bits ( $a' = a$  and  $b' = b$ ). One might think that energy is required to erase this garbage, thus nullifying the advantage of reversible computation. Fortunately, as was shown by Bennett, this is not the case. Indeed, we can perform the required computation, print the result and then run the computation backward, again using reversible gates, to recover the initial state of the computer. As a consequence, the garbage bits return to their original state without any energy consumption.

### 1.6.2 \* The billiard-ball computer

A concrete example of reversible computation is the billiard-ball computer. In this computer the value taken by a bit is associated with the absence (0) or the presence (1) of a ball in a given position. The transmission of information is performed by

means of frictionless motion of the balls on a plane surface (a billiard table) and the logic gates are implemented by means of elastic collisions between balls and against fixed obstacles. The positions of the balls on the left- and right-hand sides of the billiard table give the input and output, respectively. As an example, let us consider the collision gate depicted in Fig. 1.28. In this figure, on input we have a ball in  $a$  and another in  $b$ , that is,  $a = b = 1$ . After the collision, the balls are recovered in  $a'$  and  $b'$ ; that is,  $a' = b' = 1$ , while  $a'' = b'' = 0$ . If instead we have zero or a single ball, then there are no collisions. For instance, if  $a = 1$  and  $b = 0$ , then  $a'' = 1$  and  $a' = b' = b'' = 0$ . Therefore, the collision gate computes the following logical functions:  $a' = b' = a \wedge b$ ,  $a'' = a \wedge \bar{b}$  and  $b'' = \bar{a} \wedge b$ . To implement a universal reversible computation, we also need the possibility to change the direction of the balls. This is obtained by means of elastic bounces off a fixed obstacle (see Fig. 1.29).

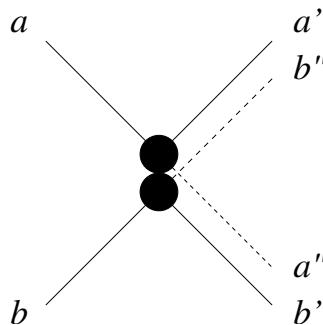


Fig. 1.28 A collision gate in a billiard-ball computer.

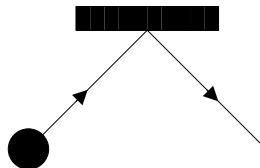


Fig. 1.29 Change of direction of a ball in a billiard-ball computer.

It is possible to combine the basic elements depicted in Figs. 1.28 and 1.29 to implement the Fredkin gate (see, for example, Feynman, 1982). Since the Fredkin gate is universal, it turns out that the billiard-ball computer may be used to compute any complex Boolean function. However, it is also clear that the actual implementation of such a computer is hindered by the instability of the motion of the balls under perturbations. Finally, it is interesting to note that the billiard-ball computer, besides being reversible, is also conservative. This means that the number of balls on input is equal to the number of balls on output; that is, the number of 0's and 1's is conserved. Actually, this is one of the properties of the Fredkin gate.

## 1.7 \* Energy dissipation in computation

The aim of this paragraph is not to describe the latest experiments, but to show that now it is possible to realize fundamental experiments on energy and information, almost unthinkable few years ago.

### 1.7.1 \* Experimental realization of a Maxwell's demon

Let us start by shortly describing the single electron transistor (SET). It is a device working at very low temperature (about 0.1 K), composed of tunnel junctions (the tunnel effect will be described in Chap. 10) and is equivalent to an electrometer so sensitive to detect reliably the charge of a single electron.

The system realizing a Maxwell's demon (Koski *et al.*, 2014) is composed of two metallic islands (of dimensions much less than a micron) separated by an insulator so thin (about a nanometer) that electrons can tunnel. Working at a temperature in the range of 0.1 K, it is possible to have the system of the two islands populated by a single electron (neglecting experimental details not essential for an intuitive understanding). So the system is surprisingly similar to the gedanken one described in Sec. 1.5. Here the two islands correspond (left and right) to the partitioned box of Fig. 1.19 and the electron corresponds to a single-molecule gas.

Through a capacitor  $C_g$  the left island is connected to a gate voltage  $V_g$  able to alter the charge of the left island, but maintaining the total charge of the two islands invariant. At first the gate voltage is set to have equal probability for the electron to be in the left or right island, so the mean number of electrons in, say, the right island, is  $n_g = \frac{1}{2}$ . A SET electrometer, connected only to right island, determines if the single electron in the device is on the left or the right island, and is exactly equivalent to a demon looking at the position of a single molecule. After reading the output of the SET, it is possible to apply a voltage  $V_g$  to the gate electrode, so that the electron is captured on the island corresponding to the output of the SET measurement. As a result of this rapid feedback, the mean charge  $n_g$  changes from  $n_g = \frac{1}{2}$  to  $n_g = 0$  (if the electron is trapped in the left island) or  $n_g = 1$  (if the electron is trapped in the right island). After that, the potential  $V_g$  is slowly driven back to the original value (ideally, in a quasistatic, reversible manner), so that the electron can tunnel to the left or right by means of thermal excitations. It is during this step that heat is extracted from the environment (and could in principle be converted into useful work). The experimental results are summarized in Fig. 1.30. It appears clearly that, slowing down the protocol, we approach more and more the ideal value  $Q = k_B T \ln 2$ . Such value corresponds to the extraction of work  $L = k_B T \ln 2$  from a bit of information, as discussed in the toy model of Sec. 1.5.3. Here, the bit of information is encoded in the position of the electron in the left or in the right island.

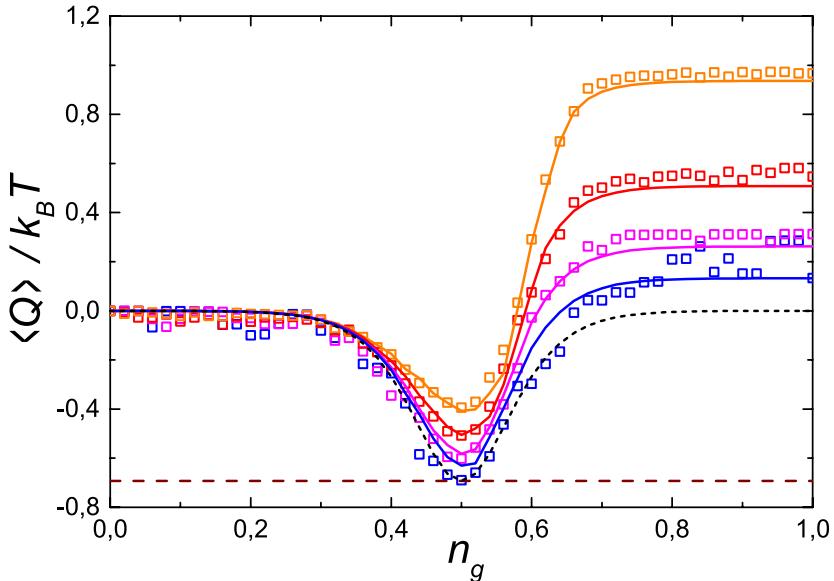


Fig. 1.30 Average total heat transferred to the environment in a ramp starting from  $n_g = 0$  to the value of  $n_g$  indicated on the  $x$  axis. Symbols (solid curves) show the measured (theoretical) values. The dashed curve gives the ideal quasistatic limit, the dashed straight line the fundamental  $-k_B T \ln 2$  limit (note the minus sign because the heat is taken negative when absorbed by the system). From top to bottom: drive rate  $\dot{n}_g = 0.22\Gamma_0, 0.11\Gamma_0, 0.055\Gamma_0, 0.027\Gamma_0$ , where  $\Gamma_0 = 22$  Hz is the tunneling rate at  $n_g = \frac{1}{2}$ . The figure is reprinted with permission from Koski *et al.* (2014).

### 1.7.2 \* Experimental verification of Landauer's principle

The experimental verification of Landauer's principle was reported by Bérut *et al.* in 2012. Such experiment demonstrates the production of heat in the erasing of a bit. The experimental system is in principle very similar to the gedanken experiment by Landauer. It consists in the realization of a single-bit memory by confining a particle (a glass bead of 2 microns diameter) in a potential well with two minima generated by means of optical tweezers. The two minima determine the 0 or 1 values of the bit. The tweezers are obtained by focusing a laser beam, and rapidly switching it between two positions to create two potential wells. By a trick equivalent to move the bead from one to the other potential well, the erasure of a bit of information is realized. That is, independently of its initial position (value 0 or 1 of the bit), the particle always ends up in the same potential well (say, value 1 of the bit). The bead is subjected to a viscous force damping its motion. Measuring the velocity of transfer from one to the other well, the force is estimated, and so the energy transferred to the bath (averaging over situations in which the bit is either initially already in 1 or is switched from 0 to 1). The results of the experiment are reported in Fig. 1.31. This figure shows clearly that Landauer's limit is achieved in the limit of long times of the erasure protocol.

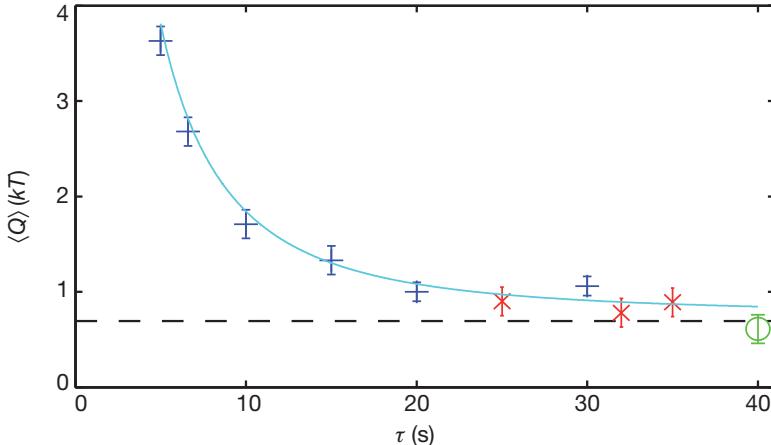


Fig. 1.31 Mean dissipated heat  $\langle Q \rangle$  for an erasure cycle as a function of the protocol duration  $\tau$ . The dashed line corresponds to the Landauer limit. The figure is reprinted with permission from Bérut *et al.* (2012). ©(2002) Macmillan Publishers Ltd.

We remark that this is an experiment of principle, and does not show a realistic implementation of an efficient memory, as the time for erasing a bit is very long and the energy necessary to create and modify the potential barrier (optical tweezers) is not taken into account. Nevertheless, this is an excellent experimental demonstration of a case of finite-time thermodynamics: the lower bound set by Landauer's principle is achieved only for quasistatic transformations, in the infinite time limit, while the dissipated heat is higher the shorter is the protocol duration  $\tau$ .

The problem of energy consumption necessary to realize the optical tweezer is common to all concrete realization of potential barriers, and also to the experimental realization of unitary gates in quantum computers. The reactive energy is the energy used to drive the system up the potential barrier. In principle, this energy could be later recovered as the system moves down the potential barrier, if the initial and final values of the potential are the same. But in real systems it is often extremely difficult to recover this energy.

### 1.7.3 \* Energy dissipation in real classical computer

To have a concrete idea of the energy dissipation problem in conventional supercomputers, it is enough to cite that (November 2017 data) the performance per watt (a measure of the energy efficiency) is 8.3 GFLOPS/W for the TOP500 list of supercomputers and 17 GFLOPS/W for the less powerful Green500 supercomputers (1 GFLOPS corresponds to  $10^9$  floating point operations per second). The present supercomputers are of the 97 petaFLOPS level and the next generation of supercomputers under active development and scheduled for 2018-2020 are at the exascale level (i.e., capable of at least one exaFLOPS, one exaFLOPS being

$10^{18}$  floating point operations per second). It is obvious that the energy dissipation problem must be urgently addressed.

It should be noted that in modern supercomputers only about 30 % of the power dissipation is due to computing. As described below, increasing the miniaturization and the speed of the processors this percentage decreases. Dissipation in interconnects and off-state leakage currents account for the remaining 70 %. In the present supercomputers the high number of operations per unit time are invariably obtained by parallelizing a huge number (millions) of elementary processing cores. In real systems this high parallelism is the origin of memory and communication problems comparable in energy cost to the processing elements. Lastly, the high energy dissipation of the numerous processors obliges to use very powerful cooling systems, reducing the overall energy efficiency of the system.

The switching energy of individual transistors for several generations of microprocessors are shown in Fig. 1.32, both for the so-called “bottom-up” and “top-down” approaches. The bottom-up data show the switching energies of individual transistors, the top-down data refer to the total power dissipation per transistor.

The switching energy per transistor scales as  $CV^2$  (see Sec. 1.7.4), where  $C$  is the capacity of the junction of the transistors (including the stray capacitances) and  $V$  is the working voltage. The capacitance  $C$  is proportional to the characteristic length  $L$  of the devices. Since also the voltage  $V$  scales as  $L$  to maintain constant the electric field, the switching energy per transistor scales as  $L^3$ . This argument explains the bottom-up line of Fig. 1.32. Comparison of the top-down and bottom-up lines show a clear divergence of the two as the linear size  $L$  shrinks. That is, the percentage of energy consumption ascribed to the switching energy of transistors lowers with shrinking the system size. This trend is driven by increased dissipation in interconnecting devices and parasitic leakage currents through thin insulators in capacitors. Note that the dissipated power density increases as the density of devices increases. At present (2018) the power density reaches about 60 – 100 W/cm<sup>2</sup>. It is difficult to increase significantly this value, due to cooling problems.

#### 1.7.4 \* Experimental realization of reversible computers

The implementation of reversible computers is a very active area of research, often referred to as a part of adiabatic computation. The standard rules to minimize the dissipation produced during the charging and discharging of capacitors are: (i) do not close a switch (i.e., turn on a transistor) if there is a voltage difference between the leads, and (ii) do not open a switch (i.e., turn off a transistor) if there is a non-zero current flowing. It turns out that the simplest way to solve these requirements is that the logic must be reversible. The circuits satisfying such requirements are referred to as adiabatic, in the sense that the dissipated energy goes to zero as the time of operation goes to infinity.

The main problem to implement a reversible computer is to charge and discharge the capacitors in the best possible way, minimizing the energy dissipation. To grasp

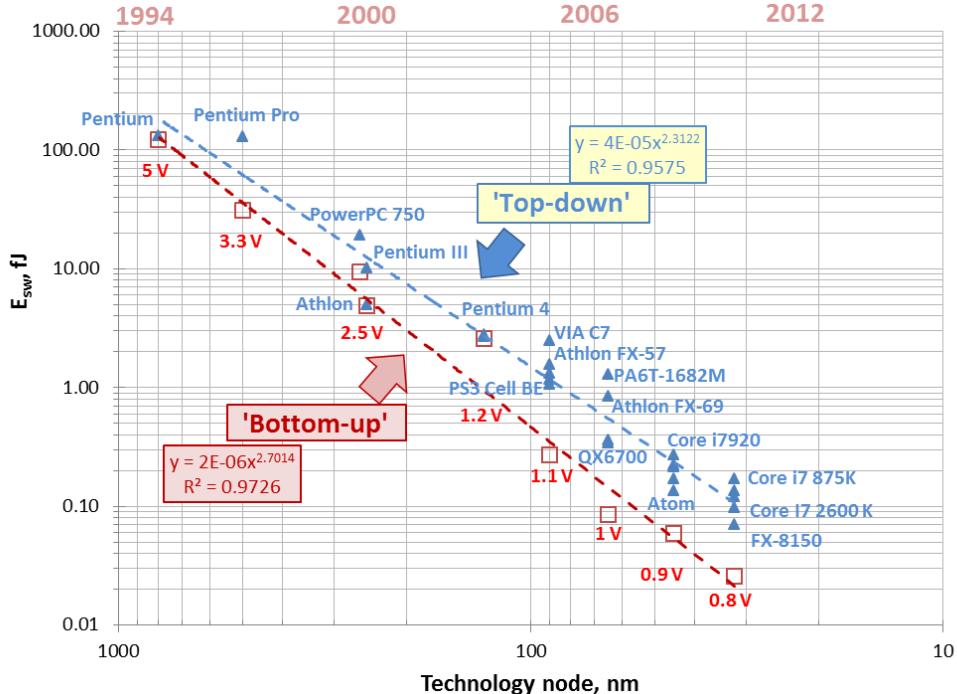


Fig. 1.32 Switching energies of individual transistors in several generations (1994–2011) of microprocessors, calculated using “bottom-up” and “top-down” approaches. The straight-line fits for the two approaches are also shown, with the determination coefficient  $R^2$  close to one and quantifying the strong correlation of data. Note, in particular, that the energy in the “bottom-up” approach scales approximately as the cubic power of the characteristic length  $L$  of the devices. The figure is reprinted from Zhirnov *et al.* (2014), CC-by-3.0 license.

in a simple way the problem, it is sufficient to model the transistor with resistors and equivalent capacitors. These capacitors describe the storage of the electric charge in the bulk of the transistor. Notwithstanding their equivalent capacitance is voltage dependent, in a first approximation it may be taken as a constant. The simplest model circuit consists of a capacitor  $C$  in series with a resistor  $R$  ( $RC$  circuit). If a constant voltage of magnitude  $V$  is applied to the circuit at time  $t = 0$ , then the charging of the capacitor is characterized by a time-dependent voltage across it:

$$V_C(t) = V [1 - \exp(-t/RC)], \quad (1.49)$$

while the voltage drop on the resistor is

$$V_R(t) = V \exp(-t/RC). \quad (1.50)$$

The total energy dissipated in the resistor is

$$E_R = \int_0^\infty \frac{V_R^2(t)}{R} dt = \frac{1}{2} CV^2. \quad (1.51)$$

Since the energy dissipated in the resistor is independent of the time constant  $RC$ , this circuit does not satisfy the adiabaticity condition. If now instead of using

a constant voltage  $V$  we use a (linear) waveform, such as to produce a constant current  $I$  in the resistor (see exercise 1.6 below), charging the capacitor from zero to a charge  $Q = CV$  in a time  $\tau$ , then the energy dissipated in the resistor is

$$E_R^{(\tau)} = RI^2\tau = R(Q/\tau)^2\tau = (RC/\tau)CV^2. \quad (1.52)$$

If  $\tau \gg RC$  the dissipation is greatly reduced as compared to the previous value, thus the capacitor should be charged with a voltage that increases linearly in time.

**Exercise 1.6** Find the time-dependent applied voltage  $V(t)$  which determines a constant current  $I$  in an  $RC$  circuit.

Note that the dissipationless limit is obtained in the limit  $\tau \rightarrow \infty$ , which is of course not of practical interest, since in that limit computations would require infinite time. It exists a direct comparison with thermodynamics, where the efficiency of the Carnot cycle is obtained for a quasi-static transformation, which requires infinite time and therefore the extracted power reduces to zero. For practical purposes, one should instead consider finite times  $\tau$ , allowing a certain amount of dissipation. This is a standard problem in finite-time thermodynamics, often referred to as a horse-carrot process. A horse is guided by waving a carrot in front of it. The problem is to move the carrot in such a way that it attracts the horse from a given point to another in a well-defined time and in an optimal way. If the distance between the horse and the carrot is too small the horse will not move very quickly; if the distance is too large, it will give up and not move at all. Here, the horse is the voltage or charge of the capacitor, the carrot is the driving voltage  $V(t)$  and the optimal process is the one that, for a given time  $\tau$ , minimizes heat dissipation.

Furthermore, to have a reversible logic the energy stored in the capacitors must be finally recovered, while discharging the capacitors. In the past years various ways were tested to satisfy this requirement of energy saving, using resonant electromagnetic circuits (and sinusoidal clocking) or mechanical resonators. All this comes with a drawback: compared to the standard dissipative logic, this reversible logic requires a factor from 8 to 16 more components. This may seem a high price to pay, but on a short time scale, it could be the only way to build energetically more efficient computers. Let us mention the pioneering realization (Wenzler *et al.*, 2013) of a Fredkin gate by means of coupled nanomechanical resonators (length of few microns). The operation speed is not sufficient for concrete applications, and the required energy, excluding losses due to resonator dissipation, is of the order of  $10^4 k_B T$  per logic operations, that is, comparable to that of the most advanced transistor-based technologies. However, this experiment represents the first demonstration of a nanomechanical universal reversible gate with a scalable technology.

### 1.7.5 \* Neuromorphic computing

In the early times of electronic digital computers it was very surprising that simple (for human beings) tasks as the recognition of images are very difficult to implement

in a standard digital computer. So, after years of developments, a very different approach emerged: neural computers that mimic as far as possible the architecture of brains. At present there is a flurry of activity by various groups to build “artificial brains” inspired by the real brain structure, with many very interesting hardware realizations of chips. The interesting thing for this chapter is that the energy dissipation of a supercomputer based on this technology may be a factor about thousand times lower than a conventional one (kW instead of MW). This is still far from the human brain’s remarkably low power consumption (few tens of Watts), but looks promising. The fundamental practical difference between a standard supercomputer and a brain-inspired computer, is that the former usually executes a fixed program, while the latter is built around “learning from examples”, to put it in a simplified way. In principle, both approaches are executable on a universal Turing machine, but the efficiency of the hardware may be very different.

## 1.8 A guide to the bibliography

The Turing machine was introduced in Turing (1936) and the Church-Turing thesis stated in Church (1936). Computing with real numbers is proposed in Siegelman (1995). The Kieu hypercomputer is reported in Kieu (2004). The undecidability of the tenth Hilbert problem is discussed in Matiyasevich (1993). A criticism to Kieu’s proposal is in Smith (2006). Hypercomputation is reviewed in Syropoulos (2008).

Classical books on algorithm design are Cormen *et al.* (2001) and Knuth (1997–1998). Textbooks on computational complexity are Garey and Johnson (1979) and Papadimitriou (1994). An informal introduction to the theory of computational complexity, directed at physicists, is Mertens (2000). The computational complexity of quantum computing is discussed in Bernstein and Vazirani (1997).

The algorithmic complexity of dynamical systems is reviewed in Alekseev and Jacobson (1981) and a very readable discussion can be found in Ford (1983).

A profound discussion of the relation between energy and information is given by Feynman (1982). Landauer’s principle was stated in Landauer (1961). Maxwell’s demon paradox has been reviewed by Bennett (1982) and Bennett (1987); for a quantum-information viewpoint, see Maruyama *et al.* (2009). Reversible computation is discussed in Bennett (1973) and Fredkin and Toffoli (1982).

A physical realization of Maxwell’s demon is reported in Koski *et al.* (2014). Landauer’s principle was experimentally tested in Bérut *et al.* (2012) and more recently in Jun *et al.* (2014). However note that this is still a debated issue in the literature, see e.g. López-Suárez *et al.* (2016) and Konopik *et al.* (2018).

Dissipation problems in real microprocessors are discussed in Zhirnov *et al.* (2014). The horse-and-carrot theorem is reported in Diósi *et al.* (1996). For a review of finite-time thermodynamics, see Andresen (2011). Two books on adiabatic logic are De Vos (2010) and Teichmann (2012). The experimental realization of a nanomechanical Fredkin gate is reported in Wenzler *et al.* (2014). For a survey of neuromorphic computing, see Schuman *et al.* (2017).

## Chapter 2

# Introduction to quantum mechanics

At the end of nineteenth century it became clear that classical physics led to predictions in disagreement with experiment. This gave rise to a profound change in the basic concepts of our understanding of Nature. A new theory, known as quantum mechanics, was constructed. This theory describes the phenomena of the microscopic world in satisfactory agreement with all present experimental data.

This chapter is an introduction to quantum mechanics. Our aim is to provide the necessary background for an understanding of the subsequent chapters. Note that no prior knowledge of quantum mechanics is required. Moreover, it is easy to learn the basic aspects of quantum mechanics. Certainly, it is much easier than one might think, considering the development of sophisticated quantum-mechanical techniques for the understanding of complex phenomena or the counter-intuitive, even paradoxical, consequences of quantum mechanics.

We begin with the description of two simple yet classic experiments: the Stern–Gerlach experiment and Young’s double-slit experiment, illustrating the distinctive features of quantum mechanics. Then we state and comment the postulates of quantum mechanics. We shall confine ourselves to the case of systems described by wave vectors residing in finite-dimensional Hilbert spaces. After that, we elucidate the unusual, non-classical properties of quantum mechanics. We discuss the EPR paradox and Bell’s inequalities, a spectacular example of the profound difference between quantum and classical physics. In the second part of the chapter, we introduce the density-matrix formalism for statistical mixtures of quantum states. As we shall see in the subsequent chapters, this will be the natural framework in which to treat open and composite quantum systems. Finally we introduce the concept of generalized measurement and discuss a simple example in which it proves to be useful.

The mathematical tools necessary for an understanding of quantum mechanics, that is, the basics of linear algebra, are reviewed in App. A.

## 2.1 The Stern–Gerlach experiment

In this section, we give a simple description of the Stern–Gerlach experiment. Perhaps, this is the experiment that illustrates most dramatically the inadequacy of classical mechanics to describe physical phenomena. It forces us to think in terms of quantum mechanics and to give up the traditional classical description of Nature. Indeed, the Stern–Gerlach experiment exhibits a typical quantum-mechanical behaviour and the predictions of classical mechanics are invalidated. It therefore shows that the basic concepts of classical mechanics must be modified in order to understand certain physical phenomena.

The Stern–Gerlach apparatus is shown schematically in Fig. 2.1: a beam of neutral atoms, having magnetic moment  $\mu$ , enters a region in which there is a magnetic field  $\mathbf{B}$  directed along the  $z$ -axis. The magnetic field is inhomogeneous, with the gradient  $\nabla B$  directed along  $z$  ( $B$  denotes the modulus of the vector  $\mathbf{B}$ ). Under these conditions, classical mechanics tells us that the atoms are subjected to a force  $\mathbf{F}$ , also directed along  $z$ . If  $F_z$  and  $\mu_z$  denote the projections of  $\mathbf{F}$  and  $\mu$  along  $z$ , we have

$$F_z = \mu_z |\nabla B| = \mu_z \frac{dB}{dz}. \quad (2.1)$$

The atoms are deflected with respect to the incoming direction by the gradient of the magnetic field and then reach a screen  $S$ . If we measure the deflection on the screen, we can derive the force  $F_z$  and therefore the component  $\mu_z$  of the magnetic moment of the atoms.

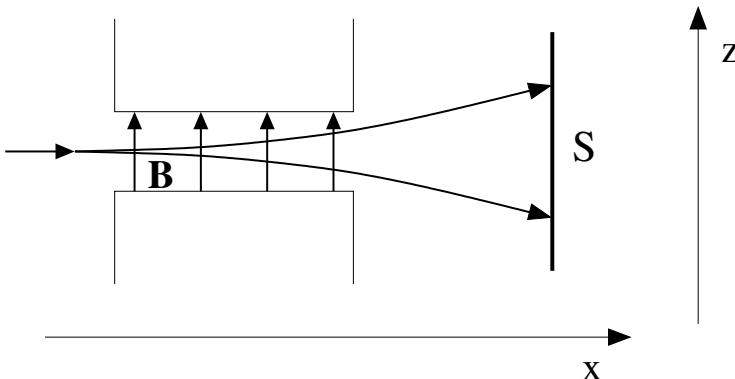


Fig. 2.1 Schematic drawing of the Stern–Gerlach experiment.

At the entrance to the region with the magnetic-field gradient, the magnetic moments of the atoms are distributed isotropically. Therefore, according to classical mechanics, all values of  $\mu_z$  between  $-m$  and  $+m$ , with  $m \equiv |\mu|$ , are allowed. As a consequence, the impact points of the atoms on the screen should be distributed continuously around the incoming direction, with maximum positive and negative

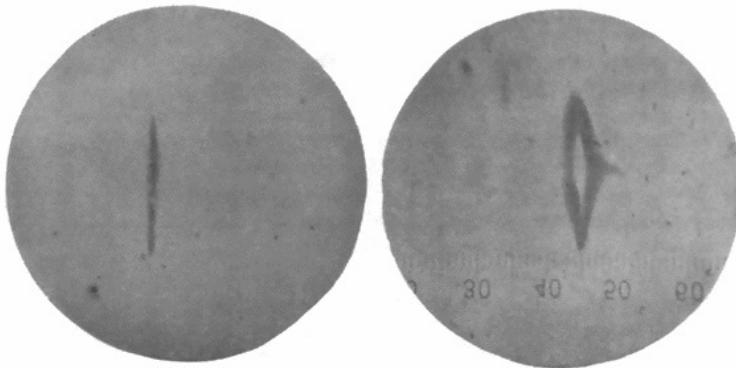


Fig. 2.2 First pictures with split and unsplit silver beams, obtained by Stern and Gerlach in 1922. Without any field there is no deflection of the beam (left image). Conversely, the quantization of the electron spin becomes clearly visible in the splitted beam obtained in the presence of a magnetic field (right image). The figure is reprinted with permission from Gerlach and Stern (1922). ©(1922) Springer-Verlag.

deflections corresponding to the values  $\mu_z = +m$  and  $\mu_z = -m$ . However, the experimental results are in clear contradiction with such predictions. Only a finite number of spots are registered on the screen. These spots are equally spaced along  $z$  and contained within the interval between the maximum positive and negative deflections corresponding to  $\mu_z = m$  and  $\mu_z = -m$ , respectively. This means that the allowed values of  $\mu_z$  are discrete. In some cases, for instance when using silver atoms, there are only two spots on the screen, corresponding to  $\mu_z = -m$  and  $\mu_z = m$  (see Fig. 2.2).

This apparently mysterious phenomenon found its explanation in the fact that electrons possess intrinsic angular momentum, known as *spin*. The magnetic moment of the atom is proportional to its angular momentum and, for atoms like silver, the angular momentum is simply equal to the spin of the outer electron. The two spots on the screen correspond to the two allowed spin states with respect to a given direction, which may be labelled *spin up* and *spin down*. The experimentally determined values of the spin angular momentum are given by  $S_z = +\frac{1}{2}\hbar$  (if the spin is up) and  $S_z = -\frac{1}{2}\hbar$  (if the spin is down), where  $\hbar = 2\pi\hbar$  is Planck's universal constant, whose value is given by  $\hbar \approx 6.626 \times 10^{-34}$  Joule s. Therefore, we say that the electron is a spin- $\frac{1}{2}$  particle. Note that the  $z$  direction in the Stern–Gerlach experiment is arbitrary; the same results are obtained if the magnetic field is oriented along any direction.

We now consider the experiment drawn schematically in Fig. 2.3. The first apparatus splits the initial beam into two components, corresponding to the spin-up and spin-down states of the electrons. Using a notation whose meaning will become clear later in this chapter, we call these two components  $|+\rangle_z$  and  $|-\rangle_z$ . Then we block the component  $|-\rangle_z$  and let the other component  $|+\rangle_z$  enter a second Stern–Gerlach apparatus analogous to the first one, namely, with the magnetic field

oriented along the  $z$  axis. A single beam, corresponding to the component  $|+\rangle_z$ , is observed to come out of the second apparatus. The  $|-\rangle_z$  component is not present, since it has been previously cut off. In short, the first Stern–Gerlach apparatus filters the atoms and only those with  $\mu_z = m$  are selected. Thus, a second Stern–Gerlach apparatus measures this same spin component.

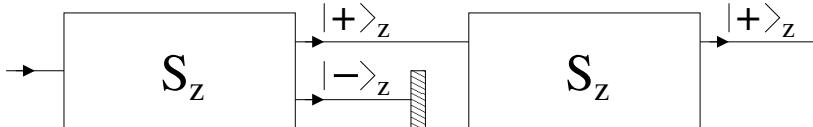


Fig. 2.3 Sketch of a Stern–Gerlach experiment. The first apparatus filters out the atoms with  $\mu_z = -m$  while the second measures  $\mu_z$ , obtaining  $\mu_z = m$ .

A different arrangement of the Stern–Gerlach experiment is illustrated in Fig. 2.4. Unlike the case of Fig. 2.3, the magnetic field in the second apparatus is directed along the  $y$  axis. Now we observe that two beams with equal intensity emerge from the second apparatus, corresponding to  $\mu_y = +m$  and  $\mu_y = -m$ , where  $\mu_y$  denotes the projection of the magnetic moment of the atoms along  $y$ . We call these two components  $|+\rangle_y$  and  $|-\rangle_y$ . This result is not surprising since the atoms enter the second Stern–Gerlach apparatus with a well-determined  $\mu_z = m$ , but the value of  $\mu_y$  is not given. Is it therefore correct to say that half of the atoms that enter the second apparatus have components  $|+\rangle_z$  and  $|+\rangle_y$  and the other half have components  $|+\rangle_z$  and  $|-\rangle_y$ ? As illustrated by the following experiment, this intuitive way of thinking is *not* valid.

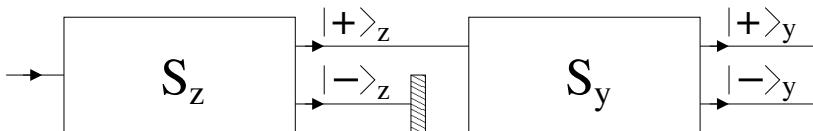


Fig. 2.4 Sketch of a Stern–Gerlach experiment. The first apparatus filters out atoms with  $\mu_z = -m$  while the second measures  $\mu_y$ .

Indeed, a very surprising result is obtained using the experimental setup drawn in Fig. 2.5. In this case, the first two Stern–Gerlach apparatuses filter out the atoms with magnetic-moment components  $\mu_z = -m$  and  $\mu_y = -m$ . The amazing result is that both components  $|+\rangle_z$  and  $|-\rangle_z$  come out with equal intensity from the third apparatus, even though the component  $|-\rangle_z$  was previously filtered out. How then is it possible that the component  $|-\rangle_z$  reappears after the third apparatus? Where does this component come from? This experiment shows that it is not correct to think that the atoms entering the third apparatus are in the state  $|+\rangle_z$  and  $|+\rangle_y$ . It is also extremely puzzling that, if the plate screening the  $|-\rangle_y$  component is removed, so that both components  $|+\rangle_y$  and  $|-\rangle_y$  enter the third apparatus, then only the component  $|+\rangle_z$  comes out. The experiment drawn in Fig. 2.5 is an impressive

illustration of a fundamental property of quantum mechanics: the final state of the system depends only on the state of the atoms that enter the last Stern–Gerlach apparatus and on the action of this apparatus; there is no memory of the previous history of the system. In short, the second apparatus singles out the state  $|+\rangle_y$  and in so doing completely destroys any information about the value of  $S_z$ . In the case of Fig. 2.5 the atoms that enter the last apparatus satisfy  $\mu_y = m$ , but there are no restrictions on  $\mu_z$ . Therefore, some of the atoms come out of the apparatus with  $\mu_z = +m$ , others with  $\mu_z = -m$ .

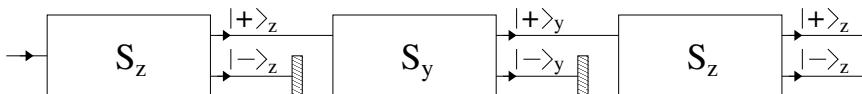


Fig. 2.5 Sketch of a Stern–Gerlach experiment. The first apparatus filters out the atoms with  $\mu_z = -m$ , the second those with  $\mu_y = -m$ , the third measures  $\mu_z$ .

## 2.2 Young's double-slit experiment

Another experiment which effectively illustrates the distinctive features of quantum mechanics is Young's double-slit experiment. As is well known, the nature of light has been the focus of deep debate over the centuries. The question was if light is a beam of particles or a wave. Newton supported the particle conception of light, since he believed that the presence of a sharp shadow behind an object could not be explained if light were a wave. Therefore, he concluded that light is not a wave like sound, which can be heard even behind objects. However, during the nineteenth century, the wave-like nature of light was demonstrated in interference experiments. Young formulated the *superposition principle*: if two waves, emitted by a single source, fall on a screen, their *amplitudes* (not their intensities, which are square moduli of the amplitudes) add up algebraically. This property is at the origin of the well-known interference fringes observed in the double-slit experiment illustrated in Fig. 2.6. Light is emitted by the source  $S$ , passes through two slits,  $O_1$  and  $O_2$ , and illuminates a screen (a photographic plate, for example). Typical interference fringes are produced on the screen, as sketched in Fig. 2.6. The main point is that the intensity  $I(x)$  of light on the screen is different from the algebraic sum of the intensities  $I_1(x)$  and  $I_2(x)$  produced by the two slits separately, that is, when we close slit  $O_2$  or  $O_1$ , respectively. Thus, we have

$$I(x) \neq I_1(x) + I_2(x). \quad (2.2)$$

In the second half of the nineteenth century, after the synthesis performed by Maxwell, it became clear that light is an electromagnetic wave. In this framework, the speed of light ( $c \approx 2.998 \times 10^8$  m/s in vacuum) is the propagation velocity of the electromagnetic field and is related to certain electric and magnetic constants. However, the energy distribution of the radiation emitted by a black body could

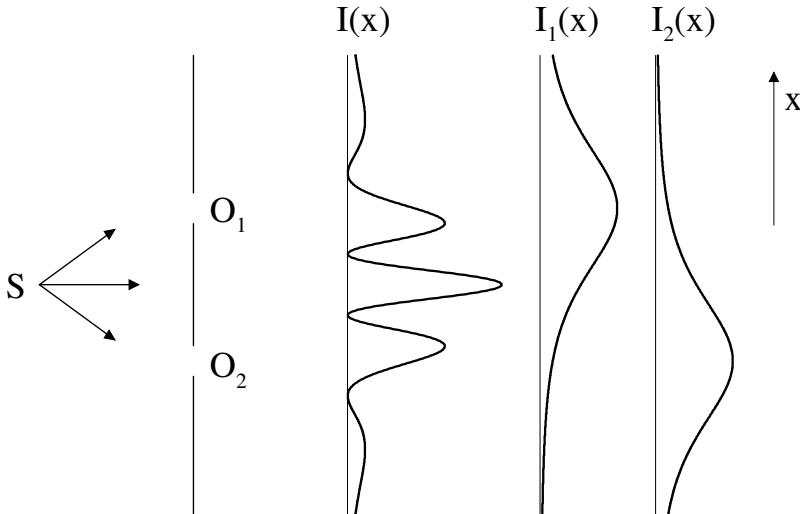


Fig. 2.6 Schematic diagram of Young's double-slit interference experiment. A source  $S$  emits light, which can pass through two slits  $O_1$  and  $O_2$ , before striking a screen. The pattern  $I_1(x)$  is produced when only slit  $O_1$  is open and pattern  $I_2(x)$  is obtained when only  $O_2$  is open. The intensity  $I(x)$  of light on the screen when both slits are open is different from the algebraic sum of the intensities  $I_1(x)$  and  $I_2(x)$ .

not be explained by the electromagnetic theory. This problem led Planck in 1900 to introduce the hypothesis that light is emitted or absorbed only in integer multiples of a basic *quantum of energy*

$$E = h\nu, \quad (2.3)$$

where  $\nu$  is the frequency of light and  $h$  is Planck's constant. Einstein (1905), in his theory of the photoelectric effect, returned to the particle theory: light consists of a beam of particles, now called *photons* (the word photon was not coined by Einstein), each possessing energy  $h\nu$ . From the study of light-matter interaction, one concludes that such particles have momentum  $p = h\nu/c$  (Einstein, 1917). Both Maxwell's theory, which describes light as an electromagnetic wave and Einstein's theory, which describes light as a beam of elementary particles called photons, received broad experimental confirmation. This again raised the question as to the particle or wave-like nature of light.

The crucial relevance of Young's double-slit experiment lies in the fact that a complete description can be obtained only by accepting simultaneously *both* the wave and particle aspects of light. The light intensity  $I(x)$  at a point  $x$  on the screen is proportional to the squared modulus of the electric field  $E(x)$  at the same point. Let us denote by  $E_1(x)$  and  $E_2(x)$  the electric field produced at  $x$  by beams passing through slits  $O_1$  and  $O_2$ , respectively. The corresponding light intensities are given by

$$I_1(x) \propto |E_1(x)|^2, \quad (2.4a)$$

which is observed when the second slit is closed, and

$$I_2(x) \propto |E_2(x)|^2, \quad (2.4b)$$

obtained when the first slit is closed. If we open both slits, the field strengths add up algebraically,

$$E(x) = E_1(x) + E_2(x), \quad (2.5)$$

and therefore the resulting light intensity is given by

$$I(x) \propto |E(x)|^2 = |E_1(x) + E_2(x)|^2, \quad (2.6)$$

and thus we find

$$I(x) \neq I_1(x) + I_2(x). \quad (2.7)$$

This is in agreement with the predictions of the wave theory of light.

What happens though if the light intensity is reduced, so that the source only emits photons one-by-one? In this case, each photon produces a *localized* impact at some point on the screen. If we expose the photographic plate for a time so short that only a few photons strike the screen, we observe a few localized impact points, but not an interference pattern. Therefore, a particle interpretation rather than a wave interpretation of light explains this experimental result.<sup>1</sup> Indeed, if we consider a wave, when its intensity diminishes, the interference fringes diminish in intensity but do not disappear. Thus, this prediction of the wave theory of light is invalidated by experimental results.

Nevertheless, if the exposure time is sufficiently long that the photographic plate can capture many photons, the interference fringes do appear. The light intensity collected at any given point is proportional to the density of photon impacts at this point. Therefore, the photons, as they arrive, build up the interference fringes and these fringes can be explained by a wave interpretation instead of a particle interpretation of light. In summary, we cannot explain the whole of the experimental results using either the predictions of the particle theory alone or those of the wave theory alone. Both the particle and wave-like nature of light are present.

Here it is important to note that, by placing a photon counter behind one of the two slits, it is indeed possible to determine through which slit each photon passes, but in so doing the interference fringes are destroyed. The crucial point is that it is physically impossible to both observe the interference pattern and determine through which slit each photon passes.

These results force us to revise some of the fundamental concepts of classical physics. The fact that interference fringes are observed if and *only* if we do not know through which slit each photon passes forces us to give up the concept of a *trajectory*. Indeed, each photon produced by the source strikes the screen at a different position. We cannot predict this position, but only give the probability  $p(x)$

---

<sup>1</sup>Note, however, that the arrival points are not as predicted by classical mechanics, but are distributed probabilistically according to the fringe-pattern intensity.

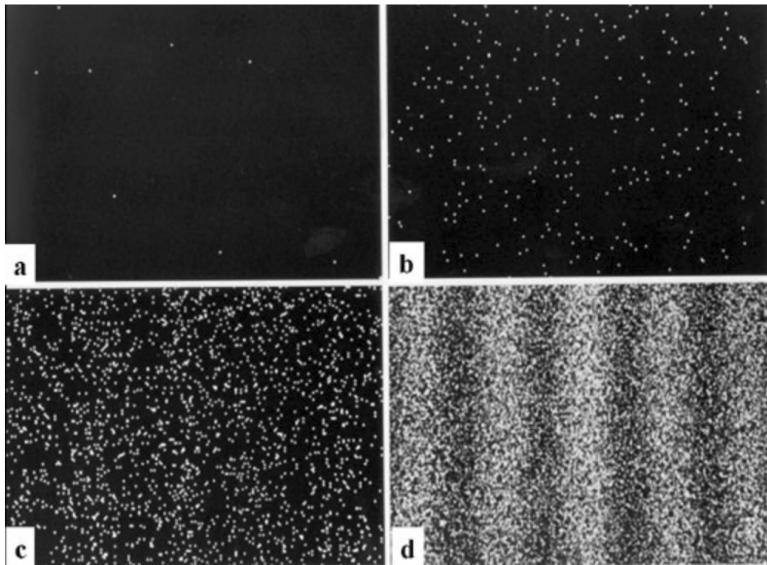


Fig. 2.7 Progressive formation of the interference pattern of an electron hologram, as obtained in the field-emission electron microscope experiment by Tonomura and coworkers (Tonomura *et al.*, 1989). Bright spots, each of them signaling the detection of a single electron, appear first at random positions (panels a and b). After the accumulation of a larger number of electrons, regular fringes start to become visible in the vertical direction (panel c). Further increasing the number of detected electrons enables to distinguish clear interference fringes (panel d). Reprinted courtesy of the Central Research Laboratory, Hitachi, Ltd., Japan.

that a photon will strike the screen at the point  $x$ . This probability is proportional to the intensity  $I(x)$ , namely, to  $|E(x)|^2$ . Even though all photons are emitted under the same conditions, we cannot know in advance where each photon will strike the screen. Therefore, we must reject the classical concept that, once the initial conditions and external forces acting on a particle are known, we can follow, at least in principle, the evolution in time of the particle's coordinates.

We must also give up the concept that the particle and wave descriptions are mutually exclusive, since both aspects are necessary to explain the experimental results. Therefore, we are inevitably led to the concept of *wave-particle duality*: light behaves simultaneously both as a wave and as a particle. We stress that, as has been shown by an overwhelming number of experiments, such duality is not restricted to the description of optical phenomena. Indeed, material particles may also exhibit wave properties and, *vice versa*, waves can be associated with particles.

Let us finally emphasize that this radical revision of the concepts of classical physics was imposed and guided by experimental results. For instance, an experimental verification of Young's experiment for electrons has been realized by Merli *et al.* (1974), and then repeated fifteen years later, by means of a coherent electron beam illuminating an object to be imaged (see Tonomura *et al.*, 1989). The electron beam intensity is sufficiently low that the chance to have more than one electron at

a time in the interference microscope is negligible. As a result, the pattern of the reconstructed hologram shown in Fig. 2.7 is formed as a consequence of the accumulation in time of single-electron quantum interference. This unveils a genuinely single-particle quantum phenomenon.

## 2.3 The postulates of quantum mechanics

In classical mechanics the state of a system of  $n$  particles at time  $t_0$  is determined by the positions  $\{\mathbf{x}_1(t_0), \mathbf{x}_2(t_0), \dots, \mathbf{x}_n(t_0)\}$  and the velocities  $\{\dot{\mathbf{x}}_1(t_0), \dot{\mathbf{x}}_2(t_0), \dots, \dot{\mathbf{x}}_n(t_0)\}$  of all the particles at this time. If these initial conditions are known, Newton's laws allow, at least in principle, to compute the state of the system at any time  $t$ . Indeed, the laws of classical mechanics lead to first-order ordinary differential equations in the variables  $x_i$  and  $\dot{x}_i$  and, once the initial conditions are set, there exists a unique solution  $\{\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_n(t); \dot{\mathbf{x}}_1(t), \dot{\mathbf{x}}_2(t), \dots, \dot{\mathbf{x}}_n(t)\}$ .

Quantum mechanics is based on a completely different mathematical framework. In the following, we shall introduce the postulates that are at the basis of quantum theory.

### 2.3.1 Dynamical evolution

We start by setting the basic formalism and answer the following questions: What is a quantum systems? How can the system's state at any time be formally represented? What are the equivalent of the classical variables and measures?

We shall see that most of the fundamental properties of quantum mechanics, such as the superposition principle, the interference between two states, and the uncertainty principle, follow directly from the underlying mathematical structure that is associated to any physical quantum system by a series of axiomatic paradigms.

**Postulate I:** *Each physical system  $\mathcal{S}$  is associated with a Hilbert space  $\mathcal{H}_{\mathcal{S}}$  on the complex field. A given state of the system is completely described by a unit vector  $|\psi\rangle$ , which is called the state vector, or wave function, on the Hilbert space  $\mathcal{H}_{\mathcal{S}}$ .<sup>2</sup>*

*The time evolution of the state vector  $|\psi\rangle$  is governed by the Schrödinger equation*

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle, \quad (2.8)$$

*where  $H(t)$  is a self-adjoint operator known as the Hamiltonian of the system and  $\hbar \equiv h/2\pi$ , with the physical constant  $h$  known as Planck's constant. Its value ( $h \approx 6.626 \times 10^{-34}$  Joule s) is determined experimentally.*

---

<sup>2</sup>For the sake of simplicity, here we shall consider only finite-dimensional Hilbert spaces (discrete variables) and postpone the discussion of infinite-dimensional Hilbert spaces (continuous variables) to Chap. 5.

It is important to note that the Schrödinger equation is a linear differential equation of first order in time. Therefore, given the initial state  $|\psi(t_0)\rangle$ , the state  $|\psi(t)\rangle$  at any time  $t$  is completely and uniquely determined by the solution to Eq. (2.8).

Since the Schrödinger equation is linear, the following *superposition principle* applies: if  $|\psi_1(t)\rangle$  and  $|\psi_2(t)\rangle$  are solutions of Eq. (2.8), then the superposition  $|\psi(t)\rangle = \alpha|\psi_1(t)\rangle + \beta|\psi_2(t)\rangle$ , where  $\alpha$  and  $\beta$  are complex numbers, is also a solution. Therefore, the *time-evolution operator*  $U$ , defined by

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle \quad (2.9)$$

is linear. The solution to the Schrödinger equation (2.8) can be written as

$$|\psi(t)\rangle = \mathcal{T} \exp\left[-\frac{i}{\hbar} \int_{t_0}^t d\tau H(\tau)\right] |\psi(t_0)\rangle, \quad (2.10)$$

where  $\mathcal{T}$  denotes the time-ordering operator, that is, for any time-dependent operators  $A(t_1)$  and  $B(t_2)$ , we have

$$\mathcal{T}[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2) & \text{if } t_1 > t_2, \\ B(t_2)A(t_1) & \text{if } t_1 < t_2. \end{cases} \quad (2.11)$$

In the case of a time-independent Hamiltonian  $H$ , the solution reduces to

$$|\psi(t)\rangle = \exp\left[-\frac{i}{\hbar} H(t - t_0)\right] |\psi(t_0)\rangle. \quad (2.12)$$

It can be shown that the time-evolution operator  $U$  is unitary. This is immediate to prove for Eq. (2.12).

**Exercise 2.1** *Stone's theorem for finite-dimensional Hilbert spaces.* Show that any unitary operator  $U$  can be written as  $U = \exp(iA)$ , where  $A$  is a Hermitian operator.

### 2.3.2 Outcomes of a measurement

**Postulate II:** *Any physical observable  $A$  is associated with a self-adjoint operator  $A$  on the Hilbert space  $\mathcal{H}_S$ . The possible outcome of a measurement of the observable  $A$  is one of the eigenvalues of the operator  $A$ . Writing the eigenvalue equation,*

$$A|i\rangle = a_i|i\rangle, \quad (2.13)$$

where  $|i\rangle$  is an orthonormal basis of eigenvectors of the operator  $A$ , and expanding the state vector  $|\psi(t)\rangle$  over this basis:

$$|\psi(t)\rangle = \sum_i c_i(t)|i\rangle, \quad (2.14)$$

then the probability that a measurement of the observable  $A$  at time  $t$  results in outcome  $a_i$  is given by

$$p_i(t) = |\langle i|\psi(t)\rangle|^2 = |c_i(t)|^2. \quad (2.15)$$

For the sake of simplicity, we have stated Postulate II for the case in which the eigenvalues of  $A$  are non-degenerate. We shall consider the case of spectral degeneracies later, before stating Postulate III.

### Comments

- (i) *Observables* are the quantum analogue of dynamical variables in classical mechanics, such as position, linear and angular momentum and so on. In contrast, other characteristics of a system, such as mass or electric charge, are not in the class of observables, but enter as parameters in the Hamiltonian of the system.
- (ii) The following argument should help grasp the reason for which self-adjoint operators are associated with physical observables: the eigenvalues of a self-adjoint operator are real (just as the possible outcomes of a measurement) and its eigenvectors form a complete orthonormal set in the Hilbert space  $\mathcal{H}_S$  associated with the system. Since  $|\psi(t)\rangle$  has unit norm, we have

$$\sum_i p_i(t) = \sum_i |c_i(t)|^2 = 1, \quad (2.16)$$

and therefore the probabilities are normalized, that is, the total probability of obtaining an outcome from the measurement of the observable  $A$  is equal to 1. It is exactly for this reason that Postulate I requires  $|\psi(t)\rangle$  to have unit norm.

- (iii) In the particular case in which the state vector  $|\psi(t_0)\rangle$  at a given time  $t_0$  coincides with an eigenvector of the operator  $A$  with eigenvalue  $a_i$ ,

$$|\psi(t_0)\rangle = |i\rangle, \quad (2.17)$$

then a measurement of the observable  $A$  at time  $t_0$  gives, with unit probability, outcome  $a_i$ . Eigenvectors of the operator  $A$  are also called *eigenstates* of  $A$ .

- (iv) Let us assume that  $|\psi_1\rangle$  and  $|\psi_2\rangle$  are two distinct, normalized eigenvectors of the operator  $A$ , with eigenvalues  $a_1$  and  $a_2$ , respectively. The superposition principle tells us that the state

$$|\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle, \quad (2.18)$$

with  $\lambda_1$  and  $\lambda_2$  complex numbers, is also an allowed state of the system, provided that  $|\lambda_1|^2 + |\lambda_2|^2 = 1$ , so that  $|\psi\rangle$  has unit norm. Therefore, if the system is described by the state vector  $|\psi\rangle$  and we perform a measurement of the observable  $A$ , we obtain outcome  $a_1$  with probability  $|\lambda_1|^2$  and outcome  $a_2$  with probability  $|\lambda_2|^2$ . However, the superposition state  $|\psi\rangle$  is not equivalent to a *naïve* statistical mixture of the states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ , taken with probabilities  $p_1 = |\lambda_1|^2$  and  $p_2 = |\lambda_2|^2$ , respectively. We will introduce and discuss more formally the statistical mixtures later, in Sec. 2.6. Here we only show that a large number  $N$  of systems, all in the same state  $|\psi\rangle$ , is not equivalent to an ensemble of  $|\lambda_1|^2 N$  systems in the state  $|\psi_1\rangle$  and  $|\lambda_2|^2 N$  systems in the state  $|\psi_2\rangle$ . Indeed, let us assume that we wish to compute the probability  $p(b_i)$  of obtaining outcome  $b_i$  for a measurement of some observable  $B$ , given that the system is described by the state  $|\psi\rangle$ . According to Postulate II, we have

$$p(b_i) = |\langle i|\psi\rangle|^2, \quad (2.19)$$

where  $|i\rangle$  is an eigenvector of the operator  $B$  (associated with the observable  $B$ ) with eigenvalue  $b_i$ . Thus, we obtain

$$\begin{aligned} p(b_i) &= |\lambda_1\langle i|\psi_1\rangle + \lambda_2\langle i|\psi_2\rangle|^2 \\ &= |\lambda_1|^2|\langle i|\psi_1\rangle|^2 + |\lambda_2|^2|\langle i|\psi_2\rangle|^2 + 2 \operatorname{Re}\{\lambda_1\lambda_2^*\langle i|\psi_1\rangle\langle i|\psi_2\rangle^*\}. \end{aligned} \quad (2.20)$$

A different result is obtained if we consider a statistical mixture of the states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ , taken with probabilities  $|\lambda_1|^2$  and  $|\lambda_2|^2$ . In such case, the probability  $p_{\text{mix}}(b_i)$  of obtaining outcome  $b_i$  for the observable  $B$  is given by

$$p_{\text{mix}}(b_i) = |\lambda_1|^2|\langle i|\psi_1\rangle|^2 + |\lambda_2|^2|\langle i|\psi_2\rangle|^2, \quad (2.21)$$

and therefore

$$p(b_i) = p_{\text{mix}}(b_i) + 2 \operatorname{Re}\{\lambda_1\lambda_2^*\langle i|\psi_1\rangle\langle i|\psi_2\rangle^*\}. \quad (2.22)$$

The last term in Eq. (2.22) is called an *interference term*. Therefore, the probability of obtaining  $b_i$  as the outcome of a measurement of  $B$ , and more generally the predictions of the quantum-mechanical theory, depend not only on the moduli  $|\lambda_1|$  and  $|\lambda_2|$  but also on the relative phase between the complex numbers  $\lambda_1$  and  $\lambda_2$ , which affects the product  $\lambda_1\lambda_2^*$ . For example, the four states

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), & |\psi_2\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \\ |\psi_3\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), & |\psi_4\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \end{aligned} \quad (2.23)$$

represent different states of a system, leading to different experimental outcomes. In contrast, a global phase has no physical significance, that is, the normalized state vectors  $|\psi\rangle$  and  $e^{i\varphi}|\psi\rangle$ , with  $\varphi$  real, give the same predictions for the outcome of any experiment.

### Conservative systems

When the Hamiltonian  $H$  of a system does not depend explicitly on time, we say that the system is conservative. In this case, we know from classical mechanics that the energy  $E$  of the system is constant in time; that is, it is a *constant of motion*. In quantum mechanics, the solution to the Schrödinger equation (2.8) can be written easily, once we know the eigenvalues  $E_n$  and eigenvectors  $|n\rangle$  of the Hamiltonian operator  $H$ . Let us consider the eigenvalue equation for  $H$ :

$$H|n\rangle = E_n|n\rangle, \quad (2.24)$$

where, for the sake of simplicity, we assume that the spectrum of the operator  $H$  is non-degenerate, that is,  $E_m \neq E_n$  for  $m \neq n$ . Since we have assumed that  $H$  does not depend on time, the eigenvalues  $E_n$  and eigenvectors  $|n\rangle$  are also time-independent. The solution  $|\psi(t)\rangle$  to the Schrödinger equation (2.8) can be expanded over the basis of the eigenfunctions of the operator  $H$  as follows:

$$|\psi(t)\rangle = \sum_n c_n(t)|n\rangle, \quad (2.25)$$

with

$$c_n(t) = \langle n | \psi(t) \rangle. \quad (2.26)$$

The solution  $|\psi(t)\rangle$  is uniquely determined by the initial condition  $|\psi(t_0)\rangle$ , where we may take  $t_0 = 0$ . The initial condition  $|\psi(0)\rangle$  is determined if the coefficients  $c_n(0) = \langle n | \psi(0) \rangle$  are fixed. As a matter of fact, if we insert expansion (2.25) into the Schrödinger equation (2.8), we obtain

$$i\hbar \frac{d}{dt} c_n = E_n c_n, \quad (2.27)$$

whose solution is

$$c_n(t) = c_n(0) \exp\left(-\frac{i}{\hbar} E_n t\right). \quad (2.28)$$

Therefore, the state vector  $|\psi(t)\rangle$  at time  $t$  is given by

$$|\psi(t)\rangle = \sum_n c_n(0) \exp\left(-\frac{i}{\hbar} E_n t\right) |n\rangle. \quad (2.29)$$

In the special case in which  $|\psi(0)\rangle$  coincides with an eigenvector  $|n\rangle$  of the Hamiltonian operator  $H$ ,  $|\psi(0)\rangle = |n\rangle$ , due to the orthogonality between the different eigenvectors, the solution (2.29) to the Schrödinger equation reduces to

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar} E_n t\right) |n\rangle. \quad (2.30)$$

Therefore, the state vectors  $|\psi(0)\rangle$  and  $|\psi(t)\rangle$  only differ by a global phase factor of no physical significance. For this reason the eigenstates of a time-independent Hamiltonian  $H$  are called *stationary states*: if a system is described by such a state, its physical properties do not change in time.

### 2.3.3 The post-measurement state

We now turn to the effect of the measurement process on the state of the system. Let us assume that the measurement of an observable  $A$  results in outcome  $a_n$ , with  $a_n$  a non-degenerate eigenvalue of the self-adjoint operator  $A$ . If the measurement does not destroy the system and a new measurement of the observable  $A$  immediately follows, we again obtain outcome  $a_n$  with unit probability. We can explain this experimental result if we admit that the wave function of the system, which immediately before the first measurement was in the state  $|\psi\rangle$ , immediately after the measurement collapses onto the eigenstate  $|n\rangle$  of  $A$  associated with the eigenvalue  $a_n$ . In the case in which there is degeneracy, we can expand the state  $|\psi\rangle$  before the measurement as follows:

$$|\psi\rangle = \sum_n \sum_{s=1}^{g_n} c_{n_s} |n_s\rangle, \quad (2.31)$$

where  $g_n$  measures the order of degeneracy of the eigenvalue  $a_n$ , that is, the dimension of the subspace spanned by the eigenvectors of  $A$  with the same eigenvalue  $a_n$ .

After a measurement giving outcome  $a_n$ , the state of the system belongs to this subspace and is given by

$$\frac{1}{\sqrt{\sum_{s=1}^{g_n} |c_{n_s}|^2}} \sum_{s=1}^{g_n} c_{n_s} |n_s\rangle. \quad (2.32)$$

This state is the normalized projection of  $|\psi\rangle$  over the subspace corresponding to the eigenvalue  $a_n$  (*i.e.*, spanned by the eigenvectors of  $A$  with eigenvalue  $a_n$ ). We may now state the following

**Postulate III:** *If a system is described by the wave vector  $|\psi\rangle$  and we measure an observable  $A$ , obtaining the outcome  $a_n$ , then immediately after the measurement the state of the system is given by*

$$\frac{P_n |\psi\rangle}{\sqrt{\langle \psi | P_n | \psi \rangle}}, \quad (2.33)$$

where  $P_n$  is the projection operator over the subspace corresponding to  $a_n$ .

If the wave vector  $|\psi\rangle$  is given by Eq. (2.31), then the projector  $P_n$  reads

$$P_n = \sum_{s=1}^{g_n} |n_s\rangle \langle n_s|. \quad (2.34)$$

Since the eigenvectors of  $A$  constitute an orthonormal basis for the Hilbert space  $\mathcal{H}_S$  associated with the system, it is easy to check that the projectors  $P_n$  satisfy the completeness relation,

$$\sum_n P_n = I, \quad (2.35)$$

and the orthogonality condition

$$P_n P_m = \delta_{mn} P_m. \quad (2.36)$$

In the case without degeneracy,  $g_n = 1$ , the wave function of the system after the measurement collapses onto the state

$$\frac{1}{|c_n|} c_n |n\rangle, \quad (2.37)$$

and therefore (neglecting a global phase factor of no physical significance) onto the eigenstate  $|n\rangle$  corresponding to the eigenvalue  $a_n$ .

If the system is described by the wave vector (2.31), then upon measuring the observable  $A$ , the probability of obtaining any given outcome  $a_n$  is given by

$$p_n = \langle \psi | P_n | \psi \rangle. \quad (2.38)$$

It is easy to check that for  $p_n$  we recover the statement of Postulate II, namely, Eq. (2.15), in the non-degenerate case ( $g_n = 1$ ).

Probability theory now tells us that the average value of the observable  $A$  is given by

$$\langle A \rangle = \sum_n a_n p_n, \quad (2.39)$$

and therefore

$$\langle A \rangle = \sum_n a_n \langle \psi | P_n | \psi \rangle = \langle \psi | \left( \sum_n a_n P_n \right) | \psi \rangle = \langle \psi | A | \psi \rangle, \quad (2.40)$$

where we have used the spectral decomposition  $A = \sum_n a_n P_n$ .

The standard deviation  $\Delta A$  associated with observations of  $A$  is given by

$$\Delta A = \sqrt{\langle (A - \langle A \rangle)^2 \rangle} = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \quad (2.41)$$

Therefore, if we perform a large number of experiments in which the state  $|\psi\rangle$  is prepared and the observable  $A$  is measured, we obtain outcomes with mean value  $\langle A \rangle$  and standard deviation  $\Delta A$ .

### 2.3.4 Heisenberg's uncertainty principle

Starting from the analysis of a few ideal experiments, Heisenberg showed that it is not possible to simultaneously assign a well-determined position and velocity to a given particle. If we increase the precision in our measurement of the particle's velocity, then we increase the uncertainty in its position and *vice versa*. This intrinsically quantum limitation is expressed by the position–momentum uncertainty relations of Heisenberg:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \quad \Delta y \Delta p_y \geq \frac{\hbar}{2}, \quad \Delta z \Delta p_z \geq \frac{\hbar}{2}, \quad (2.42)$$

where  $\Delta x$ ,  $\Delta y$ ,  $\Delta z$  and  $\Delta p_x$ ,  $\Delta p_y$ ,  $\Delta p_z$  are the uncertainties in the position and the momentum of the particle. In the following we give the precise mathematical formulation of Heisenberg's uncertainty principle, due to Robertson (1929).

**Theorem 2.1** *Suppose that  $A$  and  $B$  are Hermitian operators associated with two observables and  $|\psi\rangle$  is a given quantum state. Then the following inequality is satisfied:*

$$\Delta A \Delta B \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}. \quad (2.43)$$

**Proof.** Let us consider the operators  $P$  and  $Q$ , defined by  $P = A - \langle A \rangle$  and  $Q = B - \langle B \rangle$ . We can always write the complex number  $\langle \psi | PQ | \psi \rangle$  as equal to  $a + ib$ , with  $a$  and  $b$  real numbers. Thus, the average values of the commutator  $[P, Q]$  and the anti-commutator  $\{P, Q\}$  are given by  $\langle \psi | [P, Q] | \psi \rangle = 2ib$  and  $\langle \psi | \{P, Q\} | \psi \rangle = 2a$ . This implies that

$$\begin{aligned} |\langle \psi | [P, Q] | \psi \rangle|^2 &\leq |\langle \psi | [P, Q] | \psi \rangle|^2 + |\langle \psi | \{P, Q\} | \psi \rangle|^2 = 4(a^2 + b^2) \\ &= 4|\langle \psi | PQ | \psi \rangle|^2 \leq 4\langle \psi | P^2 | \psi \rangle \langle \psi | Q^2 | \psi \rangle, \end{aligned} \quad (2.44)$$

where the last inequality is the Cauchy–Schwartz inequality proved in App. A.1. Finally, we consider the first and the last term in Eq. (2.44). Since  $\langle [P, Q] \rangle = \langle [A, B] \rangle$ ,  $\langle P^2 \rangle = (\Delta A)^2$  and  $\langle Q^2 \rangle = (\Delta B)^2$ , we have proved the Heisenberg inequality (2.43).  $\square$

The Heisenberg principle tells us that, given two non-commuting observables  $A$  and  $B$ , there is an intrinsic limit to the accuracy of the simultaneous measurement of both  $A$  and  $B$ . The measurement of one observable necessarily disturbs the other. For instance, if the system is prepared in an eigenstate of  $A$  associated with a well-determined eigenvalue  $a_i$ , a measurement of the observable  $A$  always results in outcome  $a_i$ . However, if we measure  $B$ , the system wave vector collapses onto an eigenstate of  $B$ , which is no longer an eigenstate of  $A$ , if  $A$  and  $B$  do not commute. Therefore, if we now measure  $A$  again, we obtain different outcomes, with probabilities determined by Postulate II. In quantum mechanics the measurement process disturbs the system: if the observable  $A$  is measured to some accuracy  $\Delta A$ , the observable  $B$  is disturbed by some amount  $\Delta B$  and  $\Delta A \Delta B$  satisfies the Heisenberg inequality (2.43). Given two non-commuting observables  $A$  and  $B$ , it is impossible to measure at the same time both  $A$  and  $B$  to an arbitrary degree of accuracy: increasing the accuracy in  $A$  implies that the accuracy in  $B$  diminishes and *vice versa*. There is no similar phenomenon in classical mechanics and we shall see in Chap. 5 that this intrinsically quantum-mechanical result finds applications in the field of cryptography.

**Exercise 2.2** Assume that the observables  $\sigma_x$  and  $\sigma_y$  are measured when a system is in the state  $|0\rangle$ , where  $|0\rangle$  denotes the eigenstate of  $\sigma_z$  corresponding to the eigenvalue +1. Show that the uncertainty principle implies that  $\Delta\sigma_x\Delta\sigma_y \geq 1$ .

The Stern–Gerlach experiment, described in Sec. 2.1, is an example of state measurement/preparation. If the apparatus is oriented along the  $z$ -axis, we obtain one out of two possible states,  $|0\rangle$  or  $|1\rangle$ . These states are the eigenvectors of the Pauli operator  $\sigma_z$ , corresponding to the eigenvalues +1 and -1 (Pauli operators are defined in App. A.1). If we block the state  $|1\rangle$ , then we are left with the eigenstate  $|0\rangle$  of the spin operator  $\sigma_z$  (see Fig. 2.3). If instead the apparatus is oriented along the  $x$ -axis, we obtain one out of two possible states, which are the eigenvectors of the Pauli operator  $\sigma_x$ , that is,  $|+\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and  $|-\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ , corresponding to the eigenvalues +1 and -1, respectively.

Note that the most general state of the system can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (2.45)$$

with  $|\alpha|^2 + |\beta|^2 = 1$ . If we introduce the spherical polar angles  $\theta$  and  $\phi$  (see Fig. 2.8), this state can equivalently be written as

$$|\psi\rangle = \cos \frac{\theta}{2} e^{-i\phi/2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi/2} |1\rangle, \quad (2.46)$$

with  $0 \leq \theta \leq \pi$  and  $0 \leq \phi < 2\pi$ . Such a state is obtained, in the Stern–Gerlach experiment, when the apparatus is directed along the axis singled out by the unit vector  $\mathbf{u} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ . Indeed, the state  $|\psi\rangle$  is an eigenstate of the operator

$$\sigma_{\mathbf{u}} = \boldsymbol{\sigma} \cdot \mathbf{u} = \sigma_x \sin \theta \cos \phi + \sigma_y \sin \theta \sin \phi + \sigma_z \cos \theta, \quad (2.47)$$

where  $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ . The operator  $\sigma_u$  has the following matrix representation in the basis of the eigenvectors of  $\sigma_z$ :

$$\sigma_u = \begin{bmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{bmatrix}. \quad (2.48)$$

It is easy to check that the matrix  $\sigma_u$  has eigenvectors

$$\begin{aligned} |+\rangle_u &= \cos \frac{\theta}{2} e^{-i\phi/2} |0\rangle + \sin \frac{\theta}{2} e^{i\phi/2} |1\rangle, \\ |-\rangle_u &= -\sin \frac{\theta}{2} e^{-i\phi/2} |0\rangle + \cos \frac{\theta}{2} e^{i\phi/2} |1\rangle, \end{aligned} \quad (2.49)$$

corresponding to the eigenvalues +1 and -1, respectively.

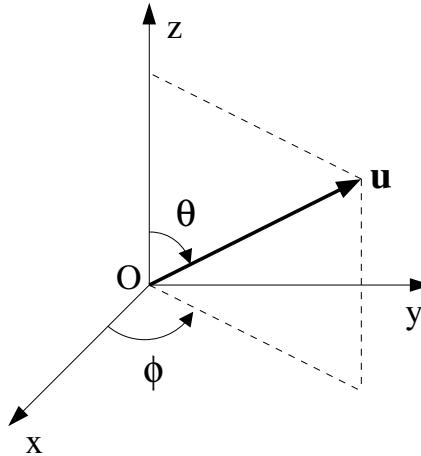


Fig. 2.8 Definition of the spherical polar coordinates  $\theta$  and  $\phi$  characterizing a unit vector  $\mathbf{u}$ .

The Stern–Gerlach apparatus can be used both to prepare a state and to measure a state. In the first case we say that the Stern–Gerlach apparatus is used as a *polarizer*, in the latter as an *analyzer*. Let us assume that a beam of atoms of spin- $\frac{1}{2}$  enters a Stern–Gerlach apparatus oriented along the  $x$  axis. As we saw in Sec. 2.1, the two components  $|+\rangle_x$  and  $|-\rangle_x$  come out of the apparatus. If we block component  $|-\rangle_x$ , then we can say that we have prepared the state  $|+\rangle_x$  and in this case the apparatus has been used as a polarizer. If a beam of atoms enters the apparatus oriented, for example, along  $z$ , we measure the value of  $\sigma_z$  and the apparatus acts as an analyzer. If the incoming state is described, for instance, by  $|\psi\rangle = |+\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , then the system is in a superposition of the two eigenstates  $|0\rangle$  and  $|1\rangle$  of  $\sigma_z$  and we can easily predict from Postulate II that the measurement of  $\sigma_z$  will give the eigenvalues +1 or -1 of  $\sigma_z$  with equal probabilities  $p_+ = p_- = \frac{1}{2}$ . Indeed, we have

$$p_+ = |\langle 0|\psi\rangle|^2 = \langle\psi|P_0|\psi\rangle = \frac{1}{2}, \quad p_- = |\langle 1|\psi\rangle|^2 = \langle\psi|P_1|\psi\rangle = \frac{1}{2}, \quad (2.50)$$

where

$$\begin{aligned} P_0 &= |0\rangle\langle 0| = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \\ P_1 &= |1\rangle\langle 1| = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \quad (2.51)$$

are the projection operators onto the subspaces spanned by the vectors  $|0\rangle$  and  $|1\rangle$ . It is easy to check that

$$P_0 + P_1 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P_0 P_1 = 0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (2.52)$$

and therefore the projectors  $P_0$  and  $P_1$  satisfy both the completeness relation (2.35) and the orthogonality condition (2.36).

**Exercise 2.3** Show that the results of the Stern–Gerlach experiment illustrated in Fig. 2.5 are in agreement with the predictions of quantum mechanics.

**Exercise 2.4** The Schrödinger equation describing the time evolution of the wave vector associated with a spin-half particle of magnetic moment  $\mu$  in a magnetic field  $\mathbf{H} = (H_x, H_y, H_z)$  is given by

$$i\hbar \frac{d}{dt} \begin{bmatrix} a(t) \\ b(t) \end{bmatrix} = -\mu(H_x\sigma_x + H_y\sigma_y + H_z\sigma_z) \begin{bmatrix} a(t) \\ b(t) \end{bmatrix}. \quad (2.53)$$

Solve this equation and compute the mean values of the Pauli operators as a function of time. If the initial wave vector is given by the  $\sigma_z$  eigenstate  $|0\rangle$ , what magnetic field and evolution time are required to evolve it into the other  $\sigma_z$  eigenstate  $|1\rangle$ ?

## 2.4 The EPR paradox

The most spectacular and counter-intuitive manifestation of quantum mechanics is the phenomenon of *entanglement*, observed in composite quantum systems. Let us now discuss the problem. The Hilbert space  $\mathcal{H}$  associated with a composite system is the tensor product of the Hilbert spaces  $\mathcal{H}_i$  associated with the system's components  $i$ . In the simplest case of a bipartite quantum system, we have

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \quad (2.54)$$

The most natural basis for the Hilbert space  $\mathcal{H}$  is constructed from the tensor products of the basis vectors of  $\mathcal{H}_1$  and  $\mathcal{H}_2$ . If, for example, the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are two-dimensional and

$$\{|0\rangle_1, |1\rangle_1\}, \quad \{|0\rangle_2, |1\rangle_2\} \quad (2.55)$$

denote their basis vectors, then a basis for the Hilbert space  $\mathcal{H}$  is given by the four vectors

$$\{|0\rangle_1 \otimes |0\rangle_2, |0\rangle_1 \otimes |1\rangle_2, |1\rangle_1 \otimes |0\rangle_2, |1\rangle_1 \otimes |1\rangle_2\}. \quad (2.56)$$

The superposition principle tells us that the most general state in the Hilbert space  $\mathcal{H}$  is not a tensor product of states residing in  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , but an arbitrary superposition of such states, which we can write as follows:

$$|\psi\rangle = \sum_{i=0}^1 \sum_{j=0}^1 c_{ij} |i\rangle_1 \otimes |j\rangle_2. \quad (2.57)$$

In order to simplify notation, we may also write

$$|\psi\rangle = \sum_{i,j} c_{ij} |ij\rangle, \quad (2.58)$$

where the first index in  $|ij\rangle$  refers to a state residing in the Hilbert space  $\mathcal{H}_1$  and the second to a state in  $\mathcal{H}_2$ . By definition, a state in  $\mathcal{H}$  is said to be *entangled*, or *non-separable*, if it cannot be written as a simple tensor product of a state  $|\alpha\rangle_1$  belonging to  $\mathcal{H}_1$  and a state  $|\beta\rangle_2$  belonging to  $\mathcal{H}_2$ . In contrast, if we can write

$$|\psi\rangle = |\alpha\rangle_1 \otimes |\beta\rangle_2, \quad (2.59)$$

we say that the state  $|\psi\rangle$  is *separable*. As simple examples, let us consider the state

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \quad (2.60)$$

which is entangled, and the state

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |11\rangle), \quad (2.61)$$

which is separable, since we can write

$$|\psi_2\rangle = tfrac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \otimes |1\rangle. \quad (2.62)$$

**Exercise 2.5** Show that the state (2.60) is entangled.

When two systems are entangled, it is not possible to assign them individual state vectors  $|\alpha\rangle_1$  and  $|\beta\rangle_2$ . The intriguing non-classical properties of entangled states were clearly illustrated by Einstein, Podolsky and Rosen (EPR) in 1935. These authors showed that quantum theory leads to a contradiction, provided that we accept the following two, seemingly natural, assumptions:

- (i) *Reality principle*: If we can predict with certainty the value of a physical quantity, then this value has physical reality, independently of our observation. For example, if a system's wave function  $|\psi\rangle$  is an eigenstate of an operator  $A$ , namely,  $A|\psi\rangle = a|\psi\rangle$ , then the value  $a$  of the observable  $A$  is an element of physical reality
- (ii) *Locality principle*: If two systems are causally disconnected, the result of any measurement performed on one system cannot influence the result of a measurement performed on the second system. Following the theory of special relativity, we say that two measurement events are disconnected if  $(\Delta x)^2 > c^2(\Delta t)^2$ , where  $\Delta x$  and  $\Delta t$  are the space and time separations of the two events in some inertial reference frame and  $c$  is the speed of light [the two events take place at space-time coordinates  $(x_1, t_1)$  and  $(x_2, t_2)$ , respectively, and  $\Delta x = x_2 - x_1$ ,  $\Delta t = t_2 - t_1$ ].

In quantum mechanics, if an operator  $B$  does not commute with  $A$ , then the two physical quantities corresponding to the operators  $A$  and  $B$  cannot have simultaneous reality since we cannot predict with certainty the outcome of the simultaneous measurement of both  $A$  and  $B$ . Following Heisenberg's principle, a measurement of  $A$  destroys knowledge of  $B$ .

### Examples

Let us illustrate the EPR paradox by means of two simple examples. The first one is due to Bohm. Consider a source  $S$  that emits a pair of spin- $\frac{1}{2}$  particles in the entangled state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (2.63)$$

This state is called an EPR or Bell state. We also say that the system is in a spin-singlet state. One spin- $\frac{1}{2}$  particle is sent to an observer called Alice and the second to another observer called Bob (see Fig. 2.9). Note that Alice and Bob may be located arbitrarily far away from each other. The only requirement is that the measurements performed by Alice and Bob be causally disconnected.

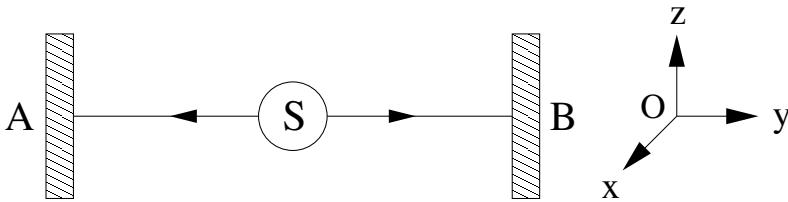


Fig. 2.9 Schematic drawing of the EPR *gedanken* experiment.

If Alice measures the  $z$  component of the spin of the particle in her possession and obtains, for instance,  $\sigma_z^{(A)} = +1$ , then the EPR state collapses onto the state  $|01\rangle$  (we remind the reader that the states  $|0\rangle$  and  $|1\rangle$  are eigenstates of  $\sigma_z$ , corresponding to the eigenvalues  $+1$  and  $-1$ , respectively). Subsequently, if Bob measures the  $z$  component of the spin for his particle, he will obtain  $\sigma_z^{(B)} = -1$  with unit probability. Therefore, the results of the measurements of Alice and Bob are perfectly anticorrelated. This result is not surprising according to our intuition, since it is easy to find analogous classical situations. As an example, let us consider two balls, one black and the other white. One ball is sent to Alice and the other to Bob. If Alice finds that her ball is black, then Bob will find with certainty that his ball is white. The surprising point comes from the observation that the spin-singlet state (2.63) can be also written as

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |--\rangle), \quad (2.64)$$

where  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and  $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  are eigenstates of  $\sigma_x$  with eigenvalues  $+1$  and  $-1$ , respectively. If Alice measures  $\sigma_x^{(A)}$  and obtains, for example, the

outcome  $\sigma_x^{(A)} = +1$ , then the EPR state collapses onto  $|+-\rangle$  and Bob will obtain with certainty from the measurement of  $\sigma_x^{(B)}$  the outcome  $\sigma_x^{(B)} = -1$ . Therefore, the state of one particle depends upon the nature of the observable measured on the other particle. If Alice measures  $\sigma_z^{(A)}$ , then the state of Bob's particle collapses onto an eigenstate of  $\sigma_z^{(B)}$ . In contrast, if Alice measures  $\sigma_x^{(A)}$ , then the state of Bob's particle collapses onto an eigenstate of  $\sigma_x^{(B)}$ . Using the EPR language, we say that in the first case we associate an element of physical reality with  $\sigma_z^{(B)}$ , in the latter with  $\sigma_x^{(B)}$ . It is impossible to assign simultaneous physical reality to both observables since they do not commute,  $[\sigma_x^{(B)}, \sigma_z^{(B)}] \neq 0$ . The main point is that Alice can choose which observable to measure even after the particles have separated. Therefore, according to the locality principle, any measurement performed by Alice cannot modify the state of Bob's particle. Thus, quantum theory leads to a contradiction if we accept the principles both of realism and locality described above.

A more intuitive explanation of the EPR paradox is provided by the following example. Let us take a stamp and cut it into two parts, which are subsequently put in two separate envelopes to be sent to the distant parties Alice and Bob. When Alice opens the envelope and looks at her half-stamp, she immediately knows the features of the complementary part, that has been sent to Bob.

Using a classical reasoning, the argument appears to be correct and does not present any interpretation problem. However in the quantum realm it is incomplete. Indeed, according to quantum theory, Alice could have chosen different and incompatible bases (*i.e.*, corresponding to non-commuting operators) along which to measure. For example, in a system of two spin- $\frac{1}{2}$  particles, she could have considered the  $\sigma_x^{(A)}$  and  $\sigma_z^{(A)}$  measures of the previous example. Forcing a classical interpretation, one could naively think that Alice cuts the stamp along a certain direction (see Fig. 2.10), after the envelopes with the two half-stamps have been

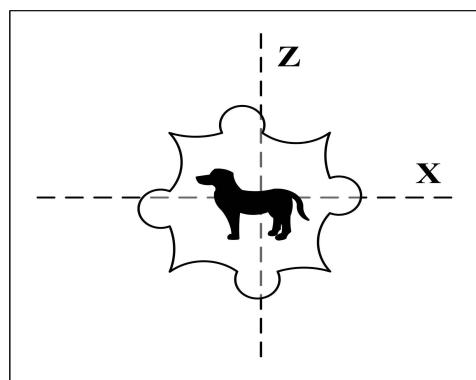


Fig. 2.10 The “quantum cut” of a stamp, according to the interpretation of the EPR paradox. Drawing courtesy of Sara Felloni.

sent. From a quantum point of view, this corresponds to choosing the measurement direction after letting the two parts of the system (*i.e.*, the two half-stamps) separate. In summary, Alice could have chosen how to cut the stamp, *after* the two envelopes have been sent to her and Bob. Forcing this classical interpretation would lead to one of the two following paradoxes:

- (i) Alice does not have free will: her decision is predetermined by the cut that has been performed before sending the two envelopes;
- (ii) Alice does have free will: her decision influences the past action of cutting the stamp.

Let us eventually stress that the EPR paradox does not violate the relativity principles. The outcome of Bob's measurement, after Alice's one, is known if and only if they both choose the same instrument (*i.e.*, they measure along the same basis, either  $\sigma_x$  or  $\sigma_z$ ). A classical communication channel between Alice and Bob is needed to make them agree on the type of measurement. This fact prevents from any faster-than-light (instantaneous) communication between the two parties.

The two above examples put forward the fundamental difference between classical and quantum mechanics, in that, from a classical point of view, the system possesses an outcome of a measure already before the measurement itself. On the other hand, quantum mechanically a completely new perspective has to be adopted: it is not possible to assess that a quantum system has a given property before having measured it. The EPR conclusion to this paradox was that quantum mechanics is an incomplete theory. It was later proposed quantum theory to be complete by introducing the so-called *hidden variables*. The suggestion was that any measurement is in reality a deterministic process, which merely appears probabilistic since some degrees of freedom (hidden variables) are not precisely known.

Let us stress that the standard interpretation of quantum mechanics does not accept Einstein's local realism. The wave function is not seen as a physical object, but just as a mathematical tool, which is useful to predict probabilities for the outcome of experiments.

**Exercise 2.6** Prove that the spin-singlet state (2.63) is rotationally-invariant, that is, that it takes the same form

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{\mathbf{u}}|-\rangle_{\mathbf{u}} - |-\rangle_{\mathbf{u}}|+\rangle_{\mathbf{u}}) \quad (2.65)$$

for any direction  $\mathbf{u}$ , the states  $|+\rangle_{\mathbf{u}}$  and  $|-\rangle_{\mathbf{u}}$  being eigenstates of  $\boldsymbol{\sigma} \cdot \mathbf{u}$ .

This result is actually rather obvious *a priori*: a spin-singlet state corresponds to zero total spin; thus, no spin vector and no preferred direction can be associated with such a state.

**Exercise 2.7** Consider the composite system of a pair of spin- $\frac{1}{2}$  particles, described by the wave function

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle \quad (|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1). \quad (2.66)$$

Assume that the spin polarization  $\sigma_z$  (or  $\sigma_x$ ) is measured for the first particle. Discuss the effect of the measurement on the system's wave function.

## 2.5 Bell's inequalities

The debate on the physical reality of quantum systems became the subject of experimental investigation after the formulation, in 1964, of Bell's inequalities. These inequalities are obtained assuming the principles of realism and locality. Since it is possible to devise situations in which quantum mechanics predicts a violation of these inequalities, any experimental observation of such a violation excludes the possibility of a local and realistic description of natural phenomena. In short, Bell showed that the principles of realism and locality lead to experimentally testable inequality relations in disagreement with the predictions of quantum mechanics.

### First derivation

It is instructive to derive Bell's inequalities following a simple model proposed by Bohm. Here we follow the presentation of Sakurai (1994). We assume that a source emits a large number of spin pairs in the singlet state (2.63). Alice and Bob each receive a member of each pair and can measure its polarization along any of three axes  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$ . We divide the particles in groups as follows. If Alice obtains, for instance, outcome +1 when she measures  $\sigma_a^{(A)}$ , +1 when she measures  $\sigma_b^{(A)}$  and -1 when she measures  $\sigma_c^{(A)}$ , then we say that the particle belongs to group  $(\mathbf{a}+, \mathbf{b}+, \mathbf{c}-)$ . We should stress that we are not saying that Alice measures  $\sigma_a^{(A)}$ ,  $\sigma_b^{(A)}$  and  $\sigma_c^{(A)}$  simultaneously. She may only measure any one of the spin components. For instance, if she measures  $\sigma_a^{(A)}$ , then she measures neither  $\sigma_b^{(A)}$  nor  $\sigma_c^{(A)}$ . However, according to the reality principle, we may assign well-defined values to the spin components along the three axes, that is, we assume that these values have physical reality *independently of our observation*. Now remember that the results of Alice's and Bob's measurements must be perfectly anticorrelated for a spin-singlet state. Thus, if Alice's particle belongs to group  $(\mathbf{a}+, \mathbf{b}+, \mathbf{c}-)$ , then Bob's particle has to be in group  $(\mathbf{a}-, \mathbf{b}-, \mathbf{c}+)$ . The eight mutually exclusive possibilities are shown in Table 2.1.

Let  $p(\mathbf{a}+, \mathbf{b}+)$  denote the probability that Alice obtains  $\sigma_a^{(A)} = +1$  and Bob obtains  $\sigma_b^{(B)} = +1$ . It is clearly seen from Table 2.1 that

$$p(\mathbf{a}+, \mathbf{b}+) = \frac{N_3 + N_4}{N_t}, \quad (2.67)$$

where  $N_t \equiv \sum_{i=1}^8 N_i$ . Similarly, we obtain

$$p(\mathbf{a}+, \mathbf{c}+) = \frac{N_2 + N_4}{N_t}, \quad p(\mathbf{c}+, \mathbf{b}+) = \frac{N_3 + N_7}{N_t}. \quad (2.68)$$

Since  $N_i \geq 0$ , we have  $N_3 + N_4 \leq (N_2 + N_4) + (N_3 + N_7)$  and therefore we obtain the following Bell inequality:

$$p(\mathbf{a}+, \mathbf{b}+) \leq p(\mathbf{a}+, \mathbf{c}+) + p(\mathbf{c}+, \mathbf{b}+). \quad (2.69)$$

Table 2.1 Division of the spin-singlet states into mutually exclusive groups.

Population	Alice's particle	Bob's particle
$N_1$	$(\mathbf{a}+, \mathbf{b}+, \mathbf{c}+)$	$(\mathbf{a}-, \mathbf{b}-, \mathbf{c}-)$
$N_2$	$(\mathbf{a}+, \mathbf{b}+, \mathbf{c}-)$	$(\mathbf{a}-, \mathbf{b}-, \mathbf{c}+)$
$N_3$	$(\mathbf{a}+, \mathbf{b}-, \mathbf{c}+)$	$(\mathbf{a}-, \mathbf{b}+, \mathbf{c}-)$
$N_4$	$(\mathbf{a}+, \mathbf{b}-, \mathbf{c}-)$	$(\mathbf{a}-, \mathbf{b}+, \mathbf{c}+)$
$N_5$	$(\mathbf{a}-, \mathbf{b}+, \mathbf{c}+)$	$(\mathbf{a}+, \mathbf{b}-, \mathbf{c}-)$
$N_6$	$(\mathbf{a}-, \mathbf{b}+, \mathbf{c}-)$	$(\mathbf{a}+, \mathbf{b}-, \mathbf{c}+)$
$N_7$	$(\mathbf{a}-, \mathbf{b}-, \mathbf{c}+)$	$(\mathbf{a}+, \mathbf{b}+, \mathbf{c}-)$
$N_8$	$(\mathbf{a}-, \mathbf{b}-, \mathbf{c}-)$	$(\mathbf{a}+, \mathbf{b}+, \mathbf{c}+)$

We point out that we have assumed the locality principle to derive this inequality. Indeed, if a pair belongs to group 1 and Alice chooses to measure  $\sigma_{\mathbf{a}}^{(A)}$ , then she will certainly obtain outcome 1, independently of the fact that Bob might choose to perform a measurement along the axes  $\mathbf{a}$ ,  $\mathbf{b}$  or  $\mathbf{c}$ .

We now evaluate the probabilities appearing in Bell's inequality (2.69) following quantum theory. Let us consider  $p(\mathbf{a}+, \mathbf{b}+)$ . If Alice finds  $\sigma_{\mathbf{a}}^{(A)} = +1$ , then the state of Bob's particle collapses onto the eigenstate  $|\rightarrow\rangle_{\mathbf{a}}$  of  $\sigma_{\mathbf{a}}^{(B)}$  with eigenvalue  $-1$ . Thus, provided that  $\sigma_{\mathbf{a}}^{(A)} = +1$ , it is easy to check that Bob obtains  $\sigma_{\mathbf{b}}^{(B)} = +1$  with probability  $|b(+|\rightarrow\rangle_{\mathbf{a}}|^2 = \sin^2(\theta_{ab}/2)$ , where  $\theta_{ab}$  is the angle between the axes  $\mathbf{a}$  and  $\mathbf{b}$ . Since Alice obtains  $\sigma_{\mathbf{a}}^{(A)} = +1$  with probability one half, we obtain

$$p(\mathbf{a}+, \mathbf{b}+) = \frac{1}{2} \sin^2\left(\frac{\theta_{ab}}{2}\right). \quad (2.70)$$

In the same way we can compute  $p(\mathbf{a}+, \mathbf{c}+)$  and  $p(\mathbf{c}+, \mathbf{b}+)$ . Hence, Bell's inequality (2.69) gives

$$\sin^2\left(\frac{\theta_{ab}}{2}\right) \leq \sin^2\left(\frac{\theta_{ac}}{2}\right) + \sin^2\left(\frac{\theta_{cb}}{2}\right). \quad (2.71)$$

If we choose the axes  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$  such that  $\theta_{ab} = 2\theta$ ,  $\theta_{ac} = \theta_{cb} = \theta$ , then this inequality is violated for  $0 < \theta < \frac{\pi}{2}$ . Therefore, quantum mechanics leads to an experimentally testable violation of Bell's inequalities.

### Second derivation

We now give an alternative derivation of Bell's inequalities. Let us assume that there exists a hidden variable  $\lambda$  such that, for any value of  $\lambda$ , a well-defined (deterministic) result  $O(\lambda)$  is obtained from the measurement of a physical observable  $O$ . We require that the distribution probability  $\rho(\lambda)$  of the variable  $\lambda$  be such that the average values predicted by quantum mechanics are recovered; that is,

$$\langle O \rangle = \int O(\lambda) \rho(\lambda) d\lambda. \quad (2.72)$$

Let us consider the EPR *gedanken* experiment drawn schematically in Fig. 2.9. We call  $A(\mathbf{a}, \lambda)$  and  $B(\mathbf{b}, \lambda)$  the results of the measurements of the (causally disconnected) spin polarizations  $\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}$  and  $\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}$  along the directions  $\mathbf{a}$  and  $\mathbf{b}$ ,

performed by Alice and Bob, respectively. Assuming the locality principle, the outcome of Alice's measurements cannot depend on the outcome of Bob's measurements. Therefore, the mean value of the correlations between their polarization measurements is given by

$$C(\mathbf{a}, \mathbf{b}) = \int A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda)\rho(\lambda)d\lambda. \quad (2.73)$$

For example, as we have seen above, quantum mechanics predicts perfect anticorrelation for the EPR state (2.63) when  $\mathbf{a} = \mathbf{b}$  and therefore

$$C(\mathbf{a}, \mathbf{a})_{\text{quantum}} = -1. \quad (2.74)$$

Let us compute

$$\begin{aligned} C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}') &= \int [A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda \\ &= \int A(\mathbf{a}, \lambda)B(\mathbf{b}, \lambda)[1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda \\ &\quad - \int A(\mathbf{a}, \lambda)B(\mathbf{b}', \lambda)[1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}, \lambda)]\rho(\lambda)d\lambda. \end{aligned} \quad (2.75)$$

Since  $A(\mathbf{a}, \lambda)$  and  $B(\mathbf{b}, \lambda)$  are polarization measurements, we have

$$|A(\mathbf{a}, \lambda)| = 1, \quad |B(\mathbf{b}, \lambda)| = 1. \quad (2.76)$$

Moreover,  $\rho(\lambda)$  is a probability distribution and therefore is non-negative for any  $\lambda$ . Thus, we have

$$|C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| \leq \int [1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}', \lambda)]\rho(\lambda)d\lambda + \int [1 \pm A(\mathbf{a}', \lambda)B(\mathbf{b}, \lambda)]\rho(\lambda)d\lambda. \quad (2.77)$$

This implies that

$$|C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| \leq \pm[C(\mathbf{a}', \mathbf{b}') + C(\mathbf{a}', \mathbf{b})] + 2 \int \rho(\lambda)d\lambda \quad (2.78)$$

and therefore

$$|C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| \leq -|C(\mathbf{a}', \mathbf{b}') + C(\mathbf{a}', \mathbf{b})| + 2 \int \rho(\lambda)d\lambda. \quad (2.79)$$

We finally obtain

$$|C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| + |C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}', \mathbf{b}')| \leq 2, \quad (2.80)$$

where we have used the normalization of the probability distribution  $\rho(\lambda)$ , that is,  $\int \rho(\lambda)d\lambda = 1$ . Inequality (2.80) is known as the CHSH inequality, after its four discoverers (Clauser, Horne, Shimony and Holt). It is an example of a larger set known as Bell's inequalities. The main point is that there exist directions  $(\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}')$  such that, considering entangled states, quantum mechanics violates the CHSH inequality. For instance, we may consider the set of directions  $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$  shown in Fig. 2.11. For the spin-singlet state (2.63), quantum mechanics predicts

that  $C(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b} = -\cos(\theta_{ab})$ , where  $\theta_{ab}$  is the angle between the directions  $\mathbf{a}$  and  $\mathbf{b}$  (see exercise 2.8 below), thus we have

$$\begin{aligned} & \left\{ |C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| + |C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}', \mathbf{b}')| \right\}_{\text{quantum}} \\ &= |- \cos(\phi) + \cos(3\phi)| + |- \cos(\phi) - \cos(\phi)| = 2\sqrt{2} \geq 2, \end{aligned} \quad (2.81)$$

when  $\phi = \frac{\pi}{4}$ .

**Exercise 2.8** Show that the quantum-mechanical mean value of the correlation

$$C(\mathbf{a}, \mathbf{b})_{\text{quantum}} = \langle \psi | (\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}) \otimes (\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}) | \psi \rangle \quad (2.82)$$

is equal to  $-\mathbf{a} \cdot \mathbf{b}$  when  $|\psi\rangle$  is the EPR state (2.63).

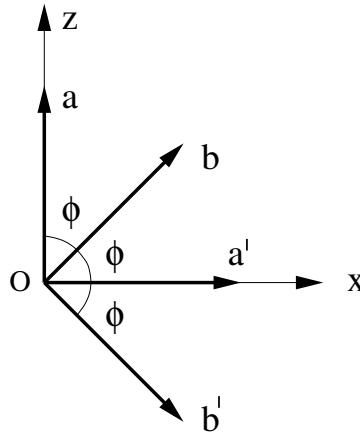


Fig. 2.11 Choice of directions leading to a violation of the CHSH inequality (2.80). The angles labelled  $\phi$  are taken equal to  $\frac{\pi}{4}$ .

Bell's inequalities represent, first of all, an experimental test of the consistency of quantum mechanics. Many experiments have been performed in order to check Bell's inequalities; the most famous involved EPR pairs of photons and was performed by Aspect and co-workers in 1981. This experiment displayed an unambiguous violation of the CHSH inequality by tens of standard deviations and an excellent agreement with quantum mechanics. More recently, other experiments have come closer to the requirements of the ideal EPR scheme and again impressive agreement with the predictions of quantum mechanics has always been found. Nonetheless, there is no general consensus as to whether or not these experiments may be considered conclusive, owing to the limited efficiency of detectors. If, for the sake of argument, we assume that the present results will not be contradicted by future experiments with high-efficiency detectors, we are led to the following *no-go* theorem (Bell's theorem): “No physical theory of local hidden variables can reproduce the predictions of quantum mechanics”. In summary, Nature does not experimentally support the EPR point of view, and the World is not locally realistic.

We should stress that there is more to learn from Bell's inequalities and Aspect's experiments than merely a consistency test of quantum mechanics. These profound results show us that entanglement is a fundamentally new resource, beyond the realm of classical physics, and that it is possible to experimentally manipulate entangled states. As we shall see in the following chapters, a major goal of quantum-information science is to exploit this resource to perform computation and communication tasks beyond classical capabilities.

**Exercise 2.9** Show that, for a state

$$|\psi\rangle = \alpha|00\rangle + \beta|11\rangle \quad (|\alpha|^2 + |\beta|^2 = 1), \quad (2.83)$$

with both  $\alpha$  and  $\beta$  different from zero, it is possible to choose the directions of  $\mathbf{a}$ ,  $\mathbf{a}'$ ,  $\mathbf{b}$  and  $\mathbf{b}'$  so that the CHSH inequality (2.80) is violated. Therefore, violation of Bell's inequalities is a typical feature of entangled states

## 2.6 The density matrix

Up to now, we have only considered cases in which the state of a physical quantum system was completely described by a single unit vector  $|\psi\rangle$  in the corresponding Hilbert space. In practice, the state vector  $|\psi\rangle$  is often not perfectly determined. For example, if we consider a beam of atoms emitted by a thermal source, we do not know the kinetic energy of each atom, but only the distribution of their kinetic energies. In this case, we say that our information on the system is *incomplete*. We only know that the system is in a state randomly taken from the ensemble

$$\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_l\rangle\}, \quad (2.84)$$

with probabilities  $\{p_1, p_2, \dots, p_l\}$ , satisfying the condition of unit total probability,  $\sum_i p_i = 1$ . We say that we have a *statistical mixture* (also known as a *mixed state*) of the states  $|\psi_k\rangle$ , with weights  $p_k$ . By contrast, the single states  $|\psi_k\rangle$  are known as *pure states*. We note that the states  $|\psi_k\rangle$  are not necessarily orthogonal.

As remarked in the comments to the Postulate II of quantum mechanics, the statistical mixture of the states  $|\psi_k\rangle$ , with weights  $p_k$ , should not be confused with the intrinsic quantum mechanical linear superposition

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle, \quad |c_k|^2 = p_k. \quad (2.85)$$

On the contrary, the random weighted superposition of the states in Eq. (2.84) has to be seen as an entirely classical probabilistic mechanism. As a matter of fact, it is actually impossible to describe a statistical mixture by means of an “average state vector”. As we shall see, it is instead possible to describe it using an “average operator”, which is commonly referred to as the density operator.

The probability  $p(i)$  that a measurement of the observable  $A$  yields outcome  $a_i$  is given by

$$p(i) = \sum_{k=1}^l p_k \langle \psi_k | P_i | \psi_k \rangle, \quad (2.86)$$

where  $P_i$  is the projector onto the subspace associated with the eigenvalue  $a_i$  of  $A$ . In this expression, the probabilities  $\langle \psi_k | P_i | \psi_k \rangle$  that  $A = a_i$  on the pure states  $|\psi_k\rangle$  are computed according to the measurement Postulate III, discussed in Sec. 2.3.3. As a result, the mean value of any observable  $A$  is

$$\langle A \rangle = \sum_{i=1}^n a_i p(i) = \sum_{k=1}^l p_k \sum_{i=1}^n a_i \langle \psi_k | P_i | \psi_k \rangle = \sum_{k=1}^l p_k \langle \psi_k | A | \psi_k \rangle. \quad (2.87)$$

Probabilities therefore appear twice:

- (i) in the initial (lack of) information on the system, expressed by the weights  $p_k$ ;
- (ii) in the measurement process, characterized by the probabilities  $\langle \psi_k | P_i | \psi_k \rangle$  to obtain outcomes  $a_i$  from the measurement of the observable  $A$  when the system is described by the state  $|\psi_k\rangle$ . These latter probabilities are intrinsically quantum mechanical.

The question now is how to take into account the partial information we have on the system and to simultaneously include in our description the laws of both quantum mechanics and classical probability theory.

To this purpose, it is very useful to introduce the density operator  $\rho$ , defined as

$$\rho \equiv \sum_k p_k |\psi_k\rangle\langle\psi_k|. \quad (2.88)$$

Given a generic orthonormal basis  $\{|i\rangle\}$ , with  $i = 1, 2, \dots, n$  ( $n$  is the dimension of the Hilbert space  $\mathcal{H}$  associated with the system), we naturally associate the operator  $\rho$  with a matrix representation. The corresponding matrix, known as the *density matrix*, has elements

$$\rho_{ij} \equiv \langle i | \rho | j \rangle. \quad (2.89)$$

It is also customary to call the density operator  $\rho$  in Eq. (2.88) a density matrix.

The mean value of any observable  $A$  can be straightforwardly computed by means of the density operator as follows:

$$\text{Tr}(\rho A) = \sum_{i=1}^n \langle i | \rho A | i \rangle = \sum_{k=1}^l \sum_{i=1}^n p_k \langle i | \psi_k \rangle \langle \psi_k | A | i \rangle, \quad (2.90)$$

which is equal to  $\langle A \rangle$ , as given in Eq. (2.87). The equality between (2.90) and (2.87) follows trivially, if we take into account the completeness relation  $\sum_i |i\rangle\langle i| = I$ . It is also easy to check that the probability  $p(i)$  that a measurement of the observable  $A$  gives outcome  $a_i$ , given by Eq. (2.86), is equal to

$$p(i) = \text{Tr}(\rho P_i). \quad (2.91)$$

Therefore, the density operator  $\rho$  completely characterizes the system in a mixed-state configuration; from it we can predict the probabilities of the possible outcomes of any experiment performed on the system.

As discussed in Sec. 2.3.3, if a system is in a pure state described by the wave function  $|\psi_k\rangle$  and we measure the observable  $A$ , obtaining outcome  $a_i$ , then the state of the system immediately after the measurement is

$$|\psi'_k\rangle = \frac{P_i |\psi_k\rangle}{\sqrt{\langle\psi_k|P_i|\psi_k\rangle}}, \quad (2.92)$$

with  $P_i$  the projector onto the subspace associated with the eigenvalue  $a_i$ . Therefore, if the system is in a mixed state described by the density matrix  $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$  and we obtain the outcome  $a_i$  from the measurement of  $A$ , then the new density matrix after the measurement is given by

$$\rho' = \sum_{k=1}^l p(k|i) |\psi'_k\rangle\langle\psi'_k| = \sum_{k=1}^l p(k|i) \frac{P_i |\psi_k\rangle\langle\psi_k| P_i}{\langle\psi_k|P_i|\psi_k\rangle}, \quad (2.93)$$

where  $p(k|i)$  is the probability that the system is described by the state vector  $|\psi'_k\rangle$ , given that the measurement of the observable  $A$  resulted in  $a_i$ . Elementary probability theory tells us that  $p(k,i) = p(i)p(k|i)$ , where  $p(k,i)$  is the joint probability to have the state  $|\psi'_k\rangle$  and the outcome  $a_i$ . Likewise, we have  $p(k,i) = p_k p(i|k)$  and therefore  $p(k|i) = p(k,i)/p(i) = p(i|k)p_k/p(i)$ . Observe now that the probability of obtaining result  $a_i$ , given that the system was in the state  $|\psi_k\rangle$ , is  $p(i|k) = \langle\psi_k|P_i|\psi_k\rangle$ . Finally, we can read  $p(i)$  from Eq. (2.91) and insert it in Eq. (2.93), together with the expression found for  $p(i|k)$ . We then obtain

$$\rho' = \frac{P_i \rho P_i}{\text{Tr}(\rho P_i)}. \quad (2.94)$$

It is also possible to describe the dynamical evolution of a mixed system in the density-operator picture, thus generalizing the Schrödinger equation (2.8) for a pure state of the Postulate I. We have

$$\frac{d\rho(t)}{dt} = \frac{d}{dt} \sum_{k=1}^l p_k |\psi_k(t)\rangle\langle\psi_k(t)| = \sum_{k=1}^l p_k \left[ \left( \frac{d}{dt} |\psi_k(t)\rangle \right) \langle\psi_k(t)| + |\psi_k(t)\rangle \left( \frac{d}{dt} \langle\psi_k(t)| \right) \right]. \quad (2.95)$$

The temporal evolution of the state vectors  $|\psi_k\rangle$  is governed by the Schrödinger equation, which reads

$$i\hbar \frac{d}{dt} |\psi_k(t)\rangle = H |\psi_k(t)\rangle. \quad (2.96)$$

and equivalently, for the dual vector  $\langle\psi_k|$ ,

$$i\hbar \frac{d}{dt} \langle\psi_k(t)| = \langle\psi_k(t)|H. \quad (2.97)$$

If we insert Eqs. (2.96) and (2.97) into Eq. (2.95), we finally obtain

$$\frac{d}{dt} \rho(t) = \frac{1}{i\hbar} (H\rho(t) - \rho(t)H) = \frac{1}{i\hbar} [H, \rho(t)]. \quad (2.98)$$

This equation, known as the *von Neumann equation*, governs the temporal evolution of the density operator  $\rho$ .

As we saw in Sec. 2.3.1, the state vector  $|\psi_k(t)\rangle$  at time  $t$  is related to the state vector  $|\psi_k(t_0)\rangle$  at time  $t_0$  by a unitary operator  $U(t, t_0)$ : we have  $|\psi_k(t)\rangle = U(t, t_0)|\psi_k(t_0)\rangle$ . Therefore, the density matrix  $\rho(t)$  at time  $t$  is related to the density matrix  $\rho(t_0)$  at time  $t_0$  as follows:

$$\begin{aligned}\rho(t) &= \sum_{k=1}^l p_k |\psi_k(t)\rangle \langle \psi_k(t)| = \sum_{k=1}^l p_k U(t, t_0) |\psi_k(t_0)\rangle \langle \psi_k(t_0)| U^\dagger(t, t_0) \\ &= U(t, t_0) \rho(t_0) U^\dagger(t, t_0).\end{aligned}\quad (2.99)$$

Since, as is clear from the above discussion, the postulates of quantum mechanics can be reformulated in the density-operator picture, it is completely equivalent to describe a pure system by means of either the wave function  $|\psi(t)\rangle$  or the density matrix  $\rho(t) = |\psi(t)\rangle \langle \psi(t)|$ . A nice property of the density-matrix picture is that the arbitrary global phase factor (associated with the wave function) disappears in this formulation: the state vectors  $|\psi(t)\rangle$  and  $e^{i\delta}|\psi(t)\rangle$ , with  $\delta$  a real number, give exactly the same density matrix. More importantly, as we shall see in the subsequent chapters, the density matrix is extremely useful in the description of mixed states and of composite quantum systems.

The density operator  $\rho$  satisfies the following properties:

- (1)  $\rho$  is *Hermitian*. Indeed, if we expand any pure state  $|\psi_k\rangle$  over an orthonormal basis  $\{|i\rangle\}$ ; that is,

$$|\psi_k\rangle = \sum_{i=1}^n c_i^{(k)} |i\rangle,\quad (2.100)$$

then we have

$$\rho_{ij} = \sum_{k=1}^l p_k \langle i | \psi_k \rangle \langle \psi_k | j \rangle = \sum_{k=1}^l p_k \sum_{l,m=1}^n c_l^{(k)} c_m^{(k)*} \langle i | l \rangle \langle m | j \rangle = \sum_{k=1}^l p_k c_i^{(k)} c_j^{(k)*}.\quad (2.101)$$

The last equality follows from the orthonormality condition, which implies  $\langle i | l \rangle = \delta_{il}$  and  $\langle m | j \rangle = \delta_{mj}$ . From the above expression, it is easy to check that  $\rho$  is Hermitian, since

$$\rho_{ji}^* = \sum_{k=1}^l p_k c_j^{(k)*} c_i^{(k)} = \rho_{ij}.\quad (2.102)$$

- (2)  $\rho$  has *unit trace*. Using expansion (2.101), we obtain

$$\text{Tr } \rho = \sum_{i=1}^n \rho_{ii} = \sum_{k=1}^l p_k \sum_{i=1}^n |c_i^{(k)}|^2 = \sum_{k=1}^l p_k = 1.\quad (2.103)$$

- (3)  $\rho$  is a *non-negative* operator; that is, for any vector  $|\varphi\rangle$  in the Hilbert space  $\mathcal{H}$ , we have  $\langle \varphi | \rho | \varphi \rangle \geq 0$ . Indeed, we have

$$\langle \varphi | \rho | \varphi \rangle = \langle \varphi | \left( \sum_{k=1}^l p_k |\psi_k\rangle \langle \psi_k| \right) | \varphi \rangle = \sum_{k=1}^l p_k |\langle \varphi | \psi_k \rangle|^2 \geq 0.\quad (2.104)$$

There is a simple criterion for determining whether a state is pure or mixed. It follows from the following theorem:

**Theorem 2.2** *For a mixed state  $\text{Tr} \rho^2 < 1$ , while for a pure state  $\text{Tr} \rho^2 = 1$ .*

**Proof.** Let us consider the spectral decomposition of the Hermitian operator  $\rho$ :

$$\rho = \sum_{j=1}^n \lambda_j |j\rangle\langle j|, \quad (2.105)$$

where the normalized vectors  $|j\rangle$  are orthogonal and  $\lambda_j \geq 0$  since  $\rho$  is non-negative. In the  $\{|j\rangle\}$  basis the matrix representation of  $\rho$  is diagonal and given by

$$\rho = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \quad (2.106)$$

Since, as we have shown above,  $\text{Tr} \rho = 1$ , then we have  $\sum_j \lambda_j = 1$ . Using the spectral decomposition (2.105), it is easy to compute  $\rho^2$  and we obtain

$$\rho^2 = \sum_{j=1}^n \lambda_j^2 |j\rangle\langle j|, \quad (2.107)$$

with matrix representation

$$\rho^2 = \begin{bmatrix} \lambda_1^2 & 0 & \dots & 0 \\ 0 & \lambda_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n^2 \end{bmatrix}. \quad (2.108)$$

Thus, we have  $\text{Tr} \rho^2 = \sum_j \lambda_j^2$ . Since  $\sum_j \lambda_j = 1$  and  $\lambda_j \geq 0$ , then  $0 \leq \lambda_j \leq 1$ . Therefore,  $\text{Tr} \rho^2 = 1$  if and only if  $\lambda_j = 1$  for just one  $j = \bar{j}$  and  $\lambda_{\bar{j}} = 0$  otherwise. This corresponds to a pure state, described by the density matrix  $\rho = |\bar{j}\rangle\langle \bar{j}|$ . It is also easy to check that for a pure state  $\rho^2 = \rho$ . By definition, we have a mixed state if the diagonal representation (2.105) of  $\rho$  involves more than one pure state and in this case  $\lambda_j < 1$  for all  $j$ . Therefore,  $\sum_j \lambda_j^2 < \sum_j \lambda_j = 1$ . This proves that  $\text{Tr} \rho^2 < 1$  for a mixed state.  $\square$

Let us now comment on the physical meaning of the matrix elements of the density operator  $\rho$ . From Eq. (2.101) we can see that the diagonal term

$$\rho_{ii} = \sum_k p_k |c_i^{(k)}|^2 = \text{Tr}(\rho P_i), \text{ where } P_i = |i\rangle\langle i|, \quad (2.109)$$

represents the probability that the system is left in the state  $|i\rangle$  after measuring the observable whose eigenstates are  $\{|i\rangle\}$ . For this reason we say that  $\rho_{ii}$  represents the *population* of the state  $|i\rangle$ .

The off-diagonal terms  $\rho_{ij}$  represent interference between the states  $|i\rangle$  and  $|j\rangle$ . Such interference is present for any state  $|\psi_k\rangle$  of the statistical mixture containing

a linear superposition of  $|i\rangle$  and  $|j\rangle$ . We can see from Eq. (2.101) that  $\rho_{ij}$  is a weighted sum of the interference terms  $c_i^{(k)} c_j^{(k)*}$ . We stress that the individual terms  $c_i^{(k)} c_j^{(k)*}$  appearing in the sum (2.101) are complex quantities and, therefore,  $\rho_{ij}$  can be equal to zero even though the individual terms are not. If  $\rho_{ij} \neq 0$ , then, even after averaging over the statistical mixture, a quantum-coherence effect between the states  $|i\rangle$  and  $|j\rangle$  will remain. For this reason the off-diagonal elements of the density matrix are known as *coherences*.

We point out that the distinction between diagonal and off-diagonal terms, that is, between populations and coherences, depends on the choice of the basis  $\{|i\rangle\}$ . Actually, since  $\rho$  is Hermitian, non-negative and has unit trace, it is always possible to diagonalize it and to find an orthonormal basis  $\{|m\rangle\}$  such that

$$\rho = \sum_m \alpha_m |m\rangle\langle m|, \quad 0 \leq \alpha_m \leq 1, \quad \sum_m \alpha_m = 1. \quad (2.110)$$

This implies that the density matrix  $\rho$  can always be seen as a statistical mixture of the states  $\{|m\rangle\}$ , without coherences between them, even though these states are not, in general, eigenstates of a physical observable.

### 2.6.1 Composite systems

We now discuss how the density matrix formalism turns out to be naturally suited to describe parts of composite quantum systems. Let us consider a pure state of a bipartite system. As we saw at the beginning of Sec. 2.4, this state resides in the Hilbert space  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ , which is the tensor product of the Hilbert spaces associated with the subsystems 1 and 2. We can therefore express a generic pure state  $|\psi\rangle \in \mathcal{H}$  as

$$|\psi\rangle = \sum_{i,\alpha} c_{i\alpha} |i\rangle_1 |\alpha\rangle_2, \quad \text{with } \sum_{i,\alpha} |c_{i\alpha}|^2 = 1, \quad (2.111)$$

where  $\{|i\rangle_1\}$  and  $\{|\alpha\rangle_2\}$  are basis sets for  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. The corresponding density operator is

$$\rho = |\psi\rangle\langle\psi| = \sum_{i,\alpha} \sum_{j,\beta} c_{i\alpha} c_{j\beta}^* |i\rangle_1 |\alpha\rangle_2 \langle j|_2 \langle \beta| = \sum_{i,\alpha} \sum_{j,\beta} \rho_{i\alpha;j\beta} |i\rangle_1 |\alpha\rangle_2 \langle j|_2 \langle \beta|, \quad (2.112)$$

with the matrix elements of  $\rho$  defined by

$$\rho_{i\alpha;j\beta} \equiv {}_1\langle i| {}_2\langle \alpha| \rho |j\rangle_1 |\beta\rangle_2. \quad (2.113)$$

Let us assume that the total system (1+2) is described by the density matrix  $\rho = |\psi\rangle\langle\psi|$ , and we wish to compute the mean value of an operator  $A_1$  acting only on subsystem 1. First of all, we trivially extend the operator  $A_1$  to the entire Hilbert space  $\mathcal{H}$  by defining the operator

$$\tilde{A} \equiv A_1 \otimes I_2, \quad (2.114)$$

with  $I_2$  the identity operator in  $\mathcal{H}_2$ . Thus, we have

$$\begin{aligned}\langle A_1 \rangle &= \text{Tr}(\rho \tilde{A}) = \sum_{k,\gamma} {}_1\langle k| {}_2\langle \gamma| \rho \tilde{A} |k\rangle_1 |\gamma\rangle_2 \\ &= \sum_{k,\gamma} {}_1\langle k| {}_2\langle \gamma| \left( \sum_{i,\alpha} \sum_{j,\beta} \rho_{i\alpha;j\beta} |i\rangle_1 |\alpha\rangle_2 {}_1\langle j| {}_2\langle \beta| \right) (A_1 \otimes I_2) |k\rangle_1 |\gamma\rangle_2.\end{aligned}\quad (2.115)$$

Taking into account the orthonormality relations  ${}_1\langle k|i\rangle_1 = \delta_{ik}$ ,  ${}_2\langle \gamma|\alpha\rangle_2 = \delta_{\alpha\gamma}$  and  ${}_2\langle \beta|\gamma\rangle_2 = \delta_{\beta\gamma}$ , we can remove three out of the six sums appearing in the above equation. This gives

$$\langle A_1 \rangle = \sum_{i,j,\alpha} \rho_{i\alpha;j\alpha} {}_1\langle j| A_1 |i\rangle_1. \quad (2.116)$$

It is now useful to introduce the *reduced* density matrix

$$\rho_1 \equiv \text{Tr}_2 \rho, \quad (2.117)$$

where  $\text{Tr}_2$  denotes the partial trace over subsystem 2:

$$\text{Tr}_2 \rho \equiv \sum_{\alpha} {}_2\langle \alpha| \rho | \alpha \rangle_2. \quad (2.118)$$

Analogously, it is also possible to define a reduced density matrix for subsystem 2:

$$\rho_2 \equiv \text{Tr}_1 \rho \equiv \sum_i {}_1\langle i| \rho |i\rangle_1. \quad (2.119)$$

The matrix elements of  $\rho_1$  in the  $\{|i\rangle_1\}$  basis are given by

$$(\rho_1)_{ij} = {}_1\langle i| \rho_1 |j\rangle_1 = \sum_{\alpha} \rho_{i\alpha;j\alpha}. \quad (2.120)$$

After insertion of this equality into Eq. (2.116), we obtain

$$\langle A_1 \rangle = \sum_{i,j} {}_1\langle i| \rho_1 |j\rangle_1 {}_1\langle j| A_1 |i\rangle_1 = \sum_i {}_1\langle i| \rho_1 A_1 |i\rangle_1 = \text{Tr}(\rho_1 A_1). \quad (2.121)$$

Therefore, it is possible to compute the expectation value of an operator acting only on subsystem 1 as if the system were isolated and described by the reduced density matrix  $\rho_1$ . We can conclude that  $\rho_1$ , obtained after partial tracing over subsystem 2, describes the state of subsystem 1.<sup>3</sup>

**Exercise 2.10** Show that the reduced density matrix is Hermitian, non-negative and has unit trace.

---

<sup>3</sup>The temporal evolution of the density matrix  $\rho(t)$  is governed by the von Neumann equation (2.98). At any given time, we have seen how to compute  $\rho_1(t)$  from  $\rho(t)$ . However, the problem of finding an equation describing the evolution of  $\rho_1(t)$  is much more complex. We shall discuss this issue in Sec 7.4.

### Reduced density matrix of Bell states

It is important to stress that, even though  $\rho$  corresponds to a pure state of the composite system, it is not assured that the reduced density matrices  $\rho_1$  and  $\rho_2$  describe a pure state. A very significant example is provided by the (entangled) states

$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \quad |\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle). \quad (2.122)$$

These four states are called EPR or Bell states. They constitute a basis, known as the Bell basis, for the Hilbert space of two spin- $\frac{1}{2}$  particles. For instance, let us consider the state  $|\psi^-\rangle$ . This state has density operator

$$\rho = |\psi^-\rangle\langle\psi^-| = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| - |01\rangle\langle 10| - |10\rangle\langle 01|). \quad (2.123)$$

Its matrix representation in the basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  is given by

$$\rho = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (2.124)$$

We can readily check that

$$\rho_1 = \text{Tr}_2\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}I_1. \quad (2.125)$$

Indeed, we have

$$\begin{aligned} (\rho_1)_{00} &= \rho_{00;00} + \rho_{01;01} = 0 + \frac{1}{2} = \frac{1}{2}, \\ (\rho_1)_{01} &= \rho_{00;10} + \rho_{01;11} = 0 + 0 = 0, \\ (\rho_1)_{10} &= \rho_{10;00} + \rho_{11;01} = 0 + 0 = 0, \\ (\rho_1)_{11} &= \rho_{10;10} + \rho_{11;11} = \frac{1}{2} + 0 = \frac{1}{2}. \end{aligned} \quad (2.126)$$

Likewise, we obtain  $\rho_2 = \frac{1}{2}I_2$ . Its matrix representation is given by

$$\begin{aligned} (\rho_2)_{00} &= \rho_{00;00} + \rho_{10;10} = 0 + \frac{1}{2} = \frac{1}{2}, \\ (\rho_2)_{01} &= \rho_{00;01} + \rho_{10;11} = 0 + 0 = 0, \\ (\rho_2)_{10} &= \rho_{01;00} + \rho_{11;10} = 0 + 0 = 0, \\ (\rho_2)_{11} &= \rho_{01;01} + \rho_{11;11} = \frac{1}{2} + 0 = \frac{1}{2}. \end{aligned} \quad (2.127)$$

The same expressions for the reduced density matrices are also obtained for the other states of the Bell basis. Thus,  $\rho_1$  and  $\rho_2$  clearly correspond to mixed states: we have  $\rho_1^2 = \rho_2^2 = \frac{1}{4}I$  and therefore  $\text{Tr}(\rho_1^2) = \text{Tr}(\rho_2^2) = \frac{1}{2} < 1$ . As is easy to check, this example also shows that the density matrix  $\rho$  for the entire system is *not* equal to the tensor product  $\rho_1 \otimes \rho_2$  of the reduced density matrices. This means that the quantum correlations between systems 1 and 2 are not included in  $\rho_1 \otimes \rho_2$ .

It is instructive to discuss a simple argument illustrating that the non-locality in the EPR phenomenon cannot be used to transmit information faster than light. Assume that Alice and Bob share the Bell state  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$  and

Alice wishes to employ it to instantaneously communicate a message to Bob, who may be located arbitrarily far away. We know that it is also possible to write  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle_x|1\rangle_x - |1\rangle_x|0\rangle_x)$ . Thus, it would be tempting for Alice to measure  $\sigma_z$  or  $\sigma_x$  on her half of a Bell state to communicate a bit of classical information. That is to say, she would measure  $\sigma_z$  to transmit 0 and  $\sigma_x$  to transmit 1. In both cases, Alice obtains outcomes 0 or 1 with equal probabilities  $\frac{1}{2}$ . Thus, her measurement generates the global density matrix

$$\rho^{(z)} = \frac{1}{2}(|0\rangle\langle 0| \otimes |1\rangle\langle 1| + |1\rangle\langle 1| \otimes |0\rangle\langle 0|), \quad (2.128)$$

if she measures  $\sigma_z$ , or

$$\rho^{(x)} = \frac{1}{2}(|0\rangle_x\langle 0| \otimes |1\rangle_x\langle 1| + |1\rangle_x\langle 1| \otimes |0\rangle_x\langle 0|), \quad (2.129)$$

if she measures  $\sigma_x$ . In the first case, the reduced density matrix for Bob is  $\rho_B = \text{Tr}_A \rho^{(z)}$ , in the latter  $\rho_B = \text{Tr}_A \rho^{(x)}$  (here  $\text{Tr}_A$  denotes the trace over Alice's degrees of freedom). In any instance, it is easy to check that  $\rho_B = \frac{1}{2}I$ . Since no measurement performed by Bob can distinguish between the two different preparations of the same density matrix  $\rho_B$ , the message sent by Alice is unreadable.

**Exercise 2.11** Show that, for a pure bipartite separable state, the reduced density matrices  $\rho_1$  and  $\rho_2$  correspond to pure states and the total density matrix of the system is given by  $\rho = \rho_1 \otimes \rho_2$ .

## 2.7 The Schmidt decomposition

In this section, we shall demonstrate the existence of a very useful decomposition, known as the Schmidt decomposition, for any pure state of a bipartite quantum system.

**Theorem 2.3** The Schmidt decomposition theorem: *Given a pure state  $|\psi\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$  of a bipartite quantum system, there exist orthonormal states  $\{|i\rangle_1\}$  for  $\mathcal{H}_1$  and  $\{|i'\rangle_2\}$  for  $\mathcal{H}_2$  such that*

$$|\psi\rangle = \sum_{i=1}^k \sqrt{p_i} |i\rangle_1 |i'\rangle_2 = \sqrt{p_1} |1\rangle_1 |1'\rangle_2 + \cdots + \sqrt{p_k} |k\rangle_1 |k'\rangle_2, \quad (2.130)$$

with  $p_i$  positive real numbers satisfying the condition  $\sum_{i=1}^k p_i = 1$ .

It is important to stress that the states  $\{|i\rangle_1\}$  and  $\{|i'\rangle_2\}$  depend on the particular state  $|\psi\rangle$  that we wish to expand.

**Proof.** Given an arbitrary state vector  $|\psi\rangle$  in  $\mathcal{H}$ , we can always write

$$|\psi\rangle = \sum_{i,\alpha} c_{i\alpha} |i\rangle_1 |\alpha\rangle_2, \quad (2.131)$$

with  $\{|i\rangle_1\}$  and  $\{|\alpha\rangle_2\}$  basis sets for  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. We can express this decomposition as

$$|\psi\rangle = \sum_i |i\rangle_1 |\tilde{i}\rangle_2, \quad (2.132)$$

where we have defined

$$|\tilde{i}\rangle_2 = \sum_{\alpha} c_{i\alpha} |\alpha\rangle_2. \quad (2.133)$$

Of course, in general, these states  $|\tilde{i}\rangle_2$  are neither orthogonal nor normalized. To prove the theorem, we must choose a basis  $\{|i\rangle_1\}$  in which the reduced density matrix  $\rho_1$  is diagonal:

$$\rho_1 = \sum_i p_i |i\rangle_1 \langle i|. \quad (2.134)$$

Since  $\rho_1$  is a density matrix, it is Hermitian, non-negative and has unit trace. Thus, we have  $p_i \geq 0$  and  $\sum_i p_i = 1$ . The expansion (2.134) involves only terms with  $p_i > 0$ , corresponding to the non-zero eigenvalues of  $\rho_1$ . We can also compute  $\rho_1$  from the partial trace

$$\begin{aligned} \rho_1 &= \text{Tr}_2(|\psi\rangle\langle\psi|) = \text{Tr}_2\left(\sum_{i,j} |i\rangle_1 |\tilde{i}\rangle_2 {}_1\langle j| {}_2\langle \tilde{j}|\right) \\ &= \sum_{i,j} |i\rangle_1 {}_1\langle j| \sum_k {}_2\langle k| \tilde{i} \rangle_2 {}_2\langle \tilde{j}| k \rangle_2 = \sum_{i,j} |i\rangle_1 {}_1\langle j| \sum_k {}_2\langle \tilde{j}| k \rangle_2 {}_2\langle k| \tilde{i} \rangle_2 \\ &= \sum_{i,j} {}_2\langle \tilde{j}| \tilde{i} \rangle_2 |i\rangle_1 {}_1\langle j|, \end{aligned} \quad (2.135)$$

where  $\{|k\rangle_2\}$  is an orthonormal basis for  $\mathcal{H}_2$  and the last equality follows by taking into account the completeness relation  $\sum_k |k\rangle_2 {}_2\langle k| = I_2$ . The equality between (2.134) and (2.135) requires  ${}_2\langle \tilde{j}| \tilde{i} \rangle_2 = p_i \delta_{ij}$ . Thus, the vectors  $|\tilde{i}\rangle_2$  are orthogonal and can be normalized as follows:

$$|i'\rangle_2 = \frac{1}{\sqrt{p_i}} |\tilde{i}\rangle_2. \quad (2.136)$$

After inserting (2.136) into (2.132), we obtain the Schmidt decomposition (2.130). Thus, we have explicitly constructed the orthonormal states  $\{|i\rangle_1\}$  and  $\{|i'\rangle_2\}$  which allow us to write down the Schmidt decomposition.

We can also take the partial trace over the first subsystem, obtaining

$$\rho_2 = \text{Tr}_1(|\psi\rangle\langle\psi|) = \text{Tr}_1\left(\sum_{i,j} \sqrt{p_i} \sqrt{p_j} |i\rangle_1 |i'\rangle_2 {}_1\langle j| {}_2\langle j'|\right) = \sum_i p_i |i'\rangle_2 {}_2\langle i'|.$$

□

Therefore, the reduced density matrices  $\rho_1$  and  $\rho_2$  have the same non-zero eigenvalues. Their number is also the number  $k$  of terms in the Schmidt decomposition (2.130) and is known as the *Schmidt number* (or the *Schmidt rank*) of the state  $|\psi\rangle$ . It is clear that a separable state, which by definition can be written as

$$|\psi\rangle = |\phi\rangle_1 |\xi\rangle_2, \quad (2.137)$$

has Schmidt number equal to one. Thus, we have the following entanglement criterion: a bipartite pure state is entangled if and only if its Schmidt number is greater than one.

We stress that the Schmidt number is a criterion for entanglement, not a measure of entanglement. In order to clarify this point, let us consider the following two states:

$$\begin{aligned} |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2), \\ |\psi\rangle &= \sqrt{1 - 2\epsilon^2}|0\rangle_1|0\rangle_2 + \epsilon|1\rangle_1|1\rangle_2 + \epsilon|2\rangle_1|2\rangle_2, \end{aligned} \quad (2.138)$$

with  $\epsilon \ll 1$ . It is clear that the Schmidt number of the Bell state  $|\phi^+\rangle$  is 2; that is, it is smaller than the Schmidt number of  $|\psi\rangle$ , which is 3. On the other hand, one sees intuitively that the entanglement content of a Bell state is much larger than that of the state  $|\psi\rangle$  since we have assumed  $\epsilon \ll 1$ . This intuition can be formalized by introducing, as a measure of entanglement, a quantity borrowed from condensed matter physics, the *participation ratio*:

$$\xi = \frac{1}{\sum_{i=1}^k p_i^2}. \quad (2.139)$$

This quantity is bounded between 1 and  $k$ : it is close to 1 (that is, to separability) if a single term dominates the Schmidt decomposition, whereas  $\xi = k$  if all terms in the decomposition have the same weight ( $p_1 = \dots = p_k = \frac{1}{k}$ ). We have  $\xi = 2$  for  $|\phi^+\rangle$ , which is larger than the value  $\xi \approx 1 + 4\epsilon^2$  obtained for the state  $|\psi\rangle$ .

**Exercise 2.12** Show that there are states

$$|\psi\rangle = \sum_{\alpha, \beta, \gamma} c_{\alpha\beta\gamma} |\alpha\rangle_1 |\beta\rangle_2 |\gamma\rangle_3 \quad (2.140)$$

that cannot be written as

$$|\psi\rangle = \sum_i \sqrt{p_i} |i\rangle_1 |i'\rangle_2 |i''\rangle_3. \quad (2.141)$$

This means that the Schmidt decomposition cannot be extended to systems composed of more than two parts.

It might be interesting to remark that the Schmidt decomposition is essentially a restatement of the singular value decomposition. Indeed, we can write  $M = W\Sigma V^\dagger$ , where  $M$  is the matrix of the coefficients  $c_{i\alpha}$  of the expansion of a pure state  $|\psi\rangle$  over an orthonormal basis  $\{|i\rangle_1 \otimes |\alpha\rangle_2\}$ ,  $W$  and  $V$  unitary matrices whose columns are given by the components of the vectors  $\{|i\rangle_1\}$  and  $\{|i'\rangle_2\}$  with respect to the bases  $\{|i\rangle_1\}$  and  $\{|\alpha\rangle_2\}$ , respectively, and the diagonal entries of  $\Sigma$  (singular values) are given by  $\Sigma_{ii} = \sqrt{p_i}$ .

## 2.8 Purification

Given a quantum system described by the density matrix  $\rho_1$ , it is possible to introduce another system, which we call system 2, such that the state  $|\psi\rangle$  of the composite system is a pure state and  $\rho_1 = \text{Tr}_2(|\psi\rangle\langle\psi|)$ . This procedure, known as *purification*, allows us to associate a pure state  $|\psi\rangle$  with a density matrix  $\rho_1$ . We note that the added system 2 does not necessarily have a physical significance.

We just have a useful mathematical tool at our disposal to work with pure states instead of density matrices.

A generic pure state of the global system 1 + 2 is given by

$$|\psi\rangle = \sum_{i,\alpha} c_{i\alpha} |i\rangle_1 |\alpha\rangle_2, \quad (2.142)$$

with  $\{|i\rangle_1\}$  and  $\{|\alpha\rangle_2\}$  basis sets for the Hilbert spaces associated with the subsystems 1 and 2. The corresponding density matrix is

$$\rho = \sum_{i,\alpha} \sum_{j,\beta} c_{i\alpha} c_{j\beta}^* |i\rangle_1 |\alpha\rangle_2 \langle j|_2 \langle \beta|. \quad (2.143)$$

Given a generic density matrix for system 1,

$$\rho_1 = \sum_{k,l} (\rho_1)_{k,l} |k\rangle_1 \langle l|, \quad (2.144)$$

we say that the state  $|\psi\rangle$  defined by Eq. (2.143) is a purification of  $\rho_1$  if

$$\begin{aligned} \rho_1 &= \text{Tr}_2(|\psi\rangle \langle \psi|) = \sum_{\gamma} {}_2\langle \gamma \left( \sum_{i,\alpha} \sum_{j,\beta} c_{i\alpha} c_{j\beta}^* |i\rangle_1 |\alpha\rangle_2 \langle j|_2 \langle \beta| \right) |\gamma\rangle_2 \\ &= \sum_{\gamma} \sum_{i,j} c_{i\gamma} c_{j\gamma}^* |i\rangle_1 \langle j|, \end{aligned} \quad (2.145)$$

where we have used the orthonormality relations  ${}_2\langle \beta | \gamma \rangle_2 = \delta_{\beta\gamma}$  and  ${}_2\langle \gamma | \alpha \rangle_2 = \delta_{\gamma\alpha}$ . The equality between (2.144) and (2.145) implies

$$(\rho_1)_{ij} = \sum_{\gamma} c_{i\gamma} c_{j\gamma}^*. \quad (2.146)$$

Here the matrix elements  $(\rho_1)_{ij}$  are given and we wish to determine the coefficients  $c_{i\gamma}$ . It is clear that system (2.146) always admits a solution, provided the Hilbert space of system 2 is large enough (we shall show below that it is sufficient to consider a system 2 whose Hilbert space dimension is the same as that of system 1).

As an example, we consider a spin- $\frac{1}{2}$  system whose density matrix  $\rho_1$  is known. It turns out that the addition of a second spin- $\frac{1}{2}$  system (known as an *ancillary* system) is sufficient for the purification of the density matrix  $\rho_1$ . Indeed, in this case the condition (2.146) gives the following set of equations:

$$\begin{aligned} (\rho_1)_{00} &= c_{00}c_{00}^* + c_{01}c_{01}^*, \\ (\rho_1)_{01} &= c_{00}c_{10}^* + c_{01}c_{11}^* = (\rho_1)_{10}^*, \\ (\rho_1)_{11} &= c_{10}c_{10}^* + c_{11}c_{11}^*. \end{aligned} \quad (2.147)$$

We can select a solution to this system if we put  $c_{01} = 0$ . It is then easy to find:

$$\begin{aligned} c_{00} &= \sqrt{(\rho_1)_{00}}, \quad c_{01} = 0, \\ c_{10} &= \frac{(\rho_1)_{01}^*}{\sqrt{(\rho_1)_{00}}}, \quad c_{11} = \sqrt{\frac{(\rho_1)_{00}(\rho_1)_{11} - |(\rho_1)_{01}|^2}{(\rho_1)_{00}}}. \end{aligned} \quad (2.148)$$

Thus, a possible purification is given by

$$|\psi\rangle = \sqrt{(\rho_1)_{00}} |0\rangle_1 |0\rangle_2 + \frac{(\rho_1)_{01}^*}{\sqrt{(\rho_1)_{00}}} |1\rangle_1 |0\rangle_2 + \sqrt{\frac{(\rho_1)_{00}(\rho_1)_{11} - |(\rho_1)_{01}|^2}{(\rho_1)_{00}}} |1\rangle_1 |1\rangle_2. \quad (2.149)$$

We point out that, given a system of two spin- $\frac{1}{2}$  particles, it is possible to generate any density matrix  $\rho_1$  for one of the two particles by means of unitary operations on the overall system. For this purpose, it is sufficient to prepare the state (2.149). Particle 1 is then described by the desired density matrix  $\rho_1$ , obtained after tracing over particle 2.

We note that if we express the reduced density matrix using its diagonal representation,

$$\rho_1 = \sum_i p_i |i\rangle_1 \langle i|, \quad (2.150)$$

it is sufficient to consider a system 2 having the same state space as system 1. Indeed, a purification for the density matrix (2.150) is given by

$$|\psi\rangle = \sum_i \sqrt{p_i} |i\rangle_1 |i'\rangle_2. \quad (2.151)$$

The close connection between purification and Schmidt decomposition is self-evident.

**Exercise 2.13** Find a purification for the density matrix describing the state of two spin- $\frac{1}{2}$ -particles (*Hint:* use two ancillary spin- $\frac{1}{2}$ -particles and assume  $c_{12} = c_{13} = c_{14} = c_{23} = c_{24} = c_{34} = 0$ ).

## 2.9 Generalized measurements

According to the Postulate III of quantum mechanics, a measure associated to an observable  $A$  is described by a set of projection operators  $P_i$  over the subspace corresponding to a given outcome  $a_i$ . Since any observable is mathematically equivalent to a Hermitian matrix, linear algebra tells us that the projectors satisfy the completeness relation (2.35) and the orthogonality condition (2.36).

It is possible to relax some of these conditions by constructing a so-called *generalized measurement*. Namely, the latter is described by a set  $\{M_i\}$  of measurement operators, not necessarily self-adjoint, that satisfy the completeness relation

$$\sum_i M_i^\dagger M_i = I. \quad (2.152)$$

If the state vector of the system before the measurement is  $|\psi\rangle$ , then with probability

$$p_i = \langle \psi | M_i^\dagger M_i | \psi \rangle \quad (2.153)$$

the measurement gives outcome  $i$  and the post-measurement state of the system is

$$|\psi'_i\rangle = \frac{M_i |\psi\rangle}{\sqrt{\langle \psi | M_i^\dagger M_i | \psi \rangle}}. \quad (2.154)$$

The completeness equation (2.152) assures the fact that the probabilities sum to unity; that is,  $\sum_i p_i = \sum_i \langle \psi | M_i^\dagger M_i | \psi \rangle = 1$ . It is also easy to recognize that the projective measurements described in Sec. 2.3.3 are a special case of generalized measurements, in which the operators  $M_i$  are orthogonal projectors; that is,  $M_i^\dagger = M_i$  and  $M_i M_j = \delta_{ij} M_i$ . Therefore, in this case, the completeness relation becomes  $\sum_i M_i = I$ . It turns out that projective measurements together with unitary operations are equivalent to generalized measurements, provided the Hilbert space is extended. This simply means that generalized measurements are equivalent to projective measurements on a larger Hilbert space. This statement is known as Neumark's theorem and is discussed, e.g., in Peres (1993).

In the following, we show that, if we restrict our attention to a subsystem of a given system, a projective measurement performed on the system cannot in general be described as a projective measurement on the subsystem. Let us consider the following unitary evolution of a composite system 1+2, initially in the state  $|\psi\rangle_1|0\rangle_2$ :

$$U|\psi\rangle_1|0\rangle_2 = \sum_k M_k |\psi\rangle_1|k\rangle_2, \quad (2.155)$$

where  $\{|k\rangle_2\}$  is an orthonormal basis for subsystem 2. A projective measurement, described by the projectors  $P_i = I_1 \otimes |i\rangle_2 \langle i|$ , with  $\sum_i P_i = I_1 \otimes I_2$ , gives outcome  $i$  with probability

$$p_i = \text{Tr}(\rho_{12} P_i) = \text{Tr}\left(\sum_{k,k'} M_k |\psi\rangle_1|k\rangle_2 \langle \psi|_2 \langle k'|M_{k'}^\dagger|i\rangle_2 \langle i|\right) = {}_1\langle \psi | M_i^\dagger M_i | \psi \rangle_1, \quad (2.156)$$

where  $\rho_{12} = U|\psi\rangle_1|0\rangle_2{}_1\langle \psi|_2 \langle 0|U^\dagger$  and the operators  $M_i$  (known as Kraus operators, see Sec. 7.1) satisfy the condition  $\sum_i M_i^\dagger M_i = I_1$  and, in general, are not projectors. Therefore, a standard projective measurement performed on the system can be described as a generalized measurement on subsystem 1.

### 2.9.1 POVM measurements

The “Positive Operator-Valued Measurements” (often shortened as POVM’s) are well suited to describing experiments where the system is measured only once and therefore we are not interested in the state of the system after the measurement. This is, for instance, the case of a photon detected by a photomultiplier: the photon is destroyed in the measurement process and therefore the measurement cannot be repeated. A POVM is described by a set of positive (more precisely, non-negative) operators  $F_i$  (POVM elements), such that

$$\sum_i F_i = I. \quad (2.157)$$

If the measurement is performed on a system described by the state vector  $|\psi\rangle$ , the probability of obtaining outcome  $i$  is

$$p_i = \langle \psi | F_i | \psi \rangle. \quad (2.158)$$

POVM's can be seen as generalized measurements, provided we define  $F_i = M_i^\dagger M_i$ . Indeed, it is evident that this definition assures that  $F_i$  is a non-negative operator. It is also clear that projective measurements are POVM's since in this case  $F_i = M_i^\dagger M_i = M_i$ , with  $M_i$  projectors and  $\sum_i F_i = \sum_i M_i = I$ . However, we stress that in the POVM formalism we do not make any assumption on the post-measurement state of the system.<sup>4</sup>

### Example

It is instructive to provide an example of POVMs. A two-dimensional system (A) is initially prepared in the state  $\rho$  and is coupled to an environment (B) which is itself a two-dimensional system, initially in the state  $|0\rangle$ . Let us assume that the overall system (A+B) evolves according to the unitary transformation

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & r & -t & 0 \\ -1 & r & -t & 0 \\ 0 & t & r & 1 \\ 0 & -t & -r & 1 \end{bmatrix}. \quad (2.159)$$

After that, two detectors perform a standard projective measurement on subsystems A and B, in both cases with possible outcomes 0 and 1. In general, we have four possible outcomes: 00, 01, 10 and 11 (in integer notation, 0, 1, 2 and 3), associated with the projectors

$$P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad P_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad P_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.160)$$

The probability of obtaining outcome  $i$  is given by

$$p_i = \text{Tr}(U \rho_{\text{in}}^{(\text{tot})} U^\dagger P_i) = \text{Tr}(\rho_{\text{in}}^{(\text{tot})} Q_i), \quad (2.161)$$

where  $\rho_{\text{in}}^{(\text{tot})}$  is the initial overall state and we have introduced the operators  $Q_i = U^\dagger P_i U$ . We assume that  $\rho_{\text{in}}^{(\text{tot})} = |0\rangle\langle 0| \otimes \rho$ , with matrix representation

$$\rho_{\text{in}}^{(\text{tot})} = \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (2.162)$$

where  $\rho$  and  $\mathbf{0}$  are  $2 \times 2$  submatrices and  $\mathbf{0}$  has all matrix elements equal to 0. Given this initial state, we have

$$p_i = \text{Tr}(\rho_{\text{in}}^{(\text{tot})} Q_i) = \text{Tr}(\rho F_i), \quad (2.163)$$

where  $F_i$  is the top-left  $2 \times 2$  submatrix of  $Q_i$ . In particular, if  $\rho = |\psi\rangle\langle\psi|$  is a pure state, then  $p_i = \langle\psi|F_i|\psi\rangle$ . We obtain

$$F_0 = \frac{1}{2} \begin{bmatrix} 1 & r \\ r & r^2 \end{bmatrix}, \quad F_1 = \frac{1}{2} \begin{bmatrix} 1 & -r \\ -r & r^2 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1-r^2 \end{bmatrix}, \quad (2.164)$$

---

<sup>4</sup>The POVM formalism can also be used when the system is prepared in a mixed state  $\rho$ . In this case, the probability of obtaining outcome  $i$  is  $p_i = \text{Tr}(F_i \rho)$ ; see, e.g., Peres (1993).

where we have added in  $F_2$  the contributions coming from  $Q_2$  and  $Q_3$  since they are identical. The  $F_i$  constitute a POVM. Indeed, they are non-negative operators and fulfill the condition  $\sum_i F_i = I$ .

POVM measurements are useful, for instance, to avoid misidentification of non-orthogonal states. Let us consider the following example: Alice sends Bob one of the following two states:

$$|\psi_1\rangle = \sin \theta |0\rangle + \cos \theta |1\rangle, \quad |\psi_2\rangle = \sin \theta |0\rangle - \cos \theta |1\rangle, \quad (2.165)$$

where we assume  $0 < \theta < \frac{\pi}{4}$ . Then Bob performs on the received state a measurement described by the POVM elements  $F_0$ ,  $F_1$  and  $F_2$  defined by Eq. (2.164). Bob's probability of obtaining outcome  $i$ , provided he received the state  $|\psi_k\rangle$  ( $k=1, 2$ ), is

$$p(i|k) = \langle \psi_k | F_i | \psi_k \rangle. \quad (2.166)$$

We choose  $r = \tan \theta$ . We have  $p(1|1) = 0$  and  $p(0|2) = 0$ . Therefore, the outcome  $i = 1$  excludes that the state  $|\psi_1\rangle$  was sent, whereas  $i = 0$  excludes  $|\psi_2\rangle$ . Finally, if we obtain outcome  $i = 2$ , we cannot conclude anything. Bob cannot always distinguish which one of the two non-orthogonal states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  was sent. However, taking advantage of POVM measurements, he can avoid misidentification.

## 2.10 A guide to the bibliography

There are many good books on quantum mechanics. An introductory level text is, for instance, Merzbacher (1997) whereas more advanced texts are Sakurai (1994) and the two volumes of Cohen-Tannoudji *et al.* (1977). These books focus on atomic physics topics not directly related to quantum computation and information. A quantum mechanics text closer to quantum information is Peres (1993).

There have been many experimental breakthroughs in creating larger and larger quantum superposition, see for instance Eibenberger *et al.* (2013) for complex molecules and Yin and Li (2017) for microorganisms.

The EPR paradox is due to Einstein *et al.* (1935), see also the comment by Bohr (1935). Bell's inequalities were first introduced by Bell (1964). We have presented a particular Bell inequality, known as the CHSH inequality after its four discoverers (Clauser *et al.*, 1969). An unambiguous violation of CHSH inequality was obtained in the experiments by Aspect *et al.* (1981). More recent experiments have come remarkably close to the ideal EPR thought experiment, see Brunner *et al.* (2014) for a review. For a discussion of the free will problem, see 't Hooft (2017).

A very useful introduction to the density operator formalism can be found in Cohen-Tannoudji *et al.* (1977).

Interesting discussions of quantum measurements can be found in Braginsky and Khalili (1992), Gardiner and Zoller (2000), Namiki *et al.* (1997) and Peres (1993).

## Chapter 3

# Quantum computation

This chapter introduces the basic principles of quantum computation. We shall adhere to the quantum circuit model of computation, with which it is easy to work and which is close to physical implementations. We shall not discuss the more abstract quantum Turing machine model, which, however, has been shown to be equivalent to the circuit model.

The elementary unit of quantum information and the basic building block of quantum computation is the qubit, a two-level quantum system that can be prepared, manipulated and measured in a controlled way. A quantum computer can be seen as a collection of  $n$  qubits and therefore its wave function resides in a  $2^n$ -dimensional complex Hilbert space. As far as coupling to the environment may be neglected, its evolution in time is unitary and governed by the Schrödinger equation.

A quantum computation is composed of three basic steps: preparation of the input state, implementation of the desired unitary transformation acting on this state and measurement of the output state. The output of the measurement process is inherently probabilistic and the probabilities of the different possible outputs are set by the basic postulates of quantum mechanics. Therefore, in a quantum algorithm we must, in general, repeat several times the algorithm to obtain the correct solution of our problem with probability as close to one as desired. In this sense, quantum algorithms are analogous to classical probabilistic algorithms. However, the superposition principle and quantum entanglement open up new possibilities for computation. Quantum computers are potentially more powerful than classical (deterministic or probabilistic) computers due to quantum interference and entanglement.

We show that, analogously to classical computation, there exists a small set of gates that are universal, that is, any unitary transformation can be decomposed into a sequence of these gates. We then discuss the implementation of the basic Boolean functions and arithmetic operations on a quantum computer.

### 3.1 The qubit

A classical bit is a system that can exist in two distinct states, which are used to represent 0 and 1, that is, a single binary digit. The only possible operations (gates) in such a system are the identity ( $0 \rightarrow 0, 1 \rightarrow 1$ ) and NOT ( $0 \rightarrow 1, 1 \rightarrow 0$ ). In contrast, a quantum bit (*qubit*) is a two-level quantum system, described by a two-dimensional complex Hilbert space. In this space, one may choose a pair of normalized and mutually orthogonal quantum states,

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (3.1)$$

to represent the values 0 and 1 of a classical bit. These two states form a *computational basis*. From the superposition principle, any state of the qubit may be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (3.2)$$

where the amplitudes  $\alpha$  and  $\beta$  are complex numbers, constrained by the normalization condition

$$|\alpha|^2 + |\beta|^2 = 1. \quad (3.3)$$

Since state vectors are defined only up to a global phase of no physical significance, one may choose  $\alpha$  real and positive (except for the basis state  $|1\rangle$ , in which  $\alpha = 0$ , and one may take  $\beta = 1$  real). Thus, the generic state of a qubit may be written as

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{bmatrix}, \quad (3.4)$$

with ( $0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi$ ). Therefore, unlike the classical bit, which can only be set equal to 0 or 1, the qubit resides in a vector space, parametrized by the continuous variables  $\alpha$  and  $\beta$  (or  $\theta$  and  $\phi$ ). Thus, a continuum of states is allowed. This contradicts our “classical” way of thinking: according to our intuition, a system with two states can only be in one state or in the other. However, as we have seen in Chap. 2, quantum mechanics is much more interesting and allows infinitely many other possibilities. At this stage, one might be tempted to say that a single qubit could be used to store an infinite amount of information. Actually, we must in general provide infinitely many bits to specify the complex numbers  $\alpha$  and  $\beta$  in (3.2). However, there is a catch: to extract this information we must perform a measurement and quantum mechanics tells us that from the measurement of the polarization state  $\sigma_n$  of a qubit along any axis  $n$ , we obtain only a single bit of information ( $\sigma_n = +1$  or  $\sigma_n = -1$ ). Infinitely many measurements on identically prepared single-qubit states are required to obtain  $\alpha$  and  $\beta$ .

A two-level quantum system can be used in practice as a qubit if it is possible to manipulate it as follows:

- (i) it can be prepared in some well-defined state, for example the state  $|0\rangle$ , which we call the *fiducial* state of the qubit;
- (ii) any state of the qubit can be transformed into any other state. Such transformations are carried out by means of unitary transformations, as we shall see in the next sections;
- (iii) the qubit state can be measured in the computational basis  $\{|0\rangle, |1\rangle\}$ . This means that we can measure the qubit polarization along the  $z$ -axis. As we have seen in Sec. 2.3, the Hermitian operator associated with this measurement is the Pauli operator  $\sigma_z$ , which has eigenstates  $|0\rangle$  and  $|1\rangle$ . Thus, if the state of the qubit is described by Eq. (3.4), as a result of the measurement one obtains 0 or 1 (that is,  $\sigma_z = +1$  or  $\sigma_z = -1$ ) with probabilities

$$p_0 = |\langle 0|\psi \rangle|^2 = \cos^2 \frac{\theta}{2}, \quad p_1 = |\langle 1|\psi \rangle|^2 = \sin^2 \frac{\theta}{2}, \quad (3.5)$$

which have been computed using Postulate II of quantum mechanics (discussed in Sec. 2.3).

It is important to stress that, as we shall discuss in detail in Chap. 10, requirements (i)–(iii) can be fulfilled nowadays in laboratory experiments. Physical implementations of a qubit are provided, for instance, by the electronic spin of a single-electron quantum dot, or by the state of an atom in a cavity ( $|0\rangle$  corresponds to the atomic ground state and  $|1\rangle$  to the first excited state), or by a Cooper pair tunnelling between two superconducting islands ( $|0\rangle$  if the pair is on one island and  $|1\rangle$  if it is on the other island). The controlled unitary evolution of the state of a qubit is then implemented by means of magnetic or laser fields and efficient measurement apparatuses have now been developed.

### 3.1.1 Pure qubit states: The Bloch sphere

The Bloch sphere representation is useful in thinking about qubits since it provides a geometric picture of the qubit and of the transformations that one can operate on the state of a qubit. Owing to the normalization condition (3.3), the qubit's state can be represented by a point on a sphere of unit radius, called the *Bloch sphere*. This sphere can be embedded in a three-dimensional space of Cartesian coordinates ( $x = \cos \phi \sin \theta$ ,  $y = \sin \phi \sin \theta$ ,  $z = \cos \theta$ ). Thus, the state (3.4) can be written as

$$|\psi\rangle = \begin{bmatrix} \sqrt{\frac{1+z}{2}} \\ \frac{x+iy}{\sqrt{2(1+z)}} \end{bmatrix}. \quad (3.6)$$

By definition, a Bloch vector is a vector whose components  $(x, y, z)$  single out a point on the Bloch sphere. Therefore, each Bloch vector must satisfy the normalization condition  $x^2 + y^2 + z^2 = 1$ . We can also say that the angles  $\theta$  and  $\phi$  define a Bloch vector, as shown in Fig. 3.1. In the same figure, we also show the sinusoidal

projection, in which the Bloch sphere is projected onto a plane.<sup>1</sup> The sinusoidal projection helps in visualizing the unitary transformations of the state of a qubit.

Another useful representation of the state (3.4) is obtained by means of the projector  $P = |\psi\rangle\langle\psi|$ . The matrix representation of the operator  $P$  on the basis  $\{|0\rangle, |1\rangle\}$  is given by

$$P = \begin{bmatrix} \cos^2 \frac{\theta}{2} & e^{-i\phi} \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \cos \frac{\theta}{2} & \sin^2 \frac{\theta}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix}, \quad (3.7)$$

where the matrix element  $P_{ij}$  ( $i, j = 0, 1$ ) is defined as  $\langle i|P|j\rangle$ .

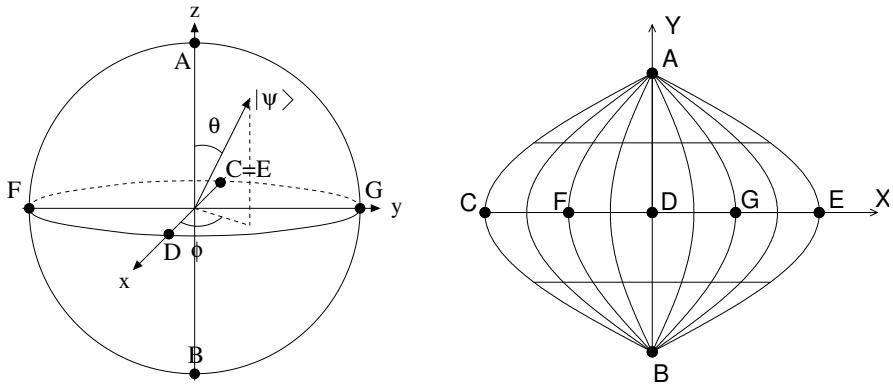


Fig. 3.1 Bloch-sphere representation of a qubit (left) and sinusoidal projection of the Bloch sphere (right). The points corresponding to the following states are shown:  $A = (\alpha=1, \beta=0)$ ,  $B = (0, 1)$ ,  $C = E = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ ,  $D = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ ,  $F = \left(\frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}\right)$ , and  $G = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right)$ . The points A (north pole of the Bloch sphere) and B (south pole) correspond to the states  $|0\rangle$  and  $|1\rangle$ , respectively.

### 3.1.2 Mixed qubit states: The Bloch ball

The above construction can be generalized by applying the density-operator formalism. As we have seen, the generic pure state of a qubit,  $|\psi(\theta, \phi)\rangle$  defined in Eq. (3.4), is represented by a Bloch vector identified by the spherical coordinates  $\theta$  and  $\phi$  on the Bloch sphere of unit radius. The corresponding density operator is given by

$$\rho(\theta, \phi) = |\psi(\theta, \phi)\rangle \langle\psi(\theta, \phi)| \quad (3.8)$$

<sup>1</sup>In this plane  $(X, Y)$  the state of a qubit has coordinates  $X = \phi \sin \theta$  and  $Y = -\theta + \frac{\pi}{2}$  (here the angle variable  $\phi$  has to be taken in the interval  $[-\pi, \pi]$ ). The sinusoidal projection is an area-preserving transformation, the parallels and the  $\phi = 0$  meridian of the Bloch sphere are straight lines, all other meridians are sinusoidal curves.

and its matrix representation in the  $\{|0\rangle, |1\rangle\}$  basis is

$$\rho(\theta, \phi) = \begin{bmatrix} \cos^2 \frac{\theta}{2} & \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} \\ \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{i\phi} & \sin^2 \frac{\theta}{2} \end{bmatrix}. \quad (3.9)$$

It is easy to check that  $\rho^2(\theta, \phi) = \rho(\theta, \phi)$ , as it must be for a pure state.

**Exercise 3.1** Show that any  $2 \times 2$  Hermitian matrix can be expanded over the basis  $\{I, \sigma_x, \sigma_y, \sigma_z\}$ , the coefficients of this expansion being real.

We now consider the density matrix  $\rho$  for the mixed state of a qubit. Since it has to be a  $2 \times 2$  Hermitian matrix, we can expand it over the basis  $\{I, \sigma_x, \sigma_y, \sigma_z\}$  (see exercise 3.1); that is,

$$\rho = aI + b\sigma_x + c\sigma_y + d\sigma_z, \quad (3.10)$$

the coefficients  $a, b, c, d$  being real. Since the condition  $\text{Tr}(\rho) = 1$  must be satisfied for a density matrix,  $\text{Tr}(I) = 2$  and  $\text{Tr}(\sigma_x) = \text{Tr}(\sigma_y) = \text{Tr}(\sigma_z) = 0$ , we have  $a = \frac{1}{2}$ . We can therefore express  $\rho$  as follows:

$$\rho = \frac{1}{2}(I + x\sigma_x + y\sigma_y + z\sigma_z) = \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix}, \quad (3.11)$$

with  $x = 2b$ ,  $y = 2c$ ,  $z = 2d$ . Any density matrix is non-negative and therefore  $\rho$  must have eigenvalues  $\lambda_1, \lambda_2 \geq 0$ . Thus, we have  $\det \rho = \lambda_1 \lambda_2 \geq 0$ . We can compute explicitly from Eq. (3.11)  $\det \rho = \frac{1}{4}(1 - |\mathbf{r}|^2)$ , with  $\mathbf{r} \equiv (x, y, z)$ . We have  $\det \rho \geq 0$  if and only if  $0 \leq |\mathbf{r}| \leq 1$ . There is a one-to-one correspondence between the density matrices for a qubit and the points on the unit ball  $0 \leq |\mathbf{r}| \leq 1$ , which is known as the *Bloch ball*. The vector  $\mathbf{r}$  is known as the Bloch vector for a generic state of the qubit. In the case of a pure state, the density matrix  $\rho$  has eigenvalues  $\lambda_1 = 1$  and  $\lambda_2 = 0$ . Thus,  $\det \rho = 0$ , which in turn implies  $|\mathbf{r}| = 1$ . We conclude that pure states are located on the boundary of the Bloch ball.

As an example of mixed state, we consider the case of a qubit which may point in any direction of the space with equal probability. We integrate over all the possible directions and obtain

$$\rho = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \begin{bmatrix} \cos^2 \frac{\theta}{2} & \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} \\ \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{i\phi} & \sin^2 \frac{\theta}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2}I, \quad (3.12)$$

where we need the normalization factor  $\frac{1}{4\pi}$  because  $\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta = 4\pi$ . We note that  $\rho^2 = \frac{1}{4}I$  and therefore  $\text{Tr}(\rho^2) = \frac{1}{2} < 1$ , as it must for a mixed state. Taking into account formula (3.11), we can easily see that the density matrix  $\rho = \frac{1}{2}I$  corresponds to the centre of the Bloch ball, identified by the zero vector  $\mathbf{r} = (0, 0, 0)$ . We say that a qubit described by this mixed state is unpolarized, because  $\langle \sigma_i \rangle = 0$ , for  $i = x, y, z$ . Indeed, we have

$$\langle \sigma_i \rangle = \text{Tr}(\rho \sigma_i) = \text{Tr}\left(\frac{1}{2} I \sigma_i\right) = \frac{1}{2} \text{Tr} \sigma_i = 0. \quad (3.13)$$

**Exercise 3.2** Show that

$$\mathrm{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}, \quad (3.14)$$

for  $i, j = x, y, z$ .

For a generic density matrix, the qubit polarization along the direction singled out by the unit vector  $\mathbf{n}$  is given by

$$\langle \sigma_{\mathbf{n}} \rangle = \mathrm{Tr}(\rho \sigma_{\mathbf{n}}) = \mathrm{Tr}\left[\frac{1}{2}(I + \mathbf{r} \cdot \boldsymbol{\sigma}) \mathbf{n} \cdot \boldsymbol{\sigma}\right]. \quad (3.15)$$

Here,  $\sigma_{\mathbf{n}} \equiv \mathbf{n} \cdot \boldsymbol{\sigma}$ , with  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  and the density matrix  $\rho$  has been expressed as in Eq. (3.11). Taking into account the result of exercise 3.2, we have

$$\langle \sigma_{\mathbf{n}} \rangle = \mathbf{n} \cdot \mathbf{r}. \quad (3.16)$$

Thus, the vector  $\mathbf{r}$  parametrizes the polarization of the qubit. As we shall see in Sec. 3.2.2, if many identically prepared systems are available; that is, the same density matrix  $\rho$  describes each system, then we can determine  $\mathbf{r}$  (and therefore the density matrix  $\rho$ ) by measuring  $\mathbf{n} \cdot \boldsymbol{\sigma}$  along three independent axes.

Finally, we note that for mixed states the decomposition of the density matrix into ensembles of pure states [Eq. (2.88)] is not unique. For example, let us consider the density matrix

$$\rho = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|. \quad (3.17)$$

This density matrix is obtained if the system is in the state  $|0\rangle$  with probability  $\frac{2}{3}$  or in the state  $|1\rangle$  with probability  $\frac{1}{3}$ . However, this is not the only ensemble of pure states giving the density matrix (3.17). There are infinitely many other possibilities, for instance we can consider the situation in which the states

$$|a\rangle = \sqrt{\frac{2}{3}}|0\rangle + \sqrt{\frac{1}{3}}|1\rangle, \quad |b\rangle = \sqrt{\frac{2}{3}}|0\rangle - \sqrt{\frac{1}{3}}|1\rangle \quad (3.18)$$

are prepared with equal probabilities  $p_a = p_b = \frac{1}{2}$ . We have

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|. \quad (3.19)$$

A further example is provided by the density matrix  $\rho = \frac{1}{2}I$ . It can correspond to the decomposition (3.12), but also to the statistical mixture of equal portions of the states  $|0\rangle_{\mathbf{n}}$  and  $|1\rangle_{\mathbf{n}}$ , where these states are the eigenvectors of the matrix  $\sigma_{\mathbf{n}}$ , with  $\mathbf{n}$  an arbitrary unit vector. For instance, considering  $\mathbf{n} = (1, 0, 0)$ , we obtain

$$\rho = \frac{1}{2}|0\rangle_x x\langle 0| + \frac{1}{2}|1\rangle_x x\langle 1| = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I. \quad (3.20)$$

Since the probabilities of the different outcomes of any conceivable experiment are governed by the density matrix  $\rho$  [see Eq. (2.91)], it is impossible to distinguish between different mixtures leading to the same  $\rho$ . Therefore, we say that, in contrast to the case of a pure state, our information on the system is *incomplete*.

**Exercise 3.3** Show that two density matrices for a qubit commute if their Bloch vectors are parallel.

### 3.2 Measuring the state of a qubit

#### 3.2.1 Pure qubit states

The state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  of a qubit can be measured in principle, provided that we have at our disposal a large number of identically prepared qubits. The Bloch-sphere representation offers a particularly well-suited framework to understand this point. In the following, we shall show that the coordinates  $x$ ,  $y$  and  $z$  of a qubit on the Bloch sphere can be measured (as we have seen, from these coordinates we can also determine  $\alpha$  and  $\beta$ , up to an overall phase factor).

Using the Pauli operators, written in the computational basis as

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (3.21)$$

one has, for the state  $|\psi\rangle$  given by (3.4),

$$\begin{aligned} \sigma_x |\psi\rangle &= e^{i\phi} \sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle, \\ \sigma_y |\psi\rangle &= -ie^{i\phi} \sin \frac{\theta}{2} |0\rangle + i \cos \frac{\theta}{2} |1\rangle, \\ \sigma_z |\psi\rangle &= \cos \frac{\theta}{2} |0\rangle - e^{i\phi} \sin \frac{\theta}{2} |1\rangle. \end{aligned} \quad (3.22)$$

Therefore, the following expectation values for the state (3.4) are obtained:

$$\begin{aligned} \langle \psi | \sigma_x | \psi \rangle &= \langle \psi | \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} | \psi \rangle = \sin \theta \cos \phi = x, \\ \langle \psi | \sigma_y | \psi \rangle &= \langle \psi | \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} | \psi \rangle = \sin \theta \sin \phi = y, \\ \langle \psi | \sigma_z | \psi \rangle &= \langle \psi | \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} | \psi \rangle = \cos \theta = z. \end{aligned} \quad (3.23)$$

The coordinates  $(x, y, z)$  can be obtained with arbitrary accuracy by means of standard projective measurements on the computational basis, that is, measuring  $\sigma_z$ . Indeed, from Eq. (3.5) we obtain

$$p_0 - p_1 = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = \cos \theta = z. \quad (3.24)$$

Thus, the coordinate  $z$  is given by the difference of the probabilities to obtain outcomes 0 or 1 from a measurement of  $\sigma_z$ . If we have at our disposal a large number  $N$  of systems identically prepared in the state (3.4), we can estimate  $z$  as  $N_0/N - N_1/N$ , where  $N_0$  and  $N_1$  count the number of outcomes 0 and 1. Therefore,  $z$  can be measured to any required accuracy, provided we measure a sufficiently large number of states.

The coordinates  $x$  and  $y$  can be obtained by using the possibility to operate a unitary transformation on the qubit. If the unitary transformation described by the matrix

$$U_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (3.25)$$

is applied to the state (3.4), we obtain the state  $|\psi^{(1)}\rangle = U_1|\psi\rangle$ . A projective measurement in the computational basis then gives outcome 0 or 1 with probabilities  $p_0^{(1)} = |\langle 0|\psi^{(1)}\rangle|^2$  and  $p_1^{(1)} = |\langle 1|\psi^{(1)}\rangle|^2$ , respectively. Therefore, we obtain

$$p_0^{(1)} - p_1^{(1)} = \cos \phi \sin \theta = x. \quad (3.26)$$

In the same way, if the state (3.4) is transformed by means of the matrix

$$U_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}, \quad (3.27)$$

we obtain the state  $|\psi^{(2)}\rangle = U_2|\psi\rangle$ . Therefore,

$$p_1^{(2)} - p_0^{(2)} = \sin \phi \sin \theta = y, \quad (3.28)$$

where  $p_0^{(2)} = |\langle 0|\psi^{(2)}\rangle|^2$  and  $p_1^{(2)} = |\langle 1|\psi^{(2)}\rangle|^2$  give the probabilities to obtain outcome 0 or 1 from the measurement of the qubit polarization along  $z$ .

We shall see in Sec. 3.4.1 that the transformation  $U_1$  corresponds to a clockwise rotation of the Bloch sphere through an angle  $\frac{\pi}{2}$  about the  $y$ -axis; in this manner, the  $x$ -axis is transformed into the  $z$ -axis and the coordinate  $x$  can be computed by measuring  $\sigma_z$  [ $U_1 = R_y(-\frac{\pi}{2})$ ]. Similarly,  $U_2$  corresponds to a clockwise rotation of the Bloch sphere through an angle  $\frac{\pi}{2}$  about the  $x$ -axis [ $U_2 = R_x(-\frac{\pi}{2})$ ].

**Exercise 3.4** The fidelity  $f$  of two pure quantum states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  is defined by  $f \equiv |\langle\psi_1|\psi_2\rangle|^2$ . It is a measure of the distance between them: we have  $0 \leq f \leq 1$ , with  $f = 1$  when  $|\psi_1\rangle$  coincides with  $|\psi_2\rangle$ , and  $f = 0$  when  $|\psi_1\rangle$  and  $|\psi_2\rangle$  are orthogonal. Show that  $f = \cos^2 \frac{\alpha}{2}$ , with  $\alpha$  being the angle between the Bloch vectors corresponding to the quantum states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ .

### 3.2.2 Mixed qubit states

Similarly to the case of pure states, it is possible to reconstruct the state of a generic mixed qubit state, provided a large number of states prepared in the same manner are available. The density matrix  $\rho$  for a qubit can be expressed using Eq. (3.11). The measurement procedure generalizes the one described in Sec. 3.2.1, and is shown in Fig. 3.2: a unitary transformation  $U$  maps  $\rho$  into a new density matrix  $\rho' = U\rho U^\dagger$  and the detector  $D$  measures  $\sigma_z$ . The possible outcomes of this measurement are  $i = 0, 1$  (we associate  $i = 0$  with  $\sigma_z = +1$  and  $i = 1$  with  $\sigma_z = -1$ ), obtained with probabilities

$$p_i = \text{Tr}(\rho' P_i), \quad (3.29)$$

where the projector operators  $P_i$  read in the  $\{|0\rangle, |1\rangle\}$  basis as follows:

$$P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad P_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.30)$$

We can also write

$$p_i = \text{Tr}(U\rho U^\dagger P_i) = \text{Tr}(\rho U^\dagger P_i U) = \text{Tr}(\rho Q_i), \quad (3.31)$$

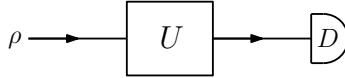


Fig. 3.2 A schematic drawing of the measurement of the density matrix for a qubit. The unitary transformation  $U$  comes before a standard measurement performed by the detector  $D$ .

where we have defined the new operators

$$Q_i \equiv U^\dagger P_i U. \quad (3.32)$$

In order to measure the coordinate  $z$ , we take  $U = I$ , so that  $Q_0 = P_0$  and  $Q_1 = P_1$ . It is easy to compute  $p_0$ ,  $p_1$  and to check that

$$p_0 - p_1 = z. \quad (3.33)$$

To compute  $x$ , we take the transformation  $U_1$  described by the matrix (3.25):

$$Q_0 = U_1^\dagger P_0 U_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad Q_1 = U_1^\dagger P_1 U_1 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (3.34)$$

In this way, it is easy to see that

$$p_0 - p_1 = x. \quad (3.35)$$

Likewise, we can compute  $y$  by taking  $U_2$  as in Eq. (3.27), so that

$$Q_0 = U_2^\dagger P_0 U_2 = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}, \quad Q_1 = U_2^\dagger P_1 U_2 = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}, \quad (3.36)$$

which implies

$$p_1 - p_0 = y. \quad (3.37)$$

Of course, we must repeat the entire procedure (preparation of the initial state, unitary transformation and measurement) a large number of times to obtain good estimates of  $x$ ,  $y$  and  $z$ . The method can be generalized to measure density matrices of larger dimensions.

### 3.3 The circuit model of quantum computation

As was shown in Chap. 1, a classical computer may be described most conveniently as a finite register of  $n$  bits. Elementary operations, such as NOT or AND, may be performed on single bits or pairs of bits, and these operations may be combined in an ordered way to produce any given complex logic function.

The circuit model can be transferred to quantum computation. The quantum computer may be thought of as a finite collection of  $n$  qubits, a *quantum register* of size  $n$ . While the state of an  $n$ -bit classical computer is described in binary notation by an integer  $i \in [0, 2^n - 1]$ ,

$$i = i_{n-1} 2^{n-1} + \cdots + i_1 2 + i_0, \quad (3.38)$$

with  $i_0, i_1, \dots, i_{n-1} \in [0, 1]$  binary digits, the state of an  $n$ -qubit quantum computer is

$$|\psi\rangle = \sum_{i=0}^{2^n-1} c_i |i\rangle = \sum_{i_{n-1}=0}^1 \cdots \sum_{i_1=0}^1 \sum_{i_0=0}^1 c_{i_{n-1}, \dots, i_1, i_0} |i_{n-1}\rangle \otimes \cdots \otimes |i_1\rangle \otimes |i_0\rangle, \quad (3.39)$$

with the complex numbers  $c_i$  constrained by the normalization condition

$$\sum_{i=0}^{2^n-1} |c_i|^2 = 1. \quad (3.40)$$

Therefore, the state of an  $n$ -qubit quantum computer is a wave function residing in a  $2^n$ -dimensional Hilbert space, constructed as the tensor product of  $n$  2-dimensional Hilbert spaces, one for each qubit. Taking into account the normalization condition (3.40) and the fact that the state of any quantum system is only defined up to a global phase of no physical significance, the state of the quantum computer is determined by  $2(2^n - 1)$  independent real parameters. As an example, we consider the case with  $n = 2$  qubits. We write the generic state of a two-qubit quantum computer as

$$\begin{aligned} |\psi\rangle &= c_0|0\rangle + c_1|1\rangle + c_2|2\rangle + c_3|3\rangle \\ &= c_{0,0}|0\rangle \otimes |0\rangle + c_{0,1}|0\rangle \otimes |1\rangle + c_{1,0}|1\rangle \otimes |0\rangle + c_{1,1}|1\rangle \otimes |1\rangle \\ &= c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle, \end{aligned} \quad (3.41)$$

where, in the last line, we have used the shorthand notation  $|i_1 i_0\rangle = |i_1\rangle \otimes |i_0\rangle$ . This notation allows us to write the state (3.39) in a simpler and compact way as

$$|\psi\rangle = \sum_{i_{n-1}, \dots, i_1, i_0=0}^1 c_{i_{n-1} \dots i_1 i_0} |i_{n-1} \dots i_1 i_0\rangle. \quad (3.42)$$

The *superposition principle* is clearly visible in Eq. (3.39): while  $n$  classical bits can store only a single integer  $i$ , the  $n$ -qubit quantum register can be prepared in the corresponding state  $|i\rangle$  of the computational basis, but also in a superposition. We stress that the number of states of the computational basis in this superposition can be as large as  $2^n$ , which grows exponentially with the number of qubits. The superposition principle opens up new possibilities for computation. When we perform a computation on a classical computer, different inputs require separate runs. In contrast, a quantum computer can perform a computation for exponentially many inputs on a single run. This huge *parallelism* is the basis of the power of quantum computation.

The superposition principle, however, is not a uniquely quantum feature and exists also for classical waves. For instance, we may consider the wave equation for a vibrating string with fixed endpoints. Its solutions  $|\varphi_i\rangle$  satisfy the superposition principle and we can write down the most general state  $|\varphi\rangle$  of a vibrating string as a linear superposition of these solutions, analogously to Eq. (3.39):

$$|\varphi\rangle = \sum_{i=0}^{2^n-1} c_i |\varphi_i\rangle. \quad (3.43)$$

It is therefore also important to point out the importance of *entanglement* for the power of quantum computation, as compared to any classical computation. Let us discuss the resources necessary to represent the superposition (3.43) in classical versus quantum physics. In order to represent the superposition of  $2^n$  levels classically, these levels must belong to the same system. Indeed, there is no entanglement in classical physics and therefore classical states of separate systems can never be superposed. Thus, we need a number of levels that grows exponentially with  $n$ . If  $\Delta$  is the typical energy separation between two consecutive levels, the amount of energy required for this computation is given by  $\Delta 2^n$ . Hence, the amount of physical resources needed for the computation grows exponentially with  $n$ .<sup>2</sup> In contrast, due to entanglement, in quantum physics a general superposition of  $2^n$  levels may be represented by means of  $n$  qubits. Thus, the amount of physical resources (energy) grows only linearly with  $n$ .

In order to perform a quantum computation, one should be able to:

- (i) *prepare* the quantum computer in a well-defined initial state  $|\psi_i\rangle$ , which we call the fiducial state of the computer, for instance the state  $|0\cdots 00\rangle$ ;
- (ii) *manipulate* the quantum-computer wave function, that is, drive any given unitary transformation  $U$ , leading to  $|\psi_f\rangle = U|\psi_i\rangle$ ;
- (iii) perform, at the end of the algorithm, a standard *measurement* in the computational basis, that is, measure the polarization  $\sigma_z$  of each qubit.

Since the quantum computer is an  $n$ -body (-qubit) quantum system, the time evolution of the wave function (3.39) is governed by the Schrödinger equation. As a result, the postulates of quantum mechanics discussed in Chap. 2 tell us that the evolution of the quantum-computer wave function is described by a unitary operator. Here we neglect non-unitary decoherence effects due to undesired coupling of the quantum computer to the environment, a problem that we shall consider in Chap. 7.

We emphasize that, even though the evolution of an  $n$ -qubit wave function is described by a  $2^n \times 2^n$  unitary matrix, this matrix can always be decomposed into a product of unitary operations acting only on one or two qubits (see Sec. 3.7). These operations are the elementary *quantum gates* of the circuit model of quantum computation.

Finally, we point out that it is possible to show that any complex collective many-qubit measurement can always be performed in the computational basis, provided that it is preceded by a suitable unitary transformation. An example of such

---

<sup>2</sup>Of course, one might imagine classical systems in which the energy levels accumulate below some upper bound. In this case, the amount of energy required for the computation could be considered constant with  $n$ . However, we would need measurement devices capable of distinguishing levels that are exponentially close in energy (their typical separation being  $\propto 2^{-n}$ ). It is reasonable to assume that exponentially large physical resources would be required for such a measurement apparatus to work.

a procedure was given in Sec. 3.2 for a single qubit. In the case of a pure state, the Bloch-sphere coordinate  $x$  (or  $y$ ) can be obtained if the unitary transformation (3.25), or (3.27), is followed by a projection onto the standard basis  $\{|0\rangle, |1\rangle\}$ . Likewise for a mixed state, one has to perform analogous unitary rotations, before measuring  $\sigma_z$ .

### 3.4 Single-qubit gates

The operations on a qubit must preserve the normalization condition (3.3) and are thus described by  $2 \times 2$  unitary matrices. In the following, we shall introduce the Hadamard and the phase-shift gates and show that they are sufficient to perform any unitary operation on a single qubit.

The *Hadamard gate* is defined as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (3.44)$$

This gate turns the computational basis  $\{|0\rangle, |1\rangle\}$  into the new basis  $\{|+\rangle, |-\rangle\}$ , whose states are a superposition of the states of the computational basis:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \equiv |+\rangle, \quad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv |-\rangle. \quad (3.45)$$

Since  $H^2 = I$ , the inverse transformation  $H^{-1} = H$ . Note that  $H$  is Hermitian. Indeed, it is evident from the matrix representation (3.44) that  $(H^T)^* = H$ .

The *phase-shift gate* is defined as

$$R_z(\delta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{bmatrix}. \quad (3.46)$$

This gate turns  $|0\rangle$  into  $|0\rangle$  and  $|1\rangle$  into  $e^{i\delta}|1\rangle$ . Since global phases have no physical meaning, the states of the computational basis,  $|0\rangle$  and  $|1\rangle$ , are unchanged. However, the action of the phase-shift gate on a generic single-qubit state  $|\psi\rangle$  [Eq. (3.4)], gives

$$R_z(\delta)|\psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta}{2} \\ e^{i(\phi+\delta)} \sin \frac{\theta}{2} \end{bmatrix}. \quad (3.47)$$

Since, as we have discussed in Sec. 2.3, relative phases are observable, the state of the qubit has been changed by the application of the phase-shift gate. It is easy to recognize from Eq. (3.47) that this gate generates a counterclockwise rotation through an angle  $\delta$  about the  $z$  axis of the Bloch sphere (see Fig. 3.1).

It turns out that any unitary operation on a single qubit can be constructed using only Hadamard and phase-shift gates. Actually, a unitary transformation moves the qubit state from one point of the Bloch sphere to another point and this can be obtained using only these two quantum gates. In particular, the generic state (3.4) can be reached starting from  $|0\rangle$  as follows:

$$R_z\left(\frac{\pi}{2} + \phi\right) H R_z(\theta) H |0\rangle = e^{i\frac{\theta}{2}} \left( \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \right). \quad (3.48)$$

**Exercise 3.5** Show that the unitary operator moving the state parametrized on the Bloch sphere by the angles  $(\theta_1, \phi_1)$  into the state  $(\theta_2, \phi_2)$  is given by

$$R_z\left(\frac{\pi}{2} + \phi_2\right) H R_z(\theta_2 - \theta_1) H R_z\left(-\frac{\pi}{2} - \phi_1\right). \quad (3.49)$$

### 3.4.1 Rotations of the Bloch sphere

We now consider a useful class of unitary transformations, the *rotations* of the Bloch sphere about an arbitrary axis. First of all, we need the following result: Let  $O$  be an operator such that  $O^2 = I$ . Thus,  $O^k = I$  for  $k$  even and  $O^k = O$  for  $k$  odd. As a consequence, from a Taylor expansion of the exponential of the operator  $O$ , one obtains

$$\begin{aligned} e^{-i\alpha O} &= \left[1 - \frac{1}{2!}\alpha^2 + \dots\right]I - i\left[\alpha - \frac{1}{3!}\alpha^3 + \dots\right]O \\ &= \cos(\alpha)I - i\sin(\alpha)O. \end{aligned} \quad (3.50)$$

Since the Pauli operators satisfy the condition  $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$ , we can apply Eq. (3.50) to exponentials of  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$ . In the case of  $\sigma_z$ , we have

$$e^{-i\frac{\delta}{2}\sigma_z} = \cos\frac{\delta}{2}I - i\sin\frac{\delta}{2}\sigma_z = e^{-i\frac{\delta}{2}} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{bmatrix} \equiv R_z(\delta). \quad (3.51)$$

We note that the above definition of  $R_z(\delta)$  differs from (3.47) only in a global phase factor of no physical significance. If we apply the phase-shift gate to a generic vector  $|\psi\rangle$  given by Eq. (3.4), we obtain, as explained in Eq. (3.47), the state

$$R_z(\delta)|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i(\phi+\delta)}\sin\frac{\theta}{2}|1\rangle. \quad (3.52)$$

Thus, if  $(x, y, z)$  denote the Cartesian coordinates of the vector  $|\psi\rangle$  and  $(x', y', z')$  the coordinates of  $R_z(\delta)|\psi\rangle$  (computed from the Bloch-sphere coordinates as explained in Sec. 3.1), we have the following coordinate transformation:

$$\begin{cases} x' = x \cos \delta - y \sin \delta, \\ y' = x \sin \delta + y \cos \delta, \\ z' = z. \end{cases} \quad (3.53)$$

Thus,  $R_z(\delta)$  corresponds to a counterclockwise rotation through an angle  $\delta$  about the  $z$ -axis of the Bloch sphere. Of course, we could equivalently say that the vector  $|\psi\rangle$  is rotated counterclockwise through an angle  $\delta$  or that the Bloch sphere itself is rotated clockwise through an angle  $\delta$ . In the first picture, we imagine the motion of the vector on a fixed Bloch sphere, in the latter picture we consider a fixed vector and moving axes. Analogously, one can obtain the unitary matrices corresponding to counterclockwise rotations about the  $x$ -axis:

$$e^{-i\frac{\delta}{2}\sigma_x} = \begin{bmatrix} \cos\frac{\delta}{2} & -i\sin\frac{\delta}{2} \\ -i\sin\frac{\delta}{2} & \cos\frac{\delta}{2} \end{bmatrix} \equiv R_x(\delta), \quad (3.54)$$

or the  $y$ -axis:

$$e^{-i\frac{\delta}{2}\sigma_y} = \begin{bmatrix} \cos \frac{\delta}{2} & -\sin \frac{\delta}{2} \\ \sin \frac{\delta}{2} & \cos \frac{\delta}{2} \end{bmatrix} \equiv R_y(\delta). \quad (3.55)$$

**Exercise 3.6** Check from the corresponding transformation of the Bloch-sphere coordinates that  $R_x(\delta)$  and  $R_y(\delta)$  correspond to counterclockwise rotations through an angle  $\delta$  about the axes  $x$  and  $y$ , respectively.

A rotation about a generic axis is obtained using the property that infinitesimal rotations can be composed as vectors. A rotation through an angle  $\epsilon \ll 1$  about the axis directed along the unit vector  $\mathbf{n} = (n_x, n_y, n_z)$  is given by the operator

$$R_{\mathbf{n}}(\epsilon) \approx_x (n_x \epsilon) R_y(n_y \epsilon) R_z(n_z \epsilon). \quad (3.56)$$

Since the Taylor expansion of Eq. (3.50) gives, for  $\epsilon \ll 1$ ,

$$R_i(n_i \epsilon) \approx I - i \frac{\epsilon}{2} n_i \sigma_i, \quad (3.57)$$

we obtain

$$R_{\mathbf{n}}(\epsilon) \approx I - i \frac{\epsilon}{2} (\mathbf{n} \cdot \boldsymbol{\sigma}), \quad (3.58)$$

where  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ . A finite rotation through an angle  $\delta$  is obtained by the composition of  $k$  infinitesimal rotations through an angle  $\epsilon = \delta/k$ :

$$R_{\mathbf{n}}(\delta) = \lim_{k \rightarrow \infty} \left[ I - i \frac{\delta}{2k} (\mathbf{n} \cdot \boldsymbol{\sigma}) \right]^k = \exp \left[ -i \frac{\delta}{2} (\mathbf{n} \cdot \boldsymbol{\sigma}) \right]. \quad (3.59)$$

Since  $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = n_x^2 \sigma_x^2 + n_y^2 \sigma_y^2 + n_z^2 \sigma_z^2 = (n_x^2 + n_y^2 + n_z^2)I = I$ , then Eq. (3.50) applies and therefore

$$R_{\mathbf{n}}(\delta) = \cos \frac{\delta}{2} I - i \sin \frac{\delta}{2} (\mathbf{n} \cdot \boldsymbol{\sigma}). \quad (3.60)$$

From this equation it is clear that we can see the Hadamard gate as a rotation through an angle  $\delta = \pi$  about the axis  $\tilde{\mathbf{n}} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ . Indeed,

$$H = \frac{1}{\sqrt{2}} (\sigma_z + \sigma_x), \quad (3.61)$$

which coincides with  $R_{\tilde{\mathbf{n}}}(\pi)$ , up to an overall phase. This transformation rotates the  $x$ -axis to  $z$  and vice versa.

**Exercise 3.7** Show that the matrices  $U_1$  and  $U_2$ , introduced in Eqs. (3.25) and (3.27), correspond to  $R_y(-\frac{\pi}{2})$  and  $R_x(-\frac{\pi}{2})$ , respectively.

**Exercise 3.8** Taking into account that a generic  $2 \times 2$  unitary matrix  $U$  can be seen (up to an overall phase factor) as a rotation of angle  $\delta$  about some axis of the Bloch sphere, compute  $\sqrt{U}$ .

**Exercise 3.9** Prove that  $(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})I + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b})$ .

### 3.5 Controlled gates and entanglement generation

Entanglement, which is the most intriguing characteristic of quantum mechanics, appears already with two qubits. Actually, a generic two-qubit state can be written in the computational basis as

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle, \quad (3.62)$$

with  $\alpha, \beta, \gamma$  and  $\delta$  complex coefficients. Taking into account the normalization condition,  $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$  and the fact that the state is only defined up to an overall phase factor, there remain 6 real degrees of freedom. Therefore, it is not possible, in general, to consider the state (3.62) as a separable state. Indeed, as discussed in Sec. 2.4, a state  $|\psi\rangle$  of a bipartite quantum system is said to be separable when it is possible to write  $|\psi\rangle = |\psi_1\rangle \otimes |\psi_0\rangle$ , with  $|\psi_1\rangle$  and  $|\psi_0\rangle$  wave functions for the two subsystems. Therefore, a separable two-qubit state has only 4 degrees of freedom since, for instance, we can take for each qubit the two parameters of its Bloch sphere. The situation becomes more complex on increasing the number of qubits. One may say that the complexity of entanglement grows exponentially with the number of qubits: while a separable state of  $n$  qubits depends only on  $2n$  real parameters, the most general (entangled) state has  $2(2^n - 1)$  degrees of freedom.

It is clear that single-qubit gates are unable to generate entanglement in an  $n$ -qubit system. Indeed, if we start from a separable state,  $|\psi\rangle = |\psi_{n-1}\rangle \otimes |\psi_{n-2}\rangle \otimes \cdots \otimes |\psi_0\rangle$ , we can move at will any qubit on its Bloch sphere, obtaining  $|\psi'\rangle = |\psi'_{n-1}\rangle \otimes |\psi'_{n-2}\rangle \otimes \cdots \otimes |\psi'_0\rangle$ . Here any state of the type  $|\psi_i\rangle$  can be transformed by gates acting on the qubit  $i$  in whatever superposition of the states  $|0\rangle$  and  $|1\rangle$ , but the  $n$ -qubit state is still separable.

To prepare an entangled state one needs inter-qubit *interactions*, that is, a two-qubit gate. The prototypical two-qubit gate that is able to generate entanglement is the controlled-NOT gate. This gate acts on the states of the computational basis,  $\{|i_1 i_0\rangle = |00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ , as the classical XOR gate:  $\text{CNOT}(|x\rangle|y\rangle) = |x\rangle|x \oplus y\rangle$ , with  $x, y = 0, 1$  and  $\oplus$  indicating addition modulo 2. The first qubit in the CNOT gate acts as a *control* and the second as a *target*. The gate flips the state of the target qubit if the control qubit is in the state  $|1\rangle$  and does nothing if the control qubit is in the state  $|0\rangle$ . We note that, as discussed in Sec. A.1, the basis vectors can be represented as column vectors:

$$|0\rangle = |00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |1\rangle = |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |2\rangle = |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |3\rangle = |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (3.63)$$

where the vector  $|i\rangle$  has the component  $i$  equal to one and all other components equal to zero. In binary notation,  $|i\rangle = |i_1 i_0\rangle$  and  $|j\rangle = |j_1 j_0\rangle$ . Therefore, we can

find a matrix representation of the CNOT gate:

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (3.64)$$

where the components  $(\text{CNOT})_{ij}$  of this matrix are given by  $(\text{CNOT})_{ij} = \langle i | \text{CNOT} | j \rangle$  (note that  $i, j = 0, \dots, 3$ ). For example, we have

$$(\text{CNOT})_{23} = \langle 2 | \text{CNOT} | 3 \rangle = \langle 10 | \text{CNOT} | 11 \rangle = 1. \quad (3.65)$$

Of course, the CNOT gate, in contrast to the classical XOR gate, can also be applied to any superposition of the computational basis states. Note that CNOT is self-inverse, since  $(\text{CNOT})^2 = I$ .

It is easy to see that CNOT can generate entangled states. For example,

$$\text{CNOT} (\alpha |0\rangle + \beta |1\rangle) |0\rangle = \alpha |00\rangle + \beta |11\rangle, \quad (3.66)$$

which is non-separable insofar as  $\alpha, \beta \neq 0$ .

**Exercise 3.10** The most general separable state of two qubits can be written, up to an overall phase, as

$$|\psi\rangle = a \{ |0\rangle + b_1 e^{i\phi_1} |1\rangle \} \otimes \{ |0\rangle + b_0 e^{i\phi_0} |1\rangle \}, \quad (3.67)$$

where  $a$  is set by the wave-function normalization. What conditions should the real coefficients  $b_0, b_1, \phi_0$  and  $\phi_1$  satisfy in order that  $\text{CNOT} |\psi\rangle$  be entangled?

It is also possible to define generalized controlled-NOT gates, depending on whether the control qubit is the first or the second qubit and whether the gate acts trivially when the control qubit is set to  $|0\rangle$  or  $|1\rangle$  (we say that a gate acts trivially if its action reduces to the identity). Correspondingly, we have the following four matrices:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.68)$$

The first of these matrices ( $A$ ) is the standard CNOT gate (3.64),  $B$  flips the second qubit if the first is set to  $|0\rangle$ ,  $C$  flips the first qubit if the second is  $|1\rangle$ , and  $D$  flips the first qubit if the second is  $|0\rangle$ . The circuit representations for the generalized CNOT gates are given in Fig. 3.3. As usual in these graphical representations, each line corresponds to a qubit and any sequence of logic gates must be read from the left (input) to the right (output). From bottom to top, qubits run from the least significant [ $i_0$ , according to binary notation (3.38)] to the most significant [ $i_{n-1}$ ]. Here a qubit is said to be more significant than another if its flip gives a larger variation in the integer number coded by the state of the  $n$  qubits.

**Exercise 3.11** Show that all four generalized CNOT gates can be constructed using only the standard CNOT gate and single-qubit gates. In particular, check that the circuit represented in Fig. 3.4 interchanges control and target qubits.

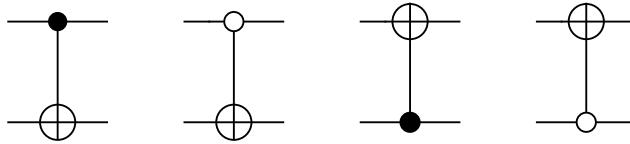


Fig. 3.3 Circuit representations for the generalized CNOT gates. From left to right:  $A$ ,  $B$ ,  $C$  and  $D$ . Note that on the control qubit we draw a full circle if the target qubit is flipped when the control is set to  $|1\rangle$ , an empty circle if instead the target is flipped when the control is  $|0\rangle$ .

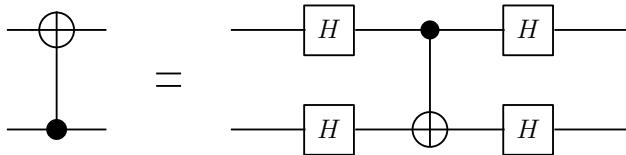


Fig. 3.4 Decomposition of the generalized CNOT gate  $C$  into a standard CNOT gate and four Hadamard gates (each of them is represented by a box with an H inside).

**Exercise 3.12** Show that all  $4! = 24$  permutations of the basis states of two qubits can be obtained using the generalized CNOT gates and draw the corresponding circuits. In particular, check that one can swap two qubits by means of the circuit of Fig. 3.5.

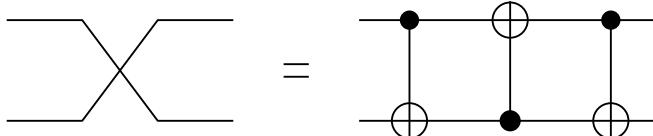


Fig. 3.5 A circuit for the SWAP gate.

Unlike the CNOT, there exist two-qubit quantum gates with no classical analog, for instance the *controlled phase shift*

$$\text{CPHASE}(\delta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{bmatrix}, \quad (3.69)$$

which applies the phase shift gate (3.46) to the target qubit only when the control qubit is in the state  $|1\rangle$ : we have  $\text{CPHASE}|11\rangle = e^{i\delta}|11\rangle$ . We show in Fig. 3.6 that a controlled phase-shift gate can be performed using CNOT gates and single-qubit phase-shift gates.

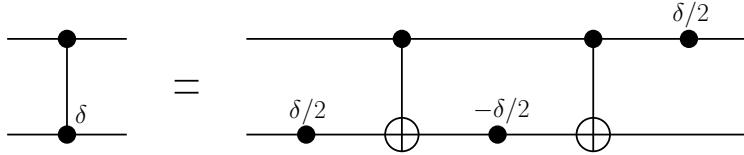


Fig. 3.6 A circuit implementing the controlled phase-shift gate (the first, the third and the last gate in the decomposition of CPHASE( $\delta$ ) are single-qubit phase-shift gates of angles  $\pm\delta/2$ ; in general a black dot with a  $\delta$  written above denotes a phase-shift gate of angle  $\delta$ ).

**Exercise 3.13** The CMINUS gate is defined as CMINUS=CPHASE( $\pi$ ), that is,

$$\text{CMINUS} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (3.70)$$

This gate is important since in some implementations it is easier to perform CMINUS rather than CNOT. Check the relation between CNOT and CMINUS shown in Fig. 3.7

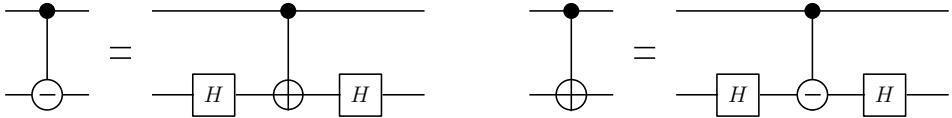


Fig. 3.7 The relation between the CNOT and CMINUS gates. The symbol on the left-hand side of the top circuit denotes CMINUS.

**Exercise 3.14 Backward sign propagation.** We define the amplitude and phase errors as follows. Given an arbitrary state of a qubit,  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , the amplitude error performs the transformation

$$|\psi\rangle \rightarrow |\psi_a\rangle = \beta|0\rangle + \alpha|1\rangle, \quad (3.71)$$

while the effect of the phase error is

$$|\psi\rangle \rightarrow |\psi_p\rangle = \alpha|0\rangle - \beta|1\rangle. \quad (3.72)$$

Discuss the effect of amplitude and phase errors, acting on the control or target qubit, on the CNOT gate. In particular, consider the initial state

$$(\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad (3.73)$$

and show that a phase error acting on the target qubit is also transferred, after application of the CNOT gate, to the control qubit. Note that, as we shall see in Chap. 4, this backward sign propagation is also a key ingredient in several quantum algorithms (*e.g.*, Deutsch's and Grover's algorithms).

**Exercise 3.15** It is possible to show that the  $4 \times 4$  Hermitian matrices constitute a linear vector space and that the tensor products  $\sigma_i \otimes \sigma_j$  are a basis for this space (where  $\sigma_0 \equiv I$ ,  $\sigma_1 \equiv \sigma_x$ ,  $\sigma_2 \equiv \sigma_y$  and  $\sigma_3 \equiv \sigma_z$ ). Therefore, all operators associated with two-qubit observables can be expanded over this basis. Compute the matrix representations of  $\sigma_i \otimes \sigma_j$  in the computational basis.

**Exercise 3.16** Compute the expectation values of the operators  $\sigma_i \otimes \sigma_j$  on the state vector

$$|\psi\rangle = c|00\rangle + \alpha|01\rangle + \beta|10\rangle + \gamma|11\rangle, \quad (3.74)$$

where the overall phase factor (up to which  $|\psi\rangle$  is defined) is chosen so that  $c$  is real (while  $\alpha$ ,  $\beta$  and  $\gamma$  are complex numbers).

### 3.5.1 The Bell basis

As we have shown, the CNOT gate can generate entanglement. For example, the entangled states of the so-called Bell basis,<sup>3</sup> defined by

$$\begin{aligned} |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), & |\phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\ |\psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), & |\psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \end{aligned} \quad (3.75)$$

can be obtained starting from the computational basis, by means of the circuit represented in Fig. 3.8. Indeed, it is easy to check that this circuit produces the following transformations:

$$|00\rangle \rightarrow |\phi^+\rangle, \quad |01\rangle \rightarrow |\psi^+\rangle, \quad |10\rangle \rightarrow |\phi^-\rangle, \quad |11\rangle \rightarrow |\psi^-\rangle. \quad (3.76)$$

This transformation can be inverted simply by running the circuit of Fig. 3.8 from right to left, since both CNOT and Hadamard gates are self-inverse. As a result, any state of the Bell basis is transformed into a separable state, this is possible since we have used a two-qubit gate. At this point it is possible, via a standard measurement in the computational basis, to establish which of the four Bell states was present at the beginning. The procedure discussed in this paragraph is known as *Bell measurement*.

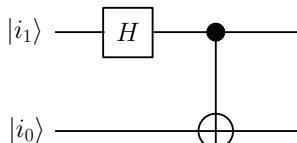


Fig. 3.8 A circuit that transforms the computational basis states  $|i_1 i_0\rangle$  to the Bell states.

---

<sup>3</sup>One of these states was introduced in Sec. 2.4 when discussing the EPR paradox.

### 3.6 Hamiltonian model for one- and two-qubit gates

To understand the basic physical principles underlying any experimental implementation of one- and two-qubit quantum gates, it is instructive to consider the following simple example. Let us assume that we have an isolated spin- $\frac{1}{2}$  particle (our qubit) in the presence of a static plus a time-dependent magnetic field. The particle could be an electron or a nuclear spin. Such a system is described by the Hamiltonian

$$H = -\mu \left\{ H_0 \sigma_z + H_1 [\cos(\omega t) \sigma_x + \sin(\omega t) \sigma_y] \right\}, \quad (3.77)$$

where  $H_0$  and  $H_1$  are the strengths of the static and the oscillating magnetic fields, respectively. Note that the static field is directed along  $z$ , while the oscillating field lies on the  $(x, y)$  plane and rotates uniformly about the  $z$  axis.

The evolution of the state  $|\psi(t)\rangle$  of the spin- $\frac{1}{2}$  particle can be computed analytically (see exercise 3.17 below). The resonance condition  $\omega = \omega_0 \equiv -2\mu H_0/\hbar$  is particularly interesting. This condition is satisfied when the angular frequency of the oscillating field (multiplied by  $\hbar$ ) is equal to the energy difference between the two spin states (in the presence of the static field only), namely, when  $\hbar\omega = -2\mu H_0$ . In this case, the solution writes

$$|\psi(t)\rangle = U |\psi(0)\rangle = e^{-i\omega\sigma_z t/2} e^{-i\Omega\sigma_x t/2} |\psi(0)\rangle, \quad (3.78)$$

where  $\Omega = -2\mu H_1/\hbar$  is the Rabi frequency. Following exercise 2.4, we can write the unitary evolution operator  $U$  of Eq. (3.78) in the computational basis. We obtain

$$U = \begin{bmatrix} e^{-i\omega t/2} & 0 \\ 0 & e^{i\omega t/2} \end{bmatrix} \begin{bmatrix} \cos(\Omega t/2) & -i \sin(\Omega t/2) \\ i \sin(\Omega t/2) & \cos(\Omega t/2) \end{bmatrix}. \quad (3.79)$$

**Exercise 3.17** Solve the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = -\mu \{ H_0 \sigma_z + H_1 [\cos(\omega t) \sigma_x + \sin(\omega t) \sigma_y] \} |\psi(t)\rangle. \quad (3.80)$$

In particular, discuss the resonance condition  $\omega = -2\mu H_0/\hbar$ .

The application, for a given period of time  $\tau$ , of an oscillating magnetic field satisfying the resonance condition  $\omega = \omega_0$ , is called a *Rabi pulse*. It is important to underline that Rabi pulses of appropriate duration implement single-qubit quantum gates. For instance, let us consider a pulse of period  $\tau$  such that  $\Omega\tau = \pi$ . It is easy to see from (3.79) that this pulse reproduces a NOT gate, up to phase factors. Indeed, if the system is initially in the state  $|0\rangle$ , it ends up in the state

$$ie^{i\omega\tau/2} |1\rangle. \quad (3.81)$$

In contrast, if the system is initially in the state  $|1\rangle$ , it ends up in the state

$$-ie^{-i\omega\tau/2} |0\rangle. \quad (3.82)$$

In order to perform exactly a NOT gate, we must eliminate the phase factors, by setting  $\omega\tau = (4n + 3)\pi$ . Similarly, we can produce arbitrary single-qubit unitary transformations.

So far, we have discussed single-qubit gates. However, the realization of two-qubit controlled gates is a necessary requirement for the implementation of universal quantum computation. It is important to point out that, for this purpose, we need *interacting* qubits. The following simple example will help clarify this concept. Let us consider a model of two coupled spin- $\frac{1}{2}$  particles, described by the Hamiltonian

$$H(t) = H_s + H_p(t), \quad (3.83)$$

where

$$H_s = -(\mu_1\sigma_1^z + \mu_2\sigma_2^z)H_0 + J\sigma_z^{(1)}\sigma_z^{(2)} \quad (3.84)$$

and  $H_p(t)$  is a time-dependent Hamiltonian describing a pulse suitable to realize a controlled gate. The first two terms in  $H_s$  describe the effect of the static magnetic field  $H_0$  on the two particles, while  $J\sigma_1^z\sigma_2^z$  represents an Ising interaction between the qubits.<sup>4</sup> The Hamiltonian  $H_s$  describes a conservative system, whose eigenstates are the states of the computational basis, namely,  $|00\rangle$ ,  $|01\rangle$ ,  $|10\rangle$  and  $|11\rangle$ , and the corresponding eigenvalues are given by

$$\begin{aligned} E_{00} &= -(\mu_1 + \mu_2)H_0 + J, & E_{01} &= -(\mu_1 - \mu_2)H_0 - J, \\ E_{10} &= (\mu_1 - \mu_2)H_0 - J, & E_{11} &= (\mu_1 + \mu_2)H_0 + J. \end{aligned} \quad (3.85)$$

Let us assume that we wish to implement a CNOT gate. This is not possible in the non-interacting case  $J = 0$ . Indeed, if in this case we apply a resonant pulse, that is, an oscillating magnetic field with frequency  $\omega = -2\mu_2H_0/\hbar$ , we induce the transition  $|0\rangle \leftrightarrow |1\rangle$  on the second qubit, *independently* of the state of the first qubit. The resonance condition is satisfied whatever the state of the first qubit is. In contrast, in the interacting case  $J \neq 0$ , we can implement a CNOT gate if the oscillating magnetic field has frequency  $\omega(J) = -2(\mu_2H_0 + J)/\hbar$ . Indeed, the resonance condition is satisfied for the transition  $|10\rangle \leftrightarrow |11\rangle$  but not for the transition  $|00\rangle \leftrightarrow |01\rangle$ . In the first case, the energy difference between the two levels involved is  $-2(\mu_2H_0 + J)$  while in the latter it is  $-2(\mu_2H_0 - J)$ .

### 3.7 Universal quantum gates

The usefulness of the circuit model in classical computation is due to the fact that a sequence of elementary operations (*e.g.*, NAND and COPY) allows one to build up arbitrarily complex computations. In this section, we shall show that a similar result exists for quantum computation, that is, any unitary operation in the Hilbert space of  $n$  qubits can be decomposed into one-qubit and two-qubit CNOT gates. In the following, we shall give a detailed proof of this important result, since it will help the reader to become familiar with quantum logic gate operations.

---

<sup>4</sup>Hereafter for systems that are composed by more than one qubit (or spin-1/2 particle), we prefer to indicate the qubit/spin label as a subscript, and place the Pauli matrix identifier ( $x$ ,  $y$ ,  $z$ ) as a superscript. In the cases where there is no need to put the qubit/spin label, we will adopt the standard convention and leave the matrix identifier as a subscript.

Let us start by defining the controlled- $U$  operation. If  $U$  is an arbitrary single-qubit unitary transformation, then controlled- $U$  means that  $U$  acts on the target qubit only if the control qubit is set to  $|1\rangle$ :

$$|i_1\rangle|i_0\rangle \rightarrow |i_1\rangle U^{i_1} |i_0\rangle. \quad (3.86)$$

**Lemma 3.1** *The controlled- $U$  gate can be implemented using only single-qubit gates and the CNOT gate.*

**Proof.** Since a matrix  $U$  is unitary if and only if its rows and columns are orthonormal, it turns out that any  $2 \times 2$  unitary matrix may be written as

$$U = \begin{bmatrix} e^{i(\delta-\alpha/2-\beta/2)} \cos \frac{\theta}{2} & -e^{i(\delta-\alpha/2+\beta/2)} \sin \frac{\theta}{2} \\ e^{i(\delta+\alpha/2-\beta/2)} \sin \frac{\theta}{2} & e^{i(\delta+\alpha/2+\beta/2)} \cos \frac{\theta}{2} \end{bmatrix}, \quad (3.87)$$

where  $\delta$ ,  $\alpha$ ,  $\beta$  and  $\theta$  are real parameters. Therefore,  $U$  can be decomposed as follows:

$$U = \Phi(\delta) R_z(\alpha) R_y(\theta) R_z(\beta), \quad (3.88)$$

where

$$\Phi(\delta) = \begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{i\delta} \end{bmatrix} \quad (3.89)$$

and  $R_y$  and  $R_z$  are the rotation matrices about the  $y$  and  $z$  axes, defined in Eqs. (3.55), (3.51). Indeed, for any  $U$  written as in Eq. (3.87), there exist three unitary matrices,  $A$ ,  $B$  and  $C$ :

$$A = R_z(\alpha) R_y\left(\frac{\theta}{2}\right), \quad B = R_y\left(-\frac{\theta}{2}\right) R_z\left(-\frac{\alpha+\beta}{2}\right), \quad C = R_z\left(\frac{\beta-\alpha}{2}\right), \quad (3.90)$$

such that

$$ABC = I \quad \text{and} \quad \Phi(\delta) A \sigma_x B \sigma_x C = U. \quad (3.91)$$

The first equality in Eq. (3.91) holds trivially, the second can be checked easily using the following properties:  $\sigma_x^2 = I$ ,  $\sigma_x R_y(\xi) \sigma_x = R_y(-\xi)$  and  $\sigma_x R_z(\xi) \sigma_x = R_z(-\xi)$ .

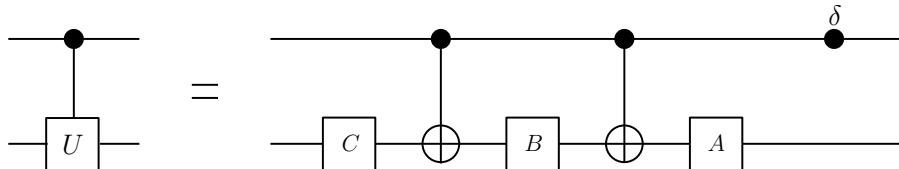


Fig. 3.9 A circuit implementing the controlled- $U$  gate.

As a consequence, it is possible to implement the controlled- $U$  operation as in Fig. 3.9. Indeed, if the value of the control qubit is 0, then  $ABC = I$  is applied to the target qubit. If the value of the control qubit is set to 1, then  $A \sigma_x B \sigma_x C = \Phi(-\delta) U$  is applied to the target qubit. Therefore, we are close to the implementation of the controlled- $U$  transformation, except for the phase factor  $\Phi(-\delta) = e^{-i\delta} I$ , which appears when the control

is set to 1. The last gate of the circuit in Fig. 3.9 removes this undesired phase factor. It acts non-trivially only on the control qubit and has the following matrix representation:

$$R_z(\delta) \otimes I = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{bmatrix} \otimes I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\delta} & 0 \\ 0 & 0 & 0 & e^{i\delta} \end{bmatrix}, \quad (3.92)$$

where the tensor product of matrices has been performed as explained in Sec. A.1.5. This gate is trivially equivalent to the controlled- $\Phi(\delta)$  gate and therefore it removes the undesired phase factor  $\Phi(-\delta)$  (as we have seen, this phase factor appears only when the control is set to 1). This completes the proof that the circuit of Fig. 3.9 implements the controlled- $U$  operation.  $\square$

We now consider the gate  $C^k\text{-}U$ , which applies a unitary transformation  $U$  to the target qubit if all the  $k$  control qubits are set to 1. We shall show that these gates can be implemented by means of elementary gates, namely, single-qubit and CNOT gates.

Of particular interest is  $C^2\text{-NOT}$  (or *Toffoli gate*), which applies a NOT operation to the target qubit only if the two control qubits are set to 1. The construction of the Toffoli gate is given in Fig. 3.10. Here  $V$  is the matrix

$$V = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}. \quad (3.93)$$

As we have seen above, since  $V$  and  $V^\dagger$  are unitary, the operations controlled- $V$  and controlled- $V^\dagger$  can be implemented using only single-qubit gates and CNOT. Thus, these elementary gates are building blocks for the Toffoli gate. This result is of particular importance for the following reasons:

- (i) since the Toffoli gate is universal for classical computation (see Chap. 1), quantum circuits having as building blocks single-qubit and CNOT gates encompass classical computation;
- (ii) unlike quantum computation, in classical computation one- and two-bit reversible gates are not universal.

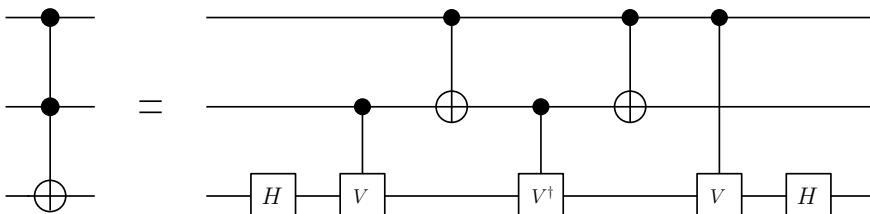


Fig. 3.10 A circuit implementing the Toffoli gate.

**Exercise 3.18** Check that for any unitary  $2 \times 2$  matrix  $U$ , the gate  $C^2\text{-}U$  can be simulated by the circuit in Fig. 3.11, where  $V$  is such that  $V^2 = U$ .

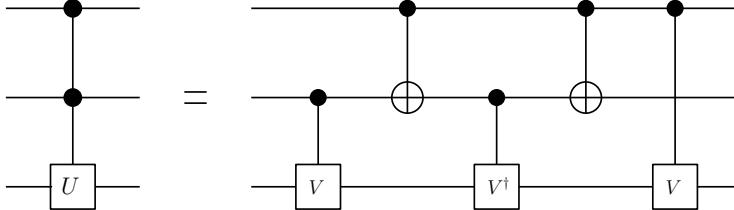


Fig. 3.11 A circuit implementing the  $C^2\text{-}U$  gate.

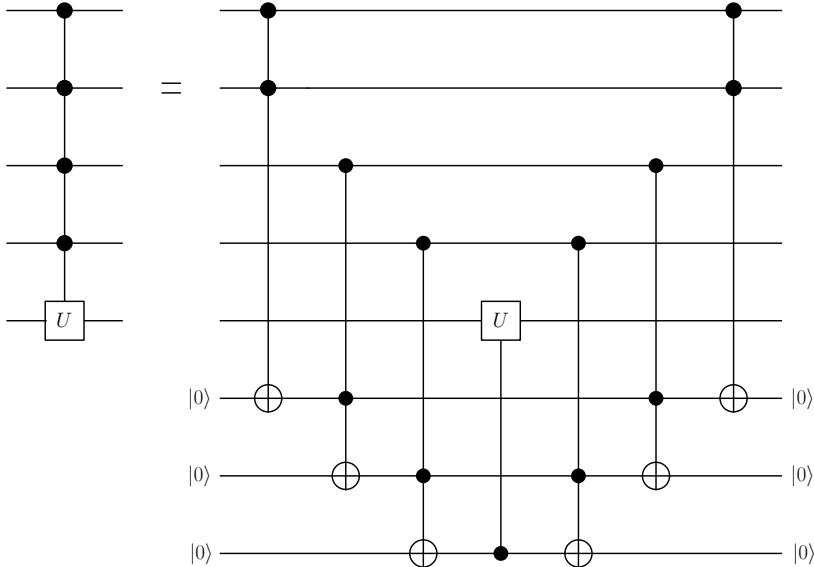
The Toffoli gate is also particularly useful in building the  $C^k\text{-}U$  gate. A simple circuit implementing this gate is shown in Fig. 3.12, for the particular case  $k = 4$ . It requires  $k - 1$  ancillary (workspace) qubits, initially set to their  $|0\rangle$  state. The first  $k - 1$  Toffoli gates change the state of the last ancillary qubit to  $|j\rangle$ , where  $j$  is given by the product  $i_{k-1}i_{k-2}i_1i_0$ , which is equal to  $|1\rangle$  if and only if all the control qubits are initially set to  $|1\rangle$ . Then a controlled- $U$  operation, having the last ancillary qubit as a control, performs the required  $C^k\text{-}U$  gate. The last  $k - 1$  Toffoli gates refresh the ancillary qubits to their initial state  $|0\rangle$ .

It is possible to show that  $C^k\text{-}U$  can be implemented without ancillary qubits (this can be achieved by means of a generalization of the circuit of Fig. 3.11, see Barenco *et al.*, 1995). The price to pay is that the number of elementary gates required is  $O(k^2)$  instead of the  $O(k)$  elementary gates used in the circuit in Fig. 3.12.

The final step in proving that single-qubit and CNOT gates are universal makes use of the decomposition formula (see Barenco *et al.*, 1995)

$$U^{(n)} = \prod_{i=1}^{2^n-1} \prod_{j=0}^{i-1} V_{ij}, \quad (3.94)$$

where  $U^{(n)}$  is a generic unitary operator acting on the  $2^n$ -dimensional Hilbert space of  $n$ -qubits and  $V_{ij}$  induces a rotation of the states  $|i\rangle$  and  $|j\rangle$  according to a unitary  $2 \times 2$  matrix. Hence,  $V_{ij}$ , when applied to a generic wave vector, acts non-trivially only on two vector components, the one along  $|i\rangle$  and the one along  $|j\rangle$ . The basic idea to implement  $V_{ij}$  on a quantum computer is to reduce the rotation of the axes  $|i\rangle$  and  $|j\rangle$  to a controlled rotation of a single qubit. For this purpose, we write a *Gray code* connecting  $i$  and  $j$ , namely, a sequence of binary numbers starting with  $i$  and finishing with  $j$ , whose consecutive members differ in one bit only. For instance,

Fig. 3.12 A circuit implementing the  $C^4 U$  gate.

if  $i = 00111010$  and  $j = 00100111$ , we have the Gray code

$$\begin{aligned}
 i &= 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
 &0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
 &0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
 &0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\
 j &= 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 .
 \end{aligned} \tag{3.95}$$

Each step of the Gray code can be performed on a quantum computer through a generalized  $C^{(n-1)}$ -NOT gate. A generalized  $C^{(n-1)}$ -NOT gate is by definition a gate in which the target qubit is flipped if and only if the  $n-1$  control qubits are in a well-defined state  $|i_{n-2} \dots i_1 i_0\rangle$ . Unlike the standard  $C^{(n-1)}$ -NOT gate, it is not required that this state correspond to  $|1 \dots 11\rangle$ . Let us consider, for instance, the first step of the Gray code (3.95). Since it changes  $i = 00111010$  into  $i' \equiv 00111011$ , it can be implemented on a quantum computer by a gate swapping the states  $|i\rangle$  and  $|i'\rangle$ . A generalized  $C^7$ -NOT accomplishes this task. It flips the state of the last qubit, provided that the first seven qubits are set to  $|0011101\rangle$ .

The penultimate line of the Gray code (3.95) differs from  $j$  (the last line) in just one bit and therefore the matrix  $V_{ij}$  can now be implemented as a rotation of the corresponding qubit, controlled by the states of all the others. Finally, the above permutations are undone in reverse order. At the end of the whole procedure, we have operated only on the states  $|i\rangle$  and  $|j\rangle$ , leaving all the other states unchanged. The quantum circuit implementing the rotation  $V_{ij}$  of the states  $|i\rangle$  and  $|j\rangle$  given in the above example is shown in Fig. 3.13. The action of this circuit can be summarized as follows: to perform a rotation  $V_{ij}$  of two generic states  $|i\rangle$  and  $|j\rangle$ ,

we operate a sequence of state permutations ( $|i\rangle = |00111010\rangle \leftrightarrow |i'\rangle = |00111011\rangle$  and so on) up to a final state ( $|i_f\rangle = |00110111\rangle$ ) that differs from  $|j\rangle$  only in one qubit. The rotation  $V_{ij}$  is then performed on the states  $|i_f\rangle$  and  $|j\rangle$ . Finally, we undo permutations so that  $|i_f\rangle$  returns back to the state  $|i\rangle$ .

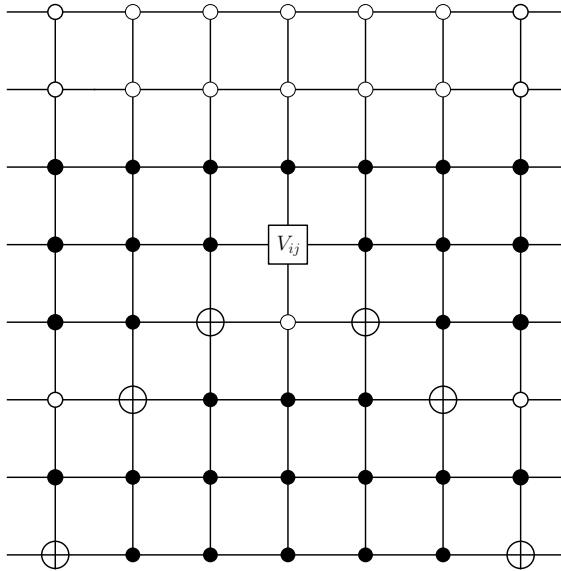


Fig. 3.13 A circuit implementing the rotation  $V_{ij}$  of the states  $|i\rangle$  and  $|j\rangle$  given in (3.95). The circuit uses 6 generalized  $C^7$ -NOT and a generalized  $C^7$ - $V_{ij}$  gate. Empty or full circles indicate that the operation on the target qubit (NOT or  $V_{ij}$ ) is active only when the control qubits are set to 0 or 1, respectively. Note that the control qubits can be both above and below the target qubit.

The results above lead to the following important result.

**Theorem 3.1** *Single-qubit plus two-qubit CNOT gates are universal gates for quantum computation.*

**Proof.** It follows directly from these main steps:

- (i) for any single-qubit rotation  $U$ , the controlled- $U$  operation can be decomposed into single-qubit and CNOT gates (Lemma 3.1);
- (ii) the  $C^2$ -NOT gate (Toffoli gate) can be implemented using CNOT, controlled- $U$  and Hadamard gates;
- (iii) any  $C^k$ - $U$  gate ( $k > 2$ ) can be decomposed into Toffoli and controlled- $U$  gates;
- (iv) a generic unitary operator  $U^{(n)}$  acting on the Hilbert space of  $n$ -qubits can be decomposed by means of  $C^k$ - $U$  gates.

□

**Exercise 3.19** Show that a generalized  $C^{(n-1)}$ -NOT gate can be obtained from the standard  $C^{(n-1)}$ -NOT gate plus single-qubit gates.

The number of elementary gates required to implement the decomposition (3.94) is  $O(n^2 4^n)$ , since there are  $O(2^n \times 2^n = 4^n)$   $V$ -terms in this product and every term requires  $O(n^2)$  elementary gates. Actually, each  $V$ -term involves at most  $2n$  permutations (elements of the Gray code plus its reverse), besides a  $C^{(n-1)}$ - $V_{ij}$  gate. Each controlled operation requires  $O(n)$  elementary gates, provided that one has at ones disposal  $n - 1$  ancillary qubits (that can be refreshed and reused each time).

We stress that the decomposition described above is in general *not efficient*, i.e., it requires a number of basic operations that scales exponentially with the number of qubits. It is also clear that to implement a generic unitary transformation  $U$  involving  $n$  qubits, one necessarily needs exponentially many elementary gates, since  $U$  is determined by  $O(4^n)$  real parameters. A fundamental and still open problem of quantum computation is to discover which special classes of unitary transformations can be computed in the quantum circuit model by means of a *polynomial* number of elementary gates.

It is interesting to give an alternative method (Tucci, 1999) to decompose a generic  $2^n \times 2^n$  unitary matrix into a sequence of elementary operations. This method utilizes the CS decomposition (C and S stand for cosine and sine, respectively). Given an  $N \times N$  unitary matrix  $U$ , where  $N$  is an even number, the CS-decomposition theorem tells us that we can always express  $U$  in the form

$$U = \begin{bmatrix} L_0 & 0 \\ 0 & L_1 \end{bmatrix} D \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix}, \quad (3.96)$$

where the matrices  $L_0$ ,  $L_1$ ,  $R_0$ , and  $R_1$  are  $(N/2) \times (N/2)$  unitary matrices and

$$D = \begin{bmatrix} D_C & -D_S \\ D_S & D_C \end{bmatrix}, \quad (3.97)$$

where  $D_C$  and  $D_S$  are diagonal matrices of the form

$$\begin{aligned} D_C &= \text{diag}(\cos \phi_1, \cos \phi_2, \dots, \cos \phi_{N/2}), \\ D_S &= \text{diag}(\sin \phi_1, \sin \phi_2, \dots, \sin \phi_{N/2}), \end{aligned} \quad (3.98)$$

with appropriate angles  $\phi_i$ . It follows from Eq. (3.96) that

$$U = \begin{bmatrix} L_0 D_C R_0 & -L_0 D_S R_1 \\ L_1 D_S R_0 & L_1 D_C R_1 \end{bmatrix}. \quad (3.99)$$

If we take  $N = 2^n$ ,  $n$  being the number of qubits, this decomposition can be iterated to matrices of smaller and smaller size. Hence, it is possible to reduce  $U$  into a sequence of elementary operations. Note that this decomposition is in general inefficient, since at the end it generates  $O(2^n)$  (controlled)  $2 \times 2$  matrices.

As a simple example, let us consider the decomposition of a  $4 \times 4$  unitary matrix  $U$ . The quantum circuit implementing Eq. (3.96) in this special case is shown in Fig. 3.14, while the decomposition of the matrix  $D$  into elementary gates is discussed in exercise 3.20.

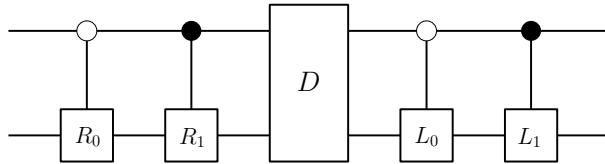


Fig. 3.14 The quantum circuit implementing the decomposition (3.96) of a  $4 \times 4$  matrix into 4 controlled operations plus a matrix  $D$ .

**Exercise 3.20** Show that the quantum circuit in Fig. 3.15 implements the unitary matrix (3.97), for a  $4 \times 4$  matrix  $D$ .

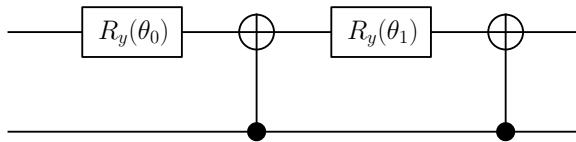


Fig. 3.15 The quantum circuit implementing the matrix  $D$  of Eq. (3.97). The rotation matrices  $R_y(\theta_i)$  are defined by Eq. (3.55).

Finally, we note that, unlike classical computation, the set of elementary gates that we have introduced is *continuous*. Indeed, this set is composed of CNOT, Hadamard and single-qubit phase-shift gates  $R_z(\delta)$ . Since  $\delta$  is a real parameter, phase-shift gates constitute a continuous set. We can easily convince ourselves that the set of elementary gates must be continuous, since the set of unitary transformations in the Hilbert space of  $n$ -qubits is continuous. However, it is possible to approximate any such transformation with arbitrary accuracy  $\epsilon$  using a *discrete* set of quantum gates. For instance, one can use Hadamard,  $R_z(\frac{\pi}{4})$  and CNOT gates. Indeed, it is possible to show that the first two gates of this set can approximate any single-qubit rotation with accuracy  $\epsilon$  in  $O(\log^c(1/\epsilon))$  steps, with the constant  $c \sim 2$  (see Nielsen and Chuang, 2000). In any case, whatever set of elementary gates, discrete or continuous, is chosen, a generic single-qubit rotation can only be simulated with finite precision by a computer operating with finite resources. For example, we should need an infinite amount of resources in order to exactly specify the real parameter  $\delta$  in the phase-shift gate  $R_z(\delta)$ . This naturally raises the question as to the stability of quantum computation in the presence of imperfect unitary operations. We shall address this problem in the next section.

### 3.7.1 \* Preparation of the initial state

In this subsection, we discuss the preparation of a generic state of the quantum computer. It turns out that this preparation in general cannot be done efficiently, since it requires a number of gates that is exponential in the number of qubits. Let us assume that the quantum computer is initially in its fiducial state  $|0\rangle$  and we

wish to prepare the state

$$|\psi\rangle = \sum_{i=0}^7 a_i |i\rangle, \quad (3.100)$$

where, for the sake of clarity, we consider the particular case corresponding to  $n = 3$  qubits.

Let us first set the amplitudes  $|a_i|$  ( $a_i = |a_i|e^{i\gamma_i}$ ). It is easy to check that this is accomplished by the circuit of Fig. 3.16 applied to the state  $|000\rangle$ . Indeed, the first gate  $R_y(2\theta_1)$ <sup>5</sup> transforms the state  $|000\rangle$  into

$$(\cos \theta_1 |0\rangle + \sin \theta_1 |1\rangle) |00\rangle. \quad (3.101)$$

Then the two generalized  $C-R_y$  gates lead to the state

$$(\cos \theta_1 \cos \theta_2 |00\rangle + \cos \theta_1 \sin \theta_2 |01\rangle + \sin \theta_1 \cos \theta_3 |10\rangle + \sin \theta_1 \sin \theta_3 |11\rangle) |0\rangle. \quad (3.102)$$

Finally, the four generalized  $C^2-R_y$  gates generate the state

$$\begin{aligned} & \cos \theta_1 \cos \theta_2 \cos \theta_4 |000\rangle + \cos \theta_1 \cos \theta_2 \sin \theta_4 |001\rangle \\ & + \cos \theta_1 \sin \theta_2 \cos \theta_5 |010\rangle + \cos \theta_1 \sin \theta_2 \sin \theta_5 |011\rangle \\ & + \sin \theta_1 \cos \theta_3 \cos \theta_6 |100\rangle + \sin \theta_1 \cos \theta_3 \sin \theta_6 |101\rangle \\ & + \sin \theta_1 \sin \theta_3 \cos \theta_7 |110\rangle + \sin \theta_1 \sin \theta_3 \sin \theta_7 |111\rangle. \end{aligned} \quad (3.103)$$

This state reproduces the amplitudes  $|a_i|$  of (3.100), provided that we set the angles  $\theta_i$  as follows:

$$\begin{aligned} |a_0| &= \cos \theta_1 \cos \theta_2 \cos \theta_4, & |a_1| &= \cos \theta_1 \cos \theta_2 \sin \theta_4, \\ |a_2| &= \cos \theta_1 \sin \theta_2 \cos \theta_5, & |a_3| &= \cos \theta_1 \sin \theta_2 \sin \theta_5, \\ |a_4| &= \sin \theta_1 \cos \theta_3 \cos \theta_6, & |a_5| &= \sin \theta_1 \cos \theta_3 \sin \theta_6, \\ |a_6| &= \sin \theta_1 \sin \theta_3 \cos \theta_7, & |a_7| &= \sin \theta_1 \sin \theta_3 \sin \theta_7. \end{aligned} \quad (3.104)$$

Thus, the angles  $\theta_i$  are determined by the amplitudes  $|a_i|$  and can be taken in the interval  $[0, \frac{\pi}{2}]$ . We note that the circuit of Fig. 3.16 requires  $2^n - 1 = 7$  (controlled) single-qubit rotations about the  $y$ -axis. This is consistent with the fact that, owing to the normalization constraint, one needs to set  $2^n - 1$  degrees of freedom to determine the  $2^n$  wave-function amplitudes.

Now we set the phases  $\gamma_i$ . One has to perform a unitary transformation  $U_D$ , whose matrix representation is diagonal in the computational basis  $\{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle\}$ :

$$U_D = \text{diag}[e^{i\gamma_0}, e^{i\gamma_1}, e^{i\gamma_2}, e^{i\gamma_3}, e^{i\gamma_4}, e^{i\gamma_5}, e^{i\gamma_6}, e^{i\gamma_7}]. \quad (3.105)$$

It is easy to check that the operator  $U_D$  is explicitly constructed in Fig. 3.17 by means of  $2^n/2$  controlled operations. In these operations,  $\Gamma_k$  is a single-qubit gate, whose matrix representation in the computational basis  $\{|0\rangle, |1\rangle\}$  is given by

$$\Gamma_k = \begin{bmatrix} e^{i\gamma_{2k}} & 0 \\ 0 & e^{i\gamma_{2k+1}} \end{bmatrix}. \quad (3.106)$$

---

<sup>5</sup>We remind the reader that the operator  $R_y$  was defined in Eq. (3.55) and that  $R_y(2\theta)$  corresponds to an anti-clockwise rotation through an angle  $2\theta$  about the  $y$ -axis of the Bloch sphere.

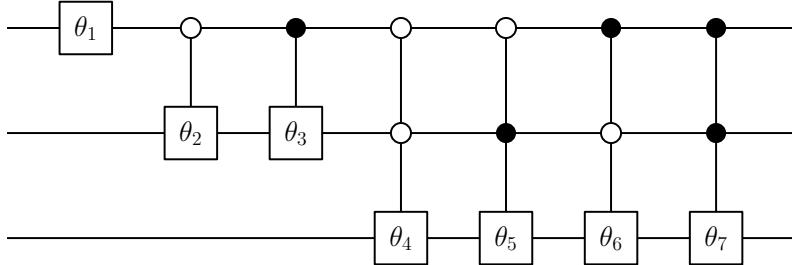


Fig. 3.16 A circuit setting the amplitudes of a generic three-qubit state. The  $\theta_i$ -symbols stand for the rotation matrices  $R_y(-2\theta_i)$  defined by Eq. (3.55).

We need  $2^N/2$  gates  $\Gamma_k$ , with  $k$  ranging from 0 to  $2^n/2 - 1$ . As can be seen from Fig. 3.17,  $\Gamma_0$  acts only when the first two qubits are in the state  $|00\rangle$  and therefore it sets the phase  $e^{i\gamma_0}$  and  $e^{i\gamma_1}$  in front of the basis vectors  $|000\rangle$  and  $|001\rangle$ , respectively. Similarly,  $\Gamma_1$  acts only when the first two qubits are in the state  $|01\rangle$  and therefore it sets the phase  $e^{i\gamma_2}$  and  $e^{i\gamma_3}$  in front of the basis vectors  $|010\rangle$  and  $|011\rangle$  and so on.

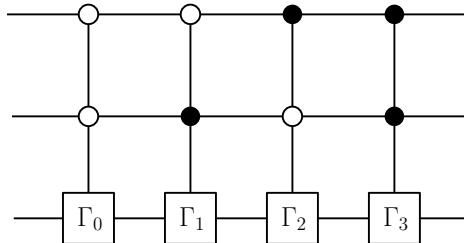


Fig. 3.17 A circuit setting the phases of a generic three-qubit state.

We remark that, while the state preparation in general requires, as in classical computation, an exponential number of operations, the quantum computer has an exponential advantage in memory requirements: A wave vector loaded in  $n$  qubits is determined by  $2^n$  complex numbers, the coefficients of its expansion over the computational basis. The classical computer needs  $O(2^n)$  bits to load  $2^n$  complex numbers (more precisely, we need  $m2^n$  bits, where  $m$  is the number of bits required to store a complex number with a given precision). The huge memory capabilities of the quantum computer appear clearly: it accomplishes this task with just  $n$  qubits.

In special cases, a given wave function can be prepared efficiently. We say that an operation in a quantum computer can be performed efficiently if it requires a number of elementary gates polynomial in the number of qubits. For instance, the equal superposition of all states of the computational basis,

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle, \quad (3.107)$$

is obtained after the application of  $n$  Hadamard gates (one for each qubit) to the state  $|0\rangle$ .

### 3.8 Unitary errors

Any quantum computation is given by a sequence of quantum gates applied to some initial state:

$$|\psi_n\rangle = \prod_{i=1}^n U_i |\psi_0\rangle. \quad (3.108)$$

Since unitary operations form a continuous set, any realistic implementation will involve some error. Let us assume for the time being that errors are unitary, while we delay the discussion of the non-unitary errors due to the unavoidable coupling to the environment until Chap. 7. Instead of the operators  $U_i$ , we apply slightly different unitary operators  $V_i$ . If we call  $|\psi_i\rangle$  the ideal state obtained after  $i$  steps, we have

$$|\psi_i\rangle = U_i |\psi_{i-1}\rangle. \quad (3.109)$$

However, if we apply the actual imperfect operation  $V_i$ , we obtain

$$V_i |\psi_{i-1}\rangle = |\psi_i\rangle + |E_i\rangle, \quad (3.110)$$

where

$$|E_i\rangle = (V_i - U_i) |\psi_{i-1}\rangle. \quad (3.111)$$

If  $|\tilde{\psi}_i\rangle$  denotes the quantum-computer wave function after  $i$  imperfect unitary transformations have been applied, then we have

$$\begin{aligned} |\tilde{\psi}_1\rangle &= |\psi_1\rangle + |E_1\rangle, \\ |\tilde{\psi}_2\rangle &= V_2 |\tilde{\psi}_1\rangle = |\psi_2\rangle + |E_2\rangle + V_2 |E_1\rangle, \end{aligned} \quad (3.112)$$

and so on. Therefore, after  $n$  iterations we obtain

$$|\tilde{\psi}_n\rangle = |\psi_n\rangle + |E_n\rangle + V_n |E_{n-1}\rangle + V_n V_{n-1} |E_{n-2}\rangle + \cdots + V_n V_{n-1} \dots V_2 |E_1\rangle. \quad (3.113)$$

In the worst case the errors are aligned and add linearly. This gives the following bound (a simple consequence of the triangle inequality)

$$\||\tilde{\psi}_n\rangle - |\psi_n\rangle\| \leq \||E_n\rangle\| + \||E_{n-1}\rangle\| + \cdots + \||E_1\rangle\|, \quad (3.114)$$

where we have used the fact that the evolution has been assumed unitary:

$$\|V_i |E_{i-1}\rangle\| = \||E_{i-1}\rangle\|. \quad (3.115)$$

We can bound the Euclidean norm of the error vector  $|E_i\rangle$  as follows:

$$\||E_i\rangle\| = \|(V_i - U_i) |\psi_{i-1}\rangle\| \leq \|V_i - U_i\|_{\sup}, \quad (3.116)$$

where  $\|V_i - U_i\|_{\sup}$  is the sup norm of the operator  $V_i - U_i$ , that is, its eigenvalue of maximum modulus. Assuming that the error is uniformly bound at each step,

$$\|V_i - U_i\|_{\sup} < \epsilon, \quad (3.117)$$

we obtain after the application of  $n$  imperfect operators

$$\|\tilde{\psi}_n\rangle - |\psi_n\rangle\| < n\epsilon. \quad (3.118)$$

Therefore, unitary errors accumulate at worst linearly with the length of the quantum computation. This growth takes place for systematic errors that line up in the same direction, while stochastic errors are randomly directed and therefore give a more favourable  $\sqrt{n}$  growth.

We note that, if a quantum computation requires  $n$  elementary gates (single-qubit and CNOT gates), it can be approximated to accuracy  $\epsilon$  using  $O(n \log^c(1/(\epsilon/n)))$  gates from the discrete set introduced in the previous section (Hadamard,  $R_z(\frac{\pi}{4})$  and CNOT). Indeed, as we stated in that section, any single-qubit rotation can be approximated with accuracy  $\epsilon/n$  in  $O(\log^c(1/(\epsilon/n)))$  gates from the above discrete set and the bound of Eq. (3.118) implies that it is sufficient to improve the accuracy of the gates linearly with the length  $n$  of the quantum computation.

It is important to connect the accuracy of the quantum-computer wave function (measured by  $\|\tilde{\psi}_n\rangle - |\psi_n\rangle\|$ ) with the accuracy of the results of a quantum computation. Quantum computation ends up with a projective measurement in the computational basis, giving outcome  $i$  with probability  $p_i = |\langle i|\psi_n\rangle|^2$ . In the presence of unitary errors, the real probability becomes  $\tilde{p}_i = |\langle i|\tilde{\psi}_n\rangle|^2$ . It is possible to show that  $\sum_i |p_i - \tilde{p}_i| \leq 2\|\tilde{\psi}_n\rangle - |\psi_n\rangle\|$  (see Preskill, 1998a).

Finally, we note that the argument developed in this section does not allow us to determine how errors scale with the number of qubits in the quantum computer.

### 3.9 Function evaluation

The basic task performed by a classical computer is the evaluation of a binary (logic) function with an  $n$ -bit input and a one-bit output:

$$f : \{0, 1\}^n \rightarrow \{0, 1\}. \quad (3.119)$$

This means that  $f$  takes the input  $(x_{n-1}, \dots, x_1, x_0)$ , with  $x_{n-1}, \dots, x_1, x_0$  binary digits, and produces the output  $f(x_{n-1}, \dots, x_1, x_0)$ , which can be equal to 0 or 1. A computer can evaluate any complicated function by assembling these binary functions (see Chap. 1). In this section, we shall discuss the implementation of such functions in a quantum computer.

Let us first discuss the construction of the binary functions for  $n = 2$  bits. There are  $2^{2^n} = 16$  two-bit logic functions, which we show in Table 3.1. These functions are not in general invertible. For instance,  $f_1(x_1, x_0) = x_1 \wedge x_0$  ( $\wedge$  denotes the logic AND) is equal to 0 for three different inputs. As we saw in Chap. 1, these functions can be evaluated in reversible computation if an ancillary bit is added. The appropriate unitary transformation that evaluates the function  $f$  on a quantum computer makes use of the ancillary qubit  $|y\rangle$  and is given by

$$U_f |x_{n-1}, x_{n-2}, \dots, x_0\rangle |y\rangle = |x_{n-1}, x_{n-2}, \dots, x_0\rangle |y \oplus f(x_{n-1}, x_{n-2}, \dots, x_0)\rangle, \quad (3.120)$$

in which, for a given output, there is a unique input.

**Exercise 3.21** Show that the transformation (3.120) is unitary.

The explicit construction of the binary functions of Table 3.1 is given by the quantum circuits of Fig. 3.18. We show only 8 functions, since  $f_{15-i} = \bar{f}_i$ , where the bar indicates NOT. Therefore,  $f_{15-i}$  can be obtained from  $f_i$  simply by application of a NOT ( $\sigma_x$ ) gate to the ancillary qubit. We note that the function  $f_6 = x_1 \oplus x_0$  could be implemented reversibly by means of a CNOT gate, with no need of any ancillary qubit. The same holds for  $f_3 = x_1$  and  $f_5 = x_0$ , which simply correspond to the input value of one of the bits, while  $f_0 = 0$  is a fully degenerate constant function.

Table 3.1 Two-bit logic functions.

$x_1 x_0$	$f_0$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$f_7$	$f_8$	$f_9$	$f_{10}$	$f_{11}$	$f_{12}$	$f_{13}$	$f_{14}$	$f_{15}$
0 0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0 1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1 0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1 1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

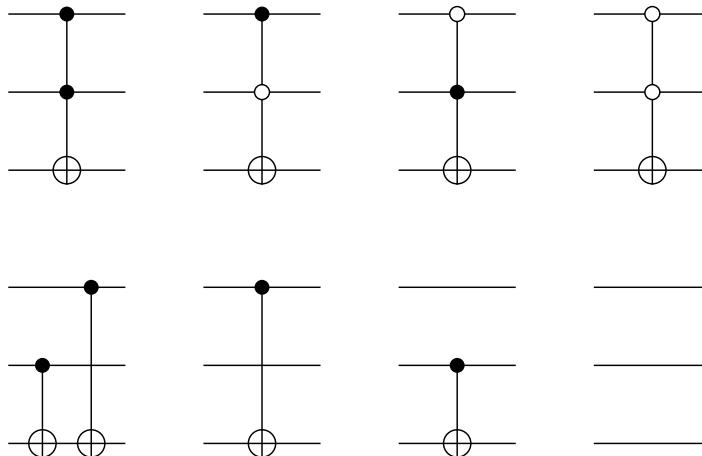


Fig. 3.18 Quantum circuits implementing the two-bit binary functions. In each circuit the three lines represent (from top to bottom) the more significant qubit, the less significant one and the ancillary, which is input in the state  $|0\rangle$ . From left to right:  $f_1 = x_1 \wedge x_0$ ,  $f_2 = x_1 \wedge x_0$ ,  $f_4 = \bar{x}_1 \wedge x_0$ ,  $f_8 = \bar{x}_1 \wedge \bar{x}_0$  (top);  $f_6 = x_1 \oplus x_0$ ,  $f_3 = x_1$ ,  $f_5 = x_0$ , and  $f_0 = 0$  (bottom).

Now let us consider a binary function with a generic number  $n$  of input bits. As we saw in Chap. 1, one way of expressing a binary function  $f(x)$  ( $x = (x_{n-1}, x_{n-2}, \dots, x_1, x_0)$ ) is to consider its *minterms*, defined, for each  $x^{(a)}$  such that  $f(x^{(a)}) = 1$ , as

$$f^{(a)}(x) = \begin{cases} 1 & \text{if } x = x^{(a)}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.121)$$

Then the function  $f(x)$  is written

$$f(x) = f^{(1)}(x) \vee f^{(2)}(x) \vee \dots \vee f^{(m)}(x), \quad (3.122)$$

where  $f$  is the logic  $\vee$  (OR) of all  $0 \leq m \leq 2^n$  minterms. Note that in Eq. (3.122) we need one minterm for each  $x$  value such that  $f(x) = 1$ . It is sufficient to compute the minterms in order to obtain  $f(x)$ .

Each minterm is implemented in a quantum computer by a generalized  $C^n$ -NOT gate. We remark that for a generic function with no structure, the number  $m$  of minterms grows exponentially with  $n$  and there is no way to evaluate  $f$  efficiently (*i.e.*, with a number of elementary gates polynomial in  $n$ ).

We now give an example of function evaluation, for the binary function  $f(x_2, x_1, x_0)$  defined by the truth table of Fig. 3.19. There are three minterms, for  $x^{(1)} = (0, 0, 1)$ ,  $x^{(2)} = (1, 0, 0)$  and  $x^{(3)} = (1, 0, 1)$ . In the corresponding quantum circuit implementing the evaluation of the function  $f$ , each generalized  $C^3$ -NOT gate corresponds to a minterm. Since the minterms  $x^{(2)}$  and  $x^{(3)}$  differ only in the value  $x_0$  of the third bit, it is possible to simplify the circuit, with a generalized  $C^2$ -NOT gate (controlled by  $x_2, x_1$ ) instead of two generalized  $C^3$ -NOT gates (controlled by  $x_2, x_1, x_0$ ). This reflects the logic identity  $x_0 + \bar{x}_0 = 1$ . We point out that the design of optimized circuits is a basic problem of computation: simplification rules of quantum logic circuits are given in Lee *et al.* (1999).

$x_2$	$x_1$	$x_0$	$f$
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

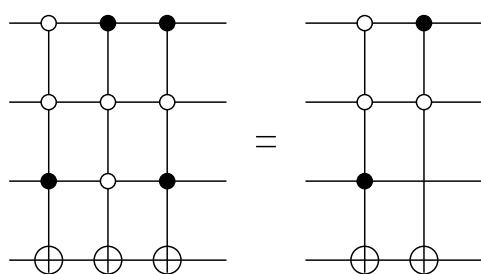


Fig. 3.19 The truth table and the quantum circuit implementation of a given binary function  $f$ . We show both the decomposition of  $f$  in minterms (left) and the simplified circuit (right). The input for the four qubits is (from top to bottom)  $|x_2\rangle$ ,  $|x_1\rangle$ ,  $|x_0\rangle$  and  $|0\rangle$ .

Let us consider a further example, the computation of the function  $x^2$  for a 2-qubit input. In general, if we need  $n = \log_2 N$  qubits to load an integer  $x \in [1, N]$ , we need  $2n = \log_2 N^2$  qubits to load  $x^2 \in [1, N^2]$ . Thus, the case  $n = 2$  requires  $2n = 4$  qubits to load the output. This corresponds to 4 binary functions, which can be evaluated reversibly using 4 ancillary qubits. The truth table for the binary function  $x^2$  and its corresponding quantum circuit are drawn in Fig. 3.20. We note that in this circuit one of the ancillary qubits is never addressed, since it gives the constant binary function  $f_0 = 0$ .

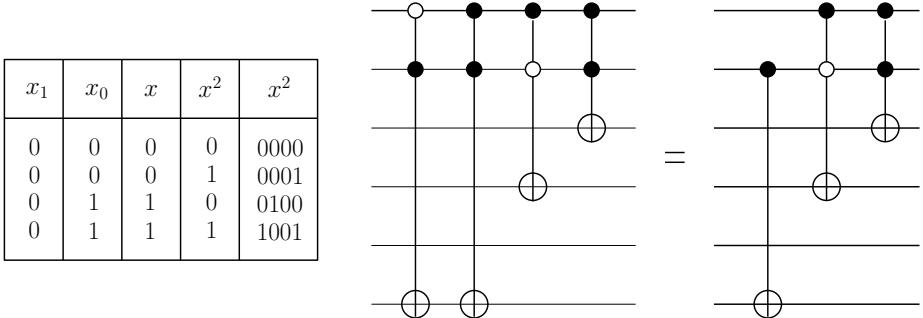


Fig. 3.20 The truth table and the quantum circuit implementation of the function  $f(x) = x^2$ , for  $n = 2$ -bit input. We show the decomposition in minterms (left) and the simplified circuit (right). The two top lines represent the input  $x$ , the four bottom lines the ancillary qubits, which are input in the state  $|0000\rangle$ . Their output loads  $x^2$  (as usual, from bottom to top, qubits run from the least significant to the most significant).

Finally, we emphasize that what makes quantum function evaluation interesting is its action on an input state given by the superposition of *exponentially many* states of the computational basis. Owing to the linearity of quantum mechanics we have:

$$U_f \sum_{x=1}^{2^n-1} c_x |x\rangle |y\rangle = \sum_{x=1}^{2^n-1} c_x |x\rangle |y \oplus f(x)\rangle, \quad (3.123)$$

which produces  $f(x)$  for all  $x$  in a single run. For instance, if we consider the function  $f(x) = x^2$  and the input

$$\frac{1}{2}(|0\rangle + |1\rangle + |2\rangle + |3\rangle)|0\rangle, \quad (3.124)$$

a single run of the circuit in Fig. 3.20 produces the output

$$\frac{1}{2}(|0\rangle|0\rangle + |1\rangle|1\rangle + |2\rangle|4\rangle + |3\rangle|9\rangle). \quad (3.125)$$

Thus, we have computed in parallel  $f(x) = x^2$  for  $x = 0, 1, 2$  and  $3$ , a possibility that is beyond the reach of the classical computer, which can only receive as input a given value of  $x$ , not a superposition of  $x$  values. This distinctive property of the quantum computer is known as *quantum parallelism*.

However, it is not an easy task to extract useful information from the superposition (3.123). The problem is that this information is, in a sense, hidden. A projective measurement of the first register in the computational basis (that is, we measure the qubit polarization along the  $z$ -axis for all the qubits in this register) yields a particular value  $x'$ , after which a measurement of the second register will necessarily give outcome  $f(x')$ . For example, if we measure the first register of the wave function (3.125), we obtain with equal probabilities  $p_0 = p_1 = p_2 = p_3 = \frac{1}{4}$  the four possible outcomes 0, 1, 2 and 3. Let us assume that the outcome is  $x' = 2$ . Then the state (3.125) collapses onto the state  $|2\rangle|4\rangle$ . Therefore, a measurement of the second register now gives outcome  $f(x') = 4$  with unit probability. Hence,

we end up with the evaluation of the function  $f(x)$  for a single value of  $x$ , exactly as with a classical computer. However, in the next chapter we shall discuss quantum algorithms that exploit quantum interference to *extract* efficiently from the superposition state (3.123) information other than just a single value of  $f$ .

### 3.10 \* The quantum adder

As in classical computation, it is important to construct quantum circuits for performing basic elementary operations. Explicit constructions for several operations, including plain addition, modular addition and modular exponentiation, can be found in Vedral *et al.* (1996). Here we describe only the plain addition of two  $n$ -bit integers  $a$  and  $b$  (in binary representation,  $a = a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + \dots + a_0 \equiv a_{n-1}a_{n-2}\dots a_0$  and analogously for  $b \equiv b_{n-1}b_{n-2}\dots b_0$ ). Following the previous section, one could consider  $n+1$  ancillary qubits and compute reversibly

$$|a, b, 0\rangle \rightarrow |a, b, a+b\rangle. \quad (3.126)$$

Since we have a  $2n$ -qubit input (encoding  $a$  and  $b$ ) and the output  $a+b$  needs  $n+1$  bits to be encoded without overflows, we must evaluate  $n+1$  binary functions with  $2n$ -bit input. This procedure is not convenient. It is more useful to compute bit by bit the following:

$$|a, b\rangle \rightarrow |a, a+b\rangle. \quad (3.127)$$

Since the input can be reconstructed from the output, this computation can be performed reversibly.

The sum is performed starting from the least significant qubit, which is the usual way to perform additions in classical computation. For the qubit  $i$ , given  $a_i$ ,  $b_i$  and the carry  $c_i$ , we need to compute the sum  $s_i = a_i \oplus b_i \oplus c_i$  and the new carry  $c_{i+1}$ . Therefore, the evaluation of two 3-bit input binary functions is required, one is called SUM and computes  $s_i$ , the other is called CARRY and computes  $c_{i+1}$ . The corresponding truth tables are given in Table 3.21. The function  $\text{SUM}(a_i, b_i, c_i)$  gives

$$s_i = a_i \oplus b_i \oplus c_i, \quad (3.128)$$

while  $\text{CARRY}(a_i, b_i, c_i)$ , as can be readily checked, produces

$$c_{i+1} = (a_i \wedge b_i) \vee (c_i \wedge a_i) \vee (c_i \wedge b_i). \quad (3.129)$$

Thus,  $c_{i+1} = 1$  when at least two of the input bits  $a_i$ ,  $b_i$  and  $c_i$  are set to one. The function SUM can be computed directly from the expression (3.128) by means of two CNOT gates. In contrast, the function CARRY involves irreversible logic functions (AND, OR) and therefore requires an ancillary qubit. The quantum circuits implementing SUM and CARRY are shown in Fig. 3.21.

The circuits implementing SUM and CARRY are the building blocks of the plain adder circuit, shown in Fig. 3.22, which computes the sum  $a+b$ . We need three

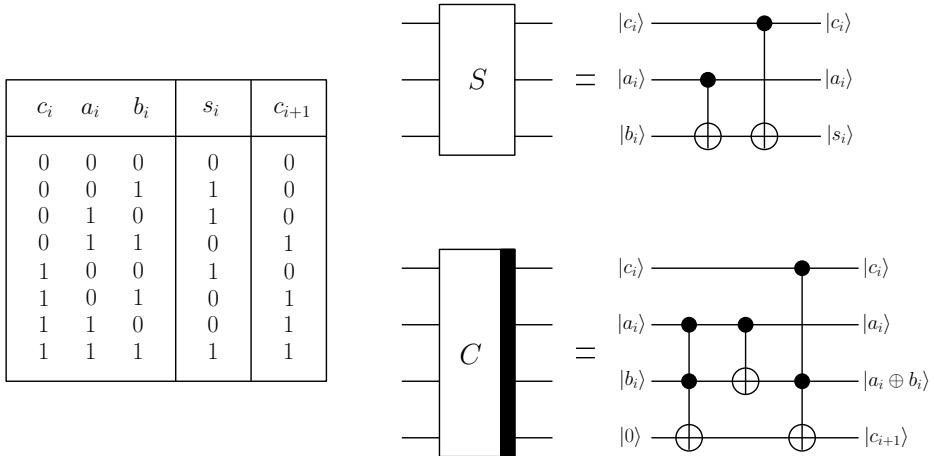


Fig. 3.21 The truth table and the quantum circuit implementation of the functions SUM (upper) and CARRY (lower).

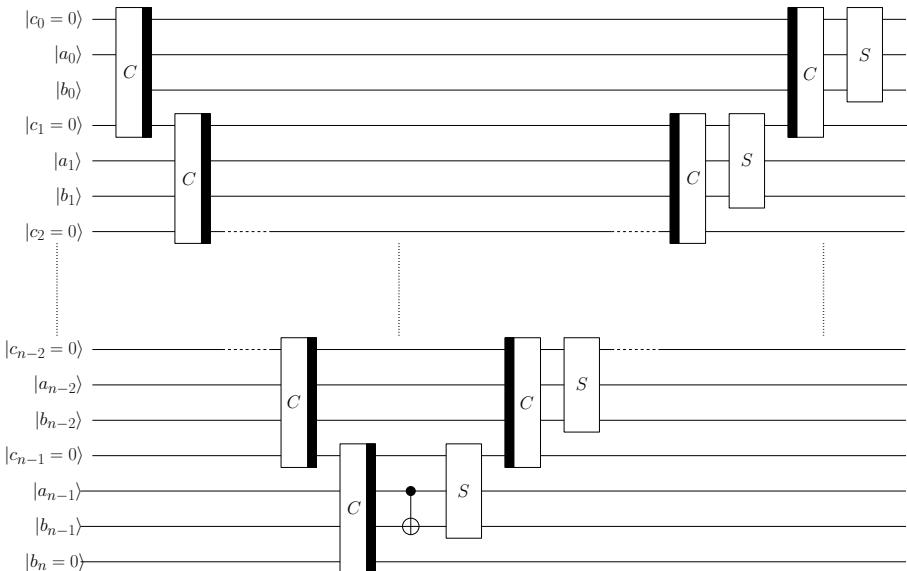


Fig. 3.22 A quantum circuit implementing the addition of two  $n$ -bit integers  $a$  and  $b$ . Note the position of the thick black bars on the CARRY circuits. A CARRY with the bar on the left represents the reverse sequence of elementary gates with respect to a CARRY with the bar on the right.

registers: the first ( $n$  qubits) with input and output  $|a_{n-1}, a_{n-2}, \dots, a_1, a_0\rangle$ , the second ( $n + 1$  qubits) with input  $|0, b_{n-1}, b_{n-2}, \dots, b_1, b_0\rangle$  and output  $|(a+b)_n, (a+b)_{n-1}, (a+b)_{n-2}, \dots, (a+b)_1, (a+b)_0\rangle$ , the third consists of  $n - 1$  qubits, initially in the state  $|0\rangle$ , to which the carries are temporarily written and which are refreshed

to the  $|0\rangle$  state at the end. Each of the first  $n$  CARRY's of the circuit transforms  $|c_i, a_i, b_i, 0\rangle$  into  $|c_i, a_i, a_i \oplus b_i, c_{i+1}\rangle$ . The last carry gives the most significant digit of the sum,  $(a + b)_n$ . Then a single CNOT gate takes  $(a_{n-1}, a_{n-1} \oplus b_{n-1})$  into  $(a_{n-1}, b_{n-1})$  and a SUM operation takes  $(c_{n-1}, a_{n-1}, b_{n-1})$  into  $(c_{n-1}, a_{n-1}, a_{n-1} + b_{n-1})$ . After this we apply CARRY and SUM  $n - 1$  times. As a result, for each qubit  $i$  the sum  $a_i + b_i$  is computed while every ancillary qubit is restored to its initial state  $|0\rangle$ . This is important since it allows us to use these ancillary qubits again for other computations.

### 3.11 Adiabatic theorem

There are certain quantum computation schemes which involve the possibility to perform controllable quantum operations on a given set of qubits, by admitting a slow variation of certain parameters of the system, and then following its dynamical evolution. The key mechanism of such kind of operations is based on the so-called *adiabatic theorem*, a fundamental concept in quantum mechanics originally put forward by Born and Fock (1928), which characterizes the time evolution of a generic conservative system described by a time-dependent Hamiltonian  $H(t)$ . Specifically, the theorem states that, if the system is subjected to a sufficiently slow change in time of the external conditions, it adapts its functional form so to remain in its instantaneous eigenstate.

To be more explicit, let us consider the time-dependent Schrödinger equation of a generic quantum system with a discrete spectrum:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle, \quad (3.130)$$

where we suppose that the Hamiltonian  $H(t)$  changes with time (such change can be large at will, although it occurs gradually and slowly enough). Here we assume to work in the absence of degeneracies for the Hamiltonian eigenvectors. We will postpone the discussion in the presence of degeneracies to Sec. 3.12. The instantaneous Hamiltonian eigenvalue problem can be written according to

$$H(t) |n(t)\rangle = E_n(t) |n(t)\rangle, \quad (3.131)$$

where the eigenstates  $|n(t)\rangle$  obviously depend on time and satisfy the completeness relation  $\langle m(t)|n(t)\rangle = \delta_{mn}$ . Note that, since the Hamiltonian explicitly depends on time, the energy itself is not conserved and the meaning of “level of energy  $E_n(t)$ ” has to be intended only as a formal concept. The generic solution to the time-dependent problem in Eq. (3.130) can be expressed as

$$|\psi(t)\rangle = \sum_n c_n(t) e^{i\theta_n(t)} |n(t)\rangle, \quad (3.132)$$

where

$$\theta_n(t) = -\frac{1}{\hbar} \int_0^t E_n(t') dt' \quad (3.133)$$

is called the *dynamical phase* of the time evolution.

Let us now substitute this solution into Eq. (3.130), so that we get

$$i\hbar \sum_n \left( \frac{dc_n(t)}{dt} |n(t)\rangle + c_n(t) \frac{d}{dt} |n(t)\rangle \right) e^{i\theta_n(t)} = 0, \quad (3.134)$$

where we have simplified the term containing the time derivative of the dynamical phase with the right hand side of Eq. (3.130), since  $d\theta_n(t)/dt = -E_n(t)/\hbar$ . By taking the inner product of Eq. (3.134) with an arbitrary eigenfunction  $\langle m(t)|$  and using the completeness relation of the eigenstates, we arrive at the following:

$$\frac{dc_m(t)}{dt} = - \sum_n c_n(t) e^{i[\theta_n(t) - \theta_m(t)]} \langle m(t)| \frac{d}{dt} |n(t)\rangle. \quad (3.135)$$

Now we can use the adiabatic approximation, under the condition that the time variation of the Hamiltonian is extremely small. In such circumstance, in the right hand side of Eq. (3.135) only the terms with  $m = n$  survive, and we can drastically simplify it to the following expression:

$$\frac{dc_m(t)}{dt} \approx -c_m(t) \langle m(t)| \frac{d}{dt} |m(t)\rangle, \quad (3.136)$$

which can be immediately integrated to give

$$c_m(t) \approx c_m(0) \exp \left\{ - \int_0^t \langle m(t')| \frac{d}{dt'} |m(t')\rangle dt' \right\}. \quad (3.137)$$

The exponential in this equation is usually expressed as  $e^{i\gamma_m(t)}$ , after defining the so-called *geometric phase*

$$\gamma_m(t) = i \int_0^t \langle m(t')| \frac{d}{dt'} |m(t')\rangle dt'. \quad (3.138)$$

It is easy to see that  $\gamma_m(t)$  is a real quantity. Indeed, after differentiating the normalization condition  $\langle m(t)|m(t)\rangle = 1$ , we get:

$$\langle m(t)| \frac{d}{dt} |m(t)\rangle + \text{H.c.} = 0, \quad \text{so that} \quad \text{Re} \left\{ \langle m(t)| \frac{d}{dt} |m(t)\rangle \right\} = 0. \quad (3.139)$$

Therefore, since the integrand function is purely imaginary,  $\gamma_m(t)$  in Eq. (3.138) needs to be a real number.

Finally, inserting the obtained expression (3.137) into the generic solution (3.132) of the time-dependent Hamiltonian eigenvalue problem, we arrive at the final form for the system wave function at time  $t$ :

$$|\psi(t)\rangle = \sum_n c_n(0) e^{i\theta_n(t)} e^{i\gamma_n(t)} |n(t)\rangle, \quad (3.140)$$

where  $\theta_n(t)$  and  $\gamma_n(t)$  respectively denote the dynamical and the geometric phase. This expression indicates that, for an adiabatic process, if the system is initialized in the  $n$ -th eigenstate, it remains there as in a time-independent processes, with the only difference that the two phase factors  $\theta_n(t)$ ,  $\gamma_n(t)$  are picked up. In general, the geometric phase  $\gamma_n(t)$  can be gauged away by a suitable choice of the eigenfunctions, unless the adiabatic evolution is cyclic. In such case,  $\gamma_n(t)$  becomes a gauge-invariant physical quantity, also known as the *Berry phase* (see Berry, 1984).

### 3.11.1 Adiabatic condition

Let us now clarify what we mean by an extremely small time variation of the Hamiltonian, a fact that we used in order to get the expression in Eq. (3.136). Coming back to Eq. (3.135), we can rewrite it as

$$\frac{dc_m(t)}{dt} = - \sum_n c_n(t) \langle m(t) | \frac{d}{dt} |n(t)\rangle e^{-\frac{i}{\hbar} \int_0^t \Delta_{nm}(t') dt'}, \quad (3.141)$$

after defining the energy gap between levels  $n$  and  $m$  at time  $t$  as

$$\Delta_{nm}(t) = E_n(t) - E_m(t). \quad (3.142)$$

It is now useful to find an alternative expression for  $\langle m(t) | \frac{d}{dt} |n(t)\rangle$ . We differentiate Eq. (3.131) with respect to the time and then take the inner product with  $\langle m(t) |$ :

$$\langle m(t) | \frac{dH(t)}{dt} |n(t)\rangle + E_m(t) \langle m(t) | \frac{d}{dt} |n(t)\rangle = E_n(t) \langle m(t) | \frac{d}{dt} |n(t)\rangle + \frac{dE_n(t)}{dt} \delta_{mn},$$

where, in the second term in the left hand side, the Hamiltonian  $H(t)$  has been applied to the left eigenstate  $\langle m(t) |$ . For  $m \neq n$ , this equation can be rewritten as

$$\langle m(t) | \frac{d}{dt} |n(t)\rangle = \frac{1}{\Delta_{nm}(t)} \left\{ \langle m(t) | \frac{dH(t)}{dt} |n(t)\rangle \right\}. \quad (3.143)$$

Therefore Eq. (3.141) can be written as

$$\frac{dc_m(t)}{dt} = -c_m \langle m(t) | \frac{d}{dt} |m(t)\rangle - \sum_{n \neq m} \frac{c_n(t)}{\Delta_{nm}(t)} \langle m(t) | \frac{dH(t)}{dt} |n(t)\rangle e^{-\frac{i}{\hbar} \int_0^t \Delta_{nm}(t') dt'}.$$

An adiabatic evolution is ensured if the instantaneous eigenvectors  $|m(t)\rangle$  of  $H(t)$  (or equivalently the amplitudes  $c_m(t)$ ) evolve independently from each other, that is, their dynamical equations do not couple. The expression that we have found guarantees this requirement, provided that

$$\max_{0 \leq t \leq T} \left| \frac{1}{\Delta_{nm}(t)} \langle m(t) | \frac{dH(t)}{dt} |n(t)\rangle \right| \ll \min_{0 \leq t \leq T} |\Delta_{nm}(t)|, \quad (3.144)$$

where  $T$  is the total evolution time. The interpretation of the adiabaticity condition (3.144) is that for all pairs of energy levels, the expectation value of the temporal rate-of-change of the Hamiltonian  $H(t)$ , in units of the gap, must be small as compared to the gap. Specializing to the case in which the system is prepared, at time  $t = 0$ , in its ground state  $|\psi_0(0)\rangle$  and is let free to evolve under the Hamiltonian  $H(t)$ , the above adiabatic condition translates into:

$$\max_{0 \leq t \leq T} \left| \langle \psi_1(t) | \frac{dH(t)}{dt} | \psi_0(t) \rangle \right| \ll \Delta E_{\min}^2. \quad (3.145)$$

In particular, this implies that the minimum energy gap  $\Delta E_{\min} = \min_{0 \leq t \leq T} \Delta_{10}(t)$  between the instantaneous ground state  $|\psi_0(t)\rangle$  and the first excited state  $|\psi_1(t)\rangle$  cannot be lower than a certain threshold.

### 3.11.2 Berry phase

We shall now be more specific on the meaning of  $\gamma_n(t)$  in Eq. (3.140). Let us take a system initialized into a stationary state of  $H(t) = H(\mathbf{r}(t))$ , *i.e.*, an instantaneous eigenstate of the Hamiltonian at time  $t$ . Here the Hamiltonian is supposed to depend on the time through a set of  $d$  parameters denoted by  $\mathbf{r}(t) = (r_1(t), r_2(t), \dots, r_d(t))$ . We consider a slow variation in time of  $\mathbf{r}(t)$ , such that the adiabatic theorem can be applied. In order to evaluate the geometric phase change of Eq. (3.138), during the dynamics, we first have to notice that a generic eigenstate  $|n(t)\rangle$  depends on  $t$  through the change in time of  $\mathbf{r}(t)$ . As a consequence we can write:

$$\frac{d}{dt}|n(t)\rangle = \frac{\partial|n(t)\rangle}{\partial r_1} \frac{dr_1}{dt} + \dots + \frac{\partial|n(t)\rangle}{\partial r_d} \frac{dr_d}{dt} = \left( \nabla_{\mathbf{r}}|n(\mathbf{r}(t))\rangle \right) \cdot \frac{d\mathbf{r}}{dt}, \quad (3.146)$$

where  $\nabla_{\mathbf{r}} = (\frac{\partial}{\partial r_1}, \frac{\partial}{\partial r_2}, \dots, \frac{\partial}{\partial r_d})$  is the gradient with respect to the parameters  $\mathbf{r}$ . Using this fact, Eq. (3.138) can be easily rewritten as

$$\gamma_n(t) = i \int_{\mathbf{r}_i}^{\mathbf{r}_f} \langle n(\mathbf{r}(t')) | \left( \nabla_{\mathbf{r}}|n(\mathbf{r}'(t))\rangle \cdot d\mathbf{r} \right). \quad (3.147)$$

Now we are interested in explicitly describing the case in which  $\mathbf{r}$  undergoes a closed curve (or loop)  $\mathcal{C}$  in the parameter space, in a given time  $T$ . Clearly after that time the Hamiltonian and its time-varying parameters will come back to its initial shape; the final state of the system will coincide with the initial one apart from the phase factor  $\gamma_n(T)$ . The net geometric phase acquired after the cyclic evolution can be thus expressed as

$$\gamma_n(\mathcal{C}) = \oint_{\mathcal{C}} \mathbf{A}(\mathbf{r})_n \cdot d\mathbf{r}, \quad (3.148)$$

where

$$\mathbf{A}_n(\mathbf{r}) = i \langle n(\mathbf{r}(t)) | \nabla_{\mathbf{r}}|n(\mathbf{r}(t))\rangle \quad (3.149)$$

is a  $d$ -dimensional vector of components  $[\mathbf{A}_n(\mathbf{r})]_\mu = i \langle n(\mathbf{r}) | \frac{\partial n(\mathbf{r})}{\partial r_\mu} \rangle$  (for the ease of notation, we dropped the time dependence in  $\mathbf{r}$ ), denoting the so-called *connection* along the curve  $\mathcal{C}$ . Equation (3.148) is a contour integral along a closed loop in the parameter space, and it is not zero in general. It denotes the so-called Berry phase. Its value only depends on the path taken, and not on how fast it is walked through, provided the adiabatic condition is respected.

To have a more practical feeling about the physical meaning of the expression (3.148), we point out that the connection  $\mathbf{A}_n(\mathbf{r})$  defined in Eq. (3.149) is a vector in the parameter space, which exhibits analogous properties of the vector potential in the electromagnetic field. It is indeed useful to define a gauge-field tensor  $\mathbb{F}_n(\mathbf{r})$  derived from the connection, whose components are:

$$[\mathbb{F}_n(\mathbf{r})]_{\mu\nu} = \frac{\partial[\mathbf{A}_n(\mathbf{r})]_\nu}{\partial r_\mu} - \frac{\partial[\mathbf{A}_n(\mathbf{r})]_\mu}{\partial r_\nu} = i \left[ \left\langle \frac{\partial n(\mathbf{r})}{\partial r_\mu} \middle| \frac{\partial n(\mathbf{r})}{\partial r_\nu} \right\rangle - \left\langle \frac{\partial n(\mathbf{r})}{\partial r_\nu} \middle| \frac{\partial n(\mathbf{r})}{\partial r_\mu} \right\rangle \right], \quad (3.150)$$

also named the *Berry curvature*. We can now cast Eq. (3.148) in a different form, after invoking Stokes' theorem: the integral of a differential form (in our case  $\mathbf{A}_n$ ) over the boundary ( $\mathcal{C}$ ) of some orientable manifold can be seen as the integral of its exterior derivative ( $\mathbb{F}_n$ ) over the whole manifold. The Berry phase can be thus written as a surface integral:

$$\gamma_n(\mathcal{C}) = \frac{1}{2} \int_{\mathcal{S}} dr_\mu \wedge dr_\nu [\mathbb{F}_n(\mathbf{r})]_{\mu\nu}, \quad (3.151)$$

where  $\wedge$  denoted the wedge product, and  $\mathcal{S}$  is an arbitrary manifold enclosed by the curve  $\mathcal{C}$ . It can be verified from Eq. (3.150) that, unlike the Berry vector potential, the Berry curvature is gauge invariant and thus observable.

For the sake of clarity, let us now give a more tangible description and specialize to the three-dimensional parameter space  $\mathbf{r} = (r_x, r_y, r_z)$ , where the previous general approach can be cast into a vector form. Supposing to perform a gauge transformation on the connection  $\mathbf{A}_n \rightarrow \mathbf{A}'_n = \mathbf{A}_n + \nabla \Lambda_n(\mathbf{r})$  (here  $\nabla$  is the usual three-dimensional gradient and  $\Lambda_n$  is a generic scalar function of  $\mathbf{r}$ ), the Berry phase is left unchanged:

$$\gamma'_n = \oint_{\mathcal{C}} \mathbf{A}'_n \cdot d\mathbf{r} = \oint_{\mathcal{C}} \mathbf{A}_n \cdot d\mathbf{r} + \oint_{\mathcal{C}} \nabla \Lambda_n \cdot d\mathbf{r} = \gamma_n, \quad (3.152)$$

since the circuitation of the gradient is zero. The invariance of the Berry phase on the gauge transformation enables us to define a vector  $\mathbf{F}_n$ , such that

$$\mathbf{F}_n = \nabla \times \mathbf{A}_n, \quad \nabla \cdot \mathbf{F}_n = 0. \quad (3.153)$$

This is related to the Berry curvature tensor by  $[\mathbb{F}_n]_{\mu\nu} = \epsilon_{\mu\nu\xi} [\mathbf{F}_n]_\xi$ , where  $\epsilon_{\mu\nu\xi}$  denotes the Levi-Civita antisymmetric tensor, with  $\epsilon_{\mu\nu\xi} = 0$  if the three indices are not all different,  $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$  and  $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$ . The three-dimensional vector form thus provides an intuitive picture of the Berry curvature: it can be viewed as the magnetic field in the parameter space.

Now invoking Stokes' theorem in three dimensions (the so-called Kelvin-Stokes theorem), the circuitation of Eq. (3.148) can be seen as the flux of the field  $\mathbf{F}_n$  across a two-dimensional surface  $\mathcal{S}$ , which is contained into the closed path  $\mathcal{C}$ :

$$\gamma_n = \oint_{\mathcal{C}} \mathbf{A}_n \cdot d\mathbf{r} = \int_{\mathcal{S}} (\nabla \times \mathbf{A}_n) \cdot d\mathbf{S} = \int_{\mathcal{S}} \mathbf{F}_n \cdot d\mathbf{S}. \quad (3.154)$$

The Berry phase can thus be seen as the flux of a pseudo-magnetic field, as it happens in the Aharonov–Bohm effect, where an electrically charged particle is affected by a magnetic field, despite being confined to a region in which such field is zero (see Aharonov and Bohm, 1959).

#### Example: a qubit in a magnetic field

We now give a simple example of physical system which exhibits a Berry phase, namely a single qubit of magnetic moment  $\mu$  that is placed in a magnetic field  $\mathbf{H}(t) = (H_x(t), H_y(t), H_z(t))$ . The system is characterized by the Hamiltonian

$$H(t) = -\mu [H_x(t)\sigma_x + H_y(t)\sigma_y + H_z(t)\sigma_z], \quad (3.155)$$

where  $\sigma_x, \sigma_y, \sigma_z$  denote the three spin-1/2 Pauli matrices, and  $\mu$  the gyromagnetic ratio of the qubit. Let us now vary cyclically the trajectory of the field  $\mathbf{H}$  on a sphere at fixed radius, as depicted in Fig. 3.23. Its three-dimensional components at the time  $t$  in the spherical coordinates  $r, \theta, \phi$  are given by:

$$\mathbf{H}(t) = (H_0, \theta, \omega t), \quad (3.156)$$

where  $H_0$  denotes its modulus, and  $\omega$  is the angular velocity. The latter quantity needs to be sufficiently small, in order to ensure the adiabatic condition. By diagonalizing  $H(t)$  (see exercise 3.17), we can find that this is guaranteed in the limit  $\omega \ll 2\mu H_0/\hbar$ .

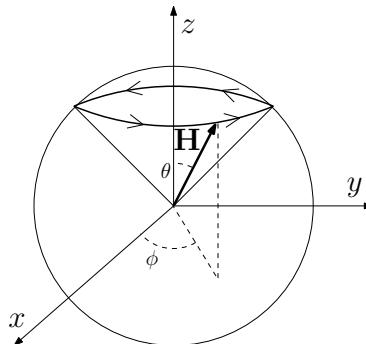


Fig. 3.23 Time evolution of the cyclic trajectory of the magnetic field  $\mathbf{H}(t)$ . Since its modulus is constant in time, the evolution is restricted on the surface of a sphere of radius  $H_0 = |\mathbf{H}|$ , and can be parametrized by the two angular coordinates  $(\theta, \phi)$ .

Moreover it is simple to evaluate its instantaneous eigenstates, which can be explicitly written in the  $\sigma_z$  basis of the qubit as

$$|\psi_1\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi(t)} \sin \frac{\theta}{2} |1\rangle, \quad (3.157)$$

$$|\psi_2\rangle = \sin \frac{\theta}{2} |0\rangle - e^{i\phi(t)} \cos \frac{\theta}{2} |1\rangle. \quad (3.158)$$

From their expression and using the definition (3.149), it is easy to calculate their connections in the parameter space  $(\theta, \phi)$ :

$$\mathbf{A}_1 = -\frac{1}{2} \begin{bmatrix} 0 \\ 1 - \cos \theta \end{bmatrix}, \quad \mathbf{A}_2 = -\frac{1}{2} \begin{bmatrix} 0 \\ 1 + \cos \theta \end{bmatrix}. \quad (3.159)$$

Using Eq. (3.148), it is thus immediate to evaluate the Berry phases associated to the two eigenstates:

$$\gamma_1 = \oint_C \mathbf{A}_1 \cdot d\mathbf{r} = \int_0^{2\pi} [\mathbf{A}_1]_\phi d\varphi = -\pi(1 - \cos \theta), \quad (3.160)$$

$$\gamma_2 = \oint_C \mathbf{A}_2 \cdot d\mathbf{r} = \int_0^{2\pi} [\mathbf{A}_2]_\phi d\varphi = -\pi(1 + \cos \theta). \quad (3.161)$$

Note that  $\gamma_1$  and  $\gamma_2$  do not depend on either the energy or the angular velocity  $\omega$ , but only on the path that has been adiabatically encircled after a time  $T = 2\pi/\omega$ .

### Discretized formulation of the Berry phase

The basic scenario where one can grasp the fundamental geometric importance of the Berry phase probably resides in its discrete formulation. Let us first consider two vectors  $|\psi_1\rangle$  and  $|\psi_2\rangle$  in the parameter space, and calculate their phase difference:

$$e^{-i\Delta\phi_{12}} = \frac{\langle\psi_1|\psi_2\rangle}{|\langle\psi_1|\psi_2\rangle|} \quad \text{such that } \Delta_{12} = -\text{Im} \ln \langle\psi_1|\psi_2\rangle. \quad (3.162)$$

The second equality comes from the fact that, for a complex number  $z = |z|e^{i\phi}$ , the expression  $\phi = \text{Im} \ln z$  just takes its complex phase discarding the magnitude. Clearly  $\Delta\phi_{12}$  depends on the arbitrary choice of phases of the individual vectors. However the situation is drastically different if we consider three vectors and calculate the change of phase in the triangle in the parameter space connecting them:

$$\Delta\phi_{\text{TOT}} = \Delta\phi_{12} + \Delta\phi_{23} + \Delta\phi_{31} = -\text{Im} \ln [\langle\psi_1|\psi_2\rangle \langle\psi_2|\psi_3\rangle \langle\psi_3|\psi_1\rangle]. \quad (3.163)$$

As a matter of fact, this latter quantity is independent of the choice made for the global phases of the three vectors, since each of them appears both as a “ket” and as a “bra” in the expression (3.163). Suppose indeed to introduce a new set of vectors  $|\tilde{\psi}_j\rangle = e^{-i\beta_j}|\psi_j\rangle$ , with  $\beta_j$  real; this is a gauge transformation, in the traditional sense where a particular choice of gauge may influence the intermediate results of a calculation, but should not affect any physically meaningful prediction. The global phase difference  $\Delta\phi_{\text{TOT}}$ , obtained after connecting the three vectors in a loop, is left unaffected by this gauge transformation, and represents the simplest discretized version of the Berry phase, which was defined in the continuum in Eq. (3.148).

At this point, it becomes straightforward to generalize this scenario to a closed polygonal path in the parameter space with  $m$  vectors  $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle$ , such that

$$\Delta\phi_{\text{TOT}}^{(m)} = \sum_{j=1}^m \Delta\phi_{j,j+1} = -\text{Im} \ln \left[ \prod_{j=1}^m \langle\psi_j|\psi_{j+1}\rangle \right], \quad (3.164)$$

where we supposed that  $\Delta\phi_{m,m+1} \equiv \Delta\phi_{m,1}$  and  $|\psi_{m+1}\rangle \equiv |\psi_1\rangle$ . The phase difference accumulated along the polygonal path does not depend on the phases of the vectors, which always cancel in a closed loop. We also mention that there is no need to assume that neither the phases nor the loop are smooth. Differentiability is requested only in the continuum limit. However the discretization of the contour integrals like the one in Eq. (3.148) remains an important step to be done for any successful numerical strategy for coping with geometric phases. When calculating the Berry phase of a stationary eigenstate  $|n(\mathbf{r}(t))\rangle$  of  $H(t)$ , in practice one has to discretize the time and to find the (approximated) eigenstates from numerical diagonalization of the Hamiltonian. The gauge is thus arbitrarily taken by the diagonalization routine; nonetheless arbitrary fluctuations of the gauge phase will not affect the final result for  $\gamma_n$ .

We finally propose an example with three quantum states in a two-dimensional Hilbert space:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |\psi_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ e^{i2\pi/3} \end{bmatrix}, \quad |\psi_3\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ e^{i4\pi/3} \end{bmatrix}. \quad (3.165)$$

In this case, it is easy to show that the Berry phase is non trivial, since

$$\Delta\phi_{\text{TOT}} = -\text{Im} \ln \left[ \frac{1}{2^{3/2}} (1 + e^{2\pi i/3})(1 + e^{2\pi i/3})(1 + e^{-4\pi i/3}) \right] = -\text{Im} \ln \left( -\frac{1}{2^{3/2}} \right) = \pi.$$

Note that, for real vectors, the global product in Eq. (3.164) would be always real and thus  $\Delta\phi_{\text{TOT}}$  would be either 0 or  $\pi$ , depending on the sign of that product. The fact that, in this specific example,  $\Delta\phi_{\text{TOT}} = \pi$  is due to the special choice of the three states in Eq. (3.165).

### 3.12 \* Non-Abelian geometric phase

Until now we considered the case in which the Hamiltonian spectrum of the system is non degenerate. In obtaining the geometric phase for an adiabatically evolving system, the assumption that the eigenspace to which the prepared state belongs is non-degenerate was crucial. Specifically from Eq. (3.140) we see that, taking as initial condition a given eigenstate  $|\psi(0)\rangle = |n\rangle$  and in the absence of degeneracies, when a loop in the parameter space is traversed, the final and initial states are proportional, so that the net effect of the evolution is merely a phase.

Conversely, if a degenerate eigenspace is assumed, we will show that it is possible to conceive a much wider variety of evolutions, with a slightly more complex structure (see Wilczek and Zee, 1984). The generalization of the geometric phase in this context is called *holonomy*, and turns out to be at the basis of the *holonomic quantum computation* as we shall see in Sec. 3.15. The word “holonomy” refers to the set of all the loops on a manifold, starting and ending in the same point. This set has the structure of group (indeed, one can define the composition of two loops by joining the end point of one loop with the starting point of the other; the identity element is the trivial loop with only one point; the inverse of a loop is the same traversed in the opposite direction). The geometric phases in the absence of level degeneracies form a representation of an holonomy group: any loop in the parameter space of an Hamiltonian is associated with a geometric phase factor, that is simply a complex number of unitary modulus. Since the phases trivially commute, that is,  $e^{i\gamma_1}e^{i\gamma_2} = e^{i\gamma_2}e^{i\gamma_1}$ , they form an Abelian representation. For this reason, the phase (3.138) derived in Sec. 3.11 is referred to as an *Abelian geometric phase*.

The non-Abelian generalizations of this construction are not represented by ordinary numbers, but by matrices. This naturally emerges in adiabatic evolving systems, when eigenspaces are degenerate. Here we do not want to give a fully comprehensive discussion of holonomies, but only sketch their basic principles which will be useful to describe the working mechanism of adiabatic quantum computation. We remind the interested reader to more specialized literature (see the guide to the bibliography at the end of the chapter).

As we did in Sec. 3.11.2, let us suppose that, along the time evolution, a set of  $d$  parameters  $\mathbf{r}(t) = (r_1(t), r_2(t), \dots, r_d(t))$  is varied. Now we admit the chance that the Hamiltonian eigenvectors can be degenerate, and denote with  $\{|n_k(\mathbf{r})\rangle\}_{k=1,\dots,g_n}$  the set of instantaneous eigenstates associated to the energy  $E_n(\mathbf{r})$  (for the sake of clarity, here we removed the implicit dependence on time of the set of parameters  $\mathbf{r}$ ). The index  $k$  is called degeneracy index (ranging from 1 to the dimension of the degenerate subspace, say  $g_n$ ), and the orthonormality relation reads  $\langle m_q(\mathbf{r})|n_k(\mathbf{r})\rangle = \delta_{mn}\delta_{qk}$ . We also introduce the projector operators  $P_n(\mathbf{r})$  on the instantaneous  $n$ -th eigenspace (with  $n = 1, \dots, N$ ):

$$P_n(\mathbf{r}) = \sum_{k=1}^{g_n} |n_k(\mathbf{r})\rangle\langle n_k(\mathbf{r})|, \quad (3.166)$$

so that the system Hamiltonian can be written in the compact form

$$H(\mathbf{r}) = \sum_{n=1}^N E_n(\mathbf{r}) P_n(\mathbf{r}). \quad (3.167)$$

Here we supposed that the instantaneous dimensions  $g_n$  of the various eigenspaces do not change in time, a possibility which is forbidden by the adiabatic requirements. This property is called *iso-spectrality* for the class of Hamiltonians  $H(\mathbf{r})$ . Provided the projectors  $P_n(\mathbf{r})$  depend smoothly on time, it implies the absence of level crossing at any time:  $n \neq m \Rightarrow E_n(\mathbf{r}(t)) \neq E_m(\mathbf{r}(t)) \forall t$ .

Considering a system with these characteristics, it is possible to prove a generalized version of the adiabatic theorem, which can be formalized according to the following. When the control parameters  $\mathbf{r}(t)$  are driven along a loop  $\mathcal{C}$  adiabatically (that is, slowly with respect to any time scale associated to the system dynamics), a generic initial state  $|\psi(0)\rangle$  will be mapped after the period  $T$  into

$$|\psi(T)\rangle = \bigoplus_{n=1}^N \left( e^{i\theta_n(T)} G_{\mathbb{A}_n}(\mathcal{C}) \right) |\psi(0)\rangle, \quad (3.168)$$

where  $\oplus$  indicates the direct sum over the different eigenspaces, so that each of them follows a distinct evolution in time. In this expression, the term

$$\theta_n(T) = -\frac{1}{\hbar} \int_0^T E_n(\mathbf{r}(t')) dt' \quad (3.169)$$

denotes the usual dynamical phase, while  $G_{\mathbb{A}_n}(\mathcal{C})$  are matrices of dimension  $g_n \times g_n$ , which represent the geometric contributions leading to the so-called *non-Abelian Berry phase*. These matrices are unitary mappings of the  $n$ -th eigenspace and can be expressed through the following path ordered integrals

$$G_{\mathbb{A}_n}(\mathcal{C}) = \overleftarrow{\mathcal{P}} \exp \left\{ \oint_{\mathcal{C}} \mathbb{A}_n \cdot d\mathbf{r} \right\}, \quad (3.170)$$

where  $\overleftarrow{\mathcal{P}}$  forces the integral along  $\mathcal{C}$  to be taken in a rigorous sequence according to the variation of the parameters along the loop, that is by ordering the operators

from the right to left. The  $d$ -dimensional array of matrices  $\mathbb{A}_n$  denotes the so-called *adiabatic connection forms*, which are given by

$$[\mathbb{A}_n]_{qk} = \langle n_q(\mathbf{r}) | \nabla_{\mathbf{r}} | n_k(\mathbf{r}) \rangle, \quad (q, k = 1, \dots, g_n), \quad (3.171)$$

thus generalizing the connection in Eq. (3.149) to the case of degenerate eigenspaces. At this point, let us stress that the path ordering in Eq. (3.170) is needed because of the non-commutativity of the operators  $\mathbb{A}_n$  for the different values of the parameters. We also remark that, due to the decomposition of the evolution operator in Eq. (3.168) with a direct sum, it is straightforward to restrict the study to a given degenerate eigenspace with fixed  $n$ .

Let us now provide a sketch of the proof leading to Eqs. (3.168)–(3.171). We will not enter the details, since they would require a rather formal discussion and are not necessary for the rest of the book. If the Hamiltonians  $H(\mathbf{r})$  satisfy the iso-spectrality property defined above, it is possible to connect them through a unitary transformation  $\mathcal{U}(\mathbf{r})$  such that

$$H(\mathbf{r}) = \mathcal{U}(\mathbf{r}) H_0 \mathcal{U}^\dagger(\mathbf{r}), \quad \text{with } H_0 = H(\mathbf{r}(0)). \quad (3.172)$$

The time-dependent Schrödinger equation (3.130) in the presence of degenerate eigenspaces can then be solved by discretizing the time interval  $[0, T]$  into  $N_T$  small segments  $\Delta t$  and trivially integrating it in the time slices. The global unitary evolution operator  $U_c(T)$  such that  $|\psi(T)\rangle = U_c(T)|\psi(0)\rangle$  can be written as

$$\begin{aligned} U_c(T) &= \overleftarrow{\mathcal{T}} \exp \left\{ -\frac{i}{\hbar} \int_0^T \mathcal{U}(\mathbf{r}) H_0 \mathcal{U}^\dagger(\mathbf{r}) dt \right\} = \overleftarrow{\mathcal{T}} \lim_{N_T \rightarrow \infty} \exp \left\{ -\frac{i}{\hbar} \sum_{j=1}^{N_T} \mathcal{U}_j H_0 \mathcal{U}_j^\dagger \Delta t \right\} \\ &= \overleftarrow{\mathcal{T}} \lim_{N_T \rightarrow \infty} \prod_{j=1}^{N_T} \mathcal{U}_j e^{-\frac{i}{\hbar} H_0 \Delta t} \mathcal{U}_j^\dagger, \end{aligned} \quad (3.173)$$

where we defined  $\mathcal{U}_j = \mathcal{U}(\mathbf{r}_j)$ , with  $\mathbf{r}_j = \mathbf{r}(t_j)$ , and we denoted with  $\overleftarrow{\mathcal{T}}$  the time ordering operator which forces the integral to be performed as a time sequence of operators which are ordered in time from the right to the left. Note that the last equality is valid in the limit  $\Delta t \rightarrow 0$ . This expression can be simplified by approximating two successive unitary operators with an infinitesimal rotation of the form  $\Delta \mathbf{r}_j = \mathbf{r}_{j+1} - \mathbf{r}_j$ :

$$\mathcal{U}_j^\dagger \mathcal{U}_{j+1} \approx I + \mathbb{A}_j \cdot \Delta \mathbf{r}_j, \quad \text{where } \mathbb{A}_j = \mathcal{U}_j^\dagger \frac{\Delta \mathcal{U}_j}{\Delta \mathbf{r}_j}. \quad (3.174)$$

In this way, we can rewrite the evolution operator as

$$U_c(T) = \overleftarrow{\mathcal{T}} \lim_{N_T \rightarrow \infty} \mathcal{U}_{N_T} \left( I - i H_0 N_T \Delta t + \sum_{j=1}^{N_T-1} \mathbb{A}_j \cdot \Delta \mathbf{r}_j \right) \mathcal{U}_1^\dagger. \quad (3.175)$$

Since we are considering a closed circuit, it is obvious that  $\mathcal{U}_1 = \mathcal{U}_{N_T}$ . After a re-parametrization, these unitaries can be taken to be equal to the identity.

Without loss in generality, let us now suppose that the initial state  $|\psi(0)\rangle$  belongs to the eigenspace associated to the energy  $E_0 = 0$  (the iso-spectrality condition forces this eigenspace to decouple from all the others). If the energy is different from zero, one can show that the results would be unaltered, apart from the insertion of a trivial

dynamical phase. Due to the time ordering in Eq. (3.175), the actions of the Hamiltonian  $H_0$  and of the connection  $\mathbb{A}$  are alternated, hence it is not possible to separate them into two exponentials. However the adiabatic condition ensures that, at each time  $t_j$ , the instantaneous wave function  $|\psi(t_j)\rangle$  will remain in the  $E_0 = 0$  eigenspace, and therefore the action of  $H_0$  can be factored out, thus obtaining

$$U_C(T) = \overleftarrow{\lim}_{N_T \rightarrow \infty} \left( I + \sum_{j=1}^{N_T-1} \mathbb{A}_j \cdot \Delta \mathbf{r}_j \right). \quad (3.176)$$

To conclude our proof, we need to show that this evolution operator is simply the time-discretized version of that in Eq. (3.170). Indeed the direct-sum structure of Eq. (3.168) directly comes from the iso-spectrality, while in the subspace  $E_{n_0} = 0$  that we are considering, the dynamical phase  $\theta_{n_0}$  is simply zero). We first note that from Eq. (3.172) we have that  $H(\mathbf{r}_{j+1}) \mathcal{U}_j = \mathcal{U}_j H(\mathbf{r}_j)$ , which implies

$$\mathcal{U}_j P_n(\mathbf{r}_j) = P_n(\mathbf{r}_{j+1}) \mathcal{U}_j. \quad (3.177)$$

In terms of the instantaneous eigenbasis on the relevant eigenspace, the operator  $\mathcal{U}_j$  can be rewritten as

$$\mathcal{U}_j = \sum_{k=1}^{g_n} |n_k(\mathbf{r}_{j+1})\rangle \langle n_k(\mathbf{r}_j)|, \quad (3.178)$$

so that we finally obtain, in the limit  $\Delta t \rightarrow 0$ , the expression

$$[\mathbb{A}_j]_{qk} = \left[ \mathcal{U}_j^\dagger \frac{\Delta \mathcal{U}_j}{\Delta \mathbf{r}_j} \right]_{qk} = \langle n_q(\mathbf{r}_{j+1}) | \nabla_{\mathbf{r}} | n_k(\mathbf{r}_{j+1}) \rangle. \quad (3.179)$$

### Discretized formulation of the non-Abelian Berry phase

In the case of a single non-degenerate Hamiltonian eigenstate  $|n(\mathbf{r})\rangle$ , we discussed previously how it is possible to generalize the notion of the Berry phase (3.148) to a discretized formulation, as in Eq. (3.164). It is possible to extend the discussion to the case of a degenerate eigenspace  $\{|n_k(\mathbf{r})\rangle\}_{k=1,\dots,g_n}$ , in such a way as to discretize the non-Abelian Berry phase of Eq. (3.170).

To this purpose, we first need to construct, at each step in the loop from point  $\mathbf{r}_j$  to point  $\mathbf{r}_{j+1}$ , the  $g_n \times g_n$  overlap matrix

$$[\mathbb{S}_j]_{qk} = \langle n_q(\mathbf{r}_j) | n_k(\mathbf{r}_{j+1}) \rangle. \quad (3.180)$$

In direct analogy with Eq. (3.164), it can be shown that the total Berry phase along a closed path in the parameter space  $\mathbf{r}(t)$  can be written as

$$\Delta\phi_{\text{TOT}}^{(m)} = -\text{Im} \ln \left[ \prod_{j=1}^m \det \mathbb{S}_j \right] = -\text{Im} \ln \det \left[ \prod_{j=1}^m \mathbb{S}_j \right], \quad (3.181)$$

where  $\det[\cdot]$  denotes the determinant of a matrix. Even in this case, it is possible to define a new matrix  $\mathbb{S}'_j = U_j^{-1} \mathbb{S}_j U_j$  (where  $U_j$  is a unitary matrix), which corresponds to a gauge transformation that leaves the cyclic product invariant and the discretization of the path integral stable.

### 3.13 Adiabatic quantum computation

The adiabatic theorem discussed in Sec. 3.11 provides the theoretical framework for the so-called *adiabatic quantum computation* (AQC) schemes. Basically these schemes assume to encode the solution of a given problem in a certain eigenstate of a suitable Hamiltonian  $H_F$ , usually the ground state. The Hamiltonian may describe a very complex system and it is usually very difficult to find its eigenmodes (for example, think about a spin glass). The goal of the computation is finding such ground state by performing an adiabatic connection (if possible) of the Hamiltonian  $H_F$  with another one  $H_I$ , which typically describes a much simpler physical system.

More specifically, suppose to initialize the system into the ground state of  $H_I$ , at  $t = 0$ . At a given time  $t$ , this is described by the instantaneous Hamiltonian

$$H(t) = \left(1 - \frac{t}{T}\right) H_I + \left(\frac{t}{T}\right) H_F, \quad \text{with } 0 \leq t \leq T. \quad (3.182)$$

If we are able to follow the quantum dynamics of  $H(t)$  for a sufficiently large control time  $T$ , the adiabatic theorem guarantees that the system will remain in its instantaneous ground state and eventually reach, at  $t = T$ , the solution (that is, the ground state of  $H_F$ ). We stress that, as discussed in Sec. 3.11.1, the adiabaticity condition (3.145) requires the instantaneous gap  $\Delta E(t)$  between the ground state and the first excited state to be not too small. This fact eventually dictates the speed of evolution, or the time  $T$  in Eq. (3.182), which has to satisfy the constraint:

$$T \gg \frac{1}{\Delta E_{\min}^2} \max_{0 \leq s \leq 1} \left| \langle \psi_1(s) | \frac{dH(s)}{ds} | \psi_0(s) \rangle \right|, \quad (3.183)$$

where we defined the rescaled variable  $s(t) = t/T$ , with  $0 \leq s \leq 1$ . If the position of the minimum gap in  $H(t)$  is known, it is possible to considerably reduce the computational time  $T$  by suitably modulating the speed of the evolution, thus modifying the linear interpolation scheme of Eq. (3.182). We will discuss more in detail this modification in Sec. 4.2.4.

The protocol described above can be cleverly employed to minimize the energy cost function of an arbitrary interacting system. There are however cases in which the adiabatic connection between the problem Hamiltonian  $H_F$  and any other “easy” Hamiltonian  $H_I$  is increasingly hard. This fact reflects into the computational complexity of AQC, an issue which can be quantified in terms of the overall time  $T$  of the evolution. It can be shown that such time typically depends on the number  $N$  of qubits the Hamiltonian  $H(t)$  acts on: the larger the number  $N$ , the longer the evolution will last. In particular, at certain points of the evolution, the energy gap may become very small (actually, it goes to zero for  $N \rightarrow \infty$ ). Under such circumstances, the system undergoes a quantum phase transition characterized by a singular behaviour.

Below we provide three basic examples of systems with a limited number of qubits, where it is possible to employ the adiabatic time evolution in order to solve some computational tasks.

Example 1

We begin by considering a single-qubit system. Our goal is to find the ground state of the following Hamiltonian, written in the computational basis  $\{|0\rangle, |1\rangle\}$ :

$$H_F^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (3.184)$$

As the initial Hamiltonian, let us take

$$H_I^{(1)} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{having a ground state } |\psi_0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (3.185)$$

The two Hamiltonians can be adiabatically connected in time by

$$H(t) = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \left(1 - \frac{t}{T}\right) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{t}{T}\right), \quad (3.186)$$

since the instantaneous gap between the ground-state and the first-excited-state energies  $E_0$  and  $E_1$  remains large, as shown in the upper panel of Fig. 3.24. Specifically, the gap satisfies the bound  $\Delta E \geq 1/\sqrt{2}$ , the equality holding at  $t/T = 1/2$ . As a consequence, the system at time  $t$  will remain in the instantaneous ground state  $|\psi(t)\rangle$  of  $H(t)$  and eventually reach the ground state of  $H_F^{(1)}$ , that is, a state in which the qubit is fully polarized along the  $z$ -axis. The working mechanism of AQC is explicitly shown in the lower panel of Fig. 3.24, where we plot the probabilities to obtain the outcome 0 or 1, respectively  $p_0 = |\langle 0|\psi(t)\rangle|^2$  and  $p_1 = |\langle 1|\psi(t)\rangle|^2$ .

Example 2

Consider now a two-qubit system. Suppose we want to find an adiabatic protocol to achieve the ground state of the Hamiltonian

$$H_F^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3.187)$$

where again we used the computational basis for the two-qubit system  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ . We consider an initial configuration where the two qubits are independently governed by the Hamiltonian defined in Eq. (3.185), that is,

$$H_I^{(2)} = H_I^{(1)} \otimes \mathbb{I}^{(1)} + \mathbb{I}^{(1)} \otimes H_I^{(1)} = \frac{1}{2} \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}, \quad (3.188)$$

where  $\mathbb{I}^{(1)}$  is the identity matrix for a single qubit. Its ground state is simply the two-qubit generalization of the previous initial state for a single qubit, that is,  $|\psi_0\rangle = \frac{1}{4}[1 \ 1 \ 1 \ 1]^T$ . Let us follow the time evolution of the system governed by the Hamiltonian in Eq. (3.182), which interpolates between  $H_I^{(2)}$  and  $H_F^{(2)}$ .

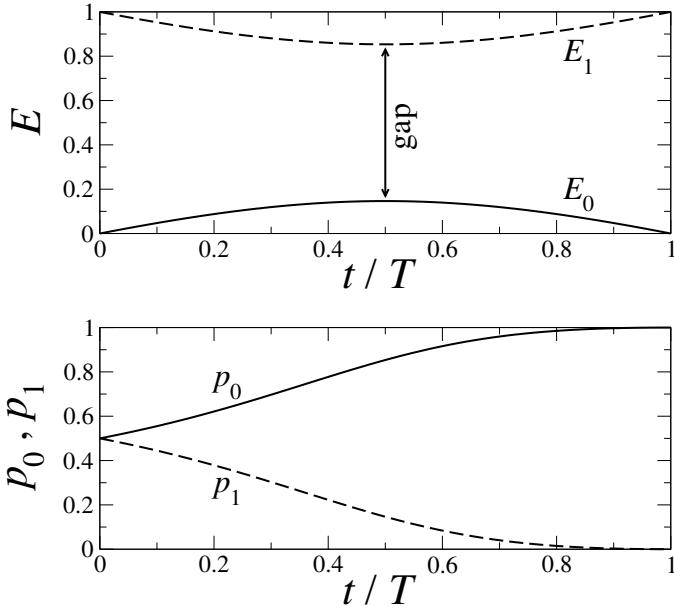


Fig. 3.24 Upper panel: the two eigenvalues of the Hamiltonian in Eq. (3.186) as a function of the time  $t$ . It is evident that there is a big energy gap, therefore it is possible to apply the adiabatic theorem. Lower panel: the probabilities  $p_1$  and  $p_2$  to obtain, respectively, outcome 0 or 1 from the measurement of the qubit polarization along the  $z$ -axis as a function of the time.

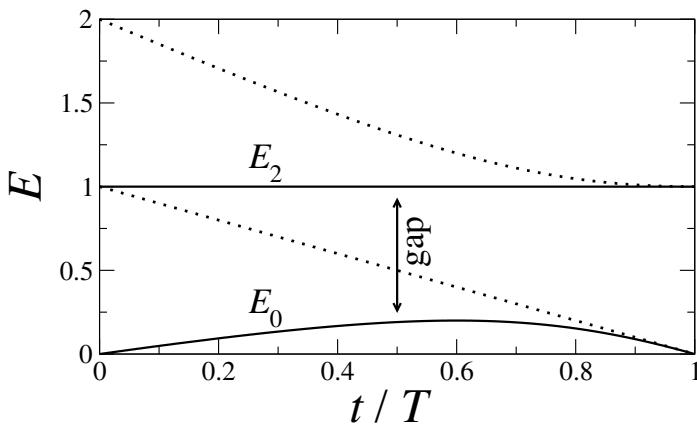


Fig. 3.25 The two eigenvalues of the Hamiltonian in Eq. (3.182) which adiabatically interpolates between (3.188) and (3.187), as a function of the time  $t$ . Dotted lines indicate the irrelevant states to the adiabatic evolution, due to symmetry. The relevant gap is that which separates the two continuous lines; it is shown that this is always large during the time evolution.

We notice that, while the ground state of  $H_I^{(2)}$  is unique, there are two ground states of  $H_F^{(2)}$ , which are  $|\phi_1\rangle = [0 \ 1 \ 0 \ 0]^T$  and  $|\phi_2\rangle = [0 \ 0 \ 1 \ 0]^T$ . This fact is clearly visible from the structure of the energy levels plotted in Fig. 3.25, where we

observe that, while at the initial point  $t = 0$  there is a single ground-state level, at the final point  $t = T$  we see the crossing of two lowest-energy levels. The system Hamiltonian however exhibits the following symmetry in the  $z$ -basis:

$$|00\rangle \rightarrow |11\rangle; \quad |01\rangle \rightarrow |10\rangle; \quad |10\rangle \rightarrow |01\rangle; \quad |11\rangle \rightarrow |00\rangle, \quad (3.189)$$

that is,  $H(t)$  commutes with the operator  $\sigma_x \otimes \sigma_x$ . Since the initial state  $|\psi_0\rangle$  is invariant under the above qubit-exchange operation, the dynamics ruled by  $H(t)$  preserves such symmetry and therefore the final state after the adiabatic time evolution needs to be the symmetric state  $|\psi(t)\rangle = \frac{1}{\sqrt{2}}(|\phi_1\rangle + |\phi_2\rangle)$ . The other final state orthogonal to it,  $|\psi(t)\rangle_{\perp} = \frac{1}{\sqrt{2}}(|\phi_1\rangle - |\phi_2\rangle)$ , follows the evolution of an antisymmetric state. The dynamics forbids any transition from the symmetric to the antisymmetric states, therefore the first excited level plotted in Fig. 3.25 is irrelevant to the adiabatic evolution of  $|\psi_0\rangle$ , and the relevant gap is  $E_2(t) - E_0(t)$ .

### Example 3

The celebrated satisfiability problem ( $k$ -SAT) can be solved by AQC, following the ideas put forward in the two previous examples. As discussed in Chap. 1, the goal is to find a set of literals which satisfies a given boolean expression  $f(\mathbf{a})$ , where  $\mathbf{a} = (a_1, \dots, a_N)$  denotes a set of  $N$  classical bits  $a_j = (0, 1)$ . The formula can be written in the conjunctive normal form as a combination of  $M$  clauses through the AND operator, each of them involving up to  $k$  bits.

We shall consider a system of  $N$  qubits. The initial Hamiltonian  $H_I^{(N)}$  is constructed by generalizing the two previous examples. Namely, we suppose that each qubit is coupled to a magnetic field along the  $x$ -direction and is governed by a Hamiltonian  $H_I^{(j)}$  as the one in Eq. (3.185), so that

$$H_I^{(N)} = \sum_{j=1}^N \mathbb{I}_1^{(1)} \otimes \mathbb{I}_2^{(1)} \otimes \dots \otimes \mathbb{I}_{j-1}^{(1)} \otimes H_I^{(1)} \otimes \mathbb{I}_{j+1}^{(1)} \otimes \dots \otimes \mathbb{I}_N^{(1)}, \quad (3.190)$$

where  $\mathbb{I}_k^{(1)}$  is the identity matrix for the  $k$ -th qubit. Thus the ground state of  $H_I^{(N)}$  is a uniform superposition of states corresponding to all possible assignments of the  $N$ -qubit values in the  $z$ -basis  $\{|0\rangle, |1\rangle\}$ :

$$|\psi_0\rangle = \frac{1}{\sqrt{2^N}} \sum_{\{a_j\}} |a_1\rangle |a_2\rangle \dots |a_N\rangle. \quad (3.191)$$

The final Hamiltonian  $H_F^{(N)}$  is chosen in such a way that its ground state encodes the solution to the instance given by  $f(\mathbf{a})$ . To this aim, we first define a classical energy function  $h(\mathbf{a})$  that is a sum of energy functions

$$h_C(a_\alpha, a_\beta, \dots) = \begin{cases} 0 & \text{if clause } C \text{ is satisfied,} \\ 1 & \text{if clause } C \text{ is violated,} \end{cases} \quad (3.192)$$

where  $a_\alpha, a_\beta, \dots$  are the bits involved in clause  $C$ . Then we construct the function

$$h(\mathbf{a}) = \sum_C h_C, \quad (3.193)$$

meaning that the energy cost of any bit assignment  $\mathbf{a}$  equals the number of clauses that the assignment violates. Clearly  $h \geq 0$  and  $h(\tilde{\mathbf{a}}) = 0$  if and only if  $\tilde{\mathbf{a}} = (\tilde{a}_1, \dots, \tilde{a}_N)$  satisfies all the clauses. The Hamiltonian  $H_F^{(N)}$  is defined by turning this classical function into a quantum operator that is diagonal in the  $z$ -basis:

$$H_F^{(N)}|a_1\rangle|a_2\rangle\dots|a_N\rangle = h(a_1, a_2, \dots, a_N)|a_1\rangle|a_2\rangle\dots|a_N\rangle, \quad (3.194)$$

so that its ground state corresponds to the bit assignment that violates the minimal number of clauses. If its energy is zero, than we know that the formula has a satisfying assignment. Note that, if more than one assignment minimizes the number of violations, then there will be more than one ground state for  $H_F^{(N)}$ , and the AQC scheme will anyway converge toward the degenerate energy minimum.

We have thus shown that, in this context, solving a k-SAT problem is equivalent to finding the ground state of a Hamiltonian. Clearly many other computationally interesting problems can be recast in this form.

**Exercise 3.22** Consider the following boolean function:

$$f(a, b) = (\bar{a} \vee b) \wedge a \wedge (\bar{a} \vee \bar{b}) \wedge \bar{b}. \quad (3.195)$$

Show that it is possible to solve the associated satisfiability problem by means of AQC. Build up explicitly the initial and the final Hamiltonians for the protocol, and compute the energy levels for the time-dependent connecting Hamiltonian.

The k-SAT problem is important in the theory of computational complexity, since many other problems can be mapped into this one. There are, however, different tasks which can be solved using AQC, like implementing the Grover's algorithm for finding a marked item in an unstructured database of  $N$  items. Later in Sec. 4.2.4, we will present the adiabatic quantum evolution protocol proposed by Roland and Cerf (2002), which, in accordance with the performance of Grover's standard algorithm (Sec. 4.2), exhibits a quadratic speedup with respect to any classical algorithm that checks all possible solutions with a complexity of order  $N$ .

### 3.14 \* Maximum speed of quantum gates

The AQC issue of the speed at which certain control parameters need to be varied in time raises a fundamental important question related to the computational limits of the physical quantum dynamics. This question can be answered at various levels of generality. Here we present a simple argument based on the estimate of the speed by looking at the maximum number of distinct states that the system can pass through, per unit of time (Margolus and Levitin, 1997). In the quantum realm, the notion of distinct states applies to orthogonal states. This criterion would correspond, for a classical computer, to the maximum number of operations per second.

### 3.14.1 \* Speed limit of an autonomous time evolution

Let us consider a generic quantum system described by a time-independent Hamiltonian  $H$ , with energy levels  $E_n$  associated to the eigenstates  $|E_n\rangle$ . For the sake of simplicity in our presentation, we focus on a system with discrete spectrum. We choose the ground-state energy  $E_0 = 0$  and suppose that  $E_0 \leq E_1 \leq \dots \leq E_{N-1}$ . The problem of determining the maximum speed of the dynamical evolution can be recast into that of finding the minimum time needed for any state of the physical system to evolve into an orthogonal state.

We start from a generic superposition of energy eigenstates:

$$|\psi(t=0)\rangle = \sum_{n=0}^{N-1} c_n |E_n\rangle, \quad \text{with } \sum_n |c_n|^2 = 1, \quad (3.196)$$

such that at time  $t$  the state evolves into

$$|\psi(t)\rangle = \sum_{n=0}^{N-1} c_n e^{-\frac{i}{\hbar} E_n t} |E_n\rangle. \quad (3.197)$$

We can define its superposition with the initial state according to

$$S(t) = \langle \psi(0)|\psi(t)\rangle = \sum_n |c_n|^2 e^{-\frac{i}{\hbar} E_n t}. \quad (3.198)$$

We would like to find the smallest value of  $t$  such that  $S(t_{\min}) = 0$ . To do this, we just note that

$$\begin{aligned} \text{Re}(S) &= \sum_n |c_n|^2 \cos\left(\frac{E_n t}{\hbar}\right) \geq \sum_n |c_n|^2 \left[1 - \frac{2}{\pi} \left(\frac{E_n t}{\hbar} + \sin\left(\frac{E_n t}{\hbar}\right)\right)\right] \\ &= 1 - \frac{2Et}{\pi\hbar} + \frac{2}{\pi} \text{Im}(S), \end{aligned} \quad (3.199)$$

where we used the inequality  $\cos x \geq 1 - \frac{2}{\pi}(x + \sin x)$ , valid for  $x \geq 0$ . The average energy of the system is

$$E = \langle \psi(t)|H|\psi(t)\rangle = \sum_{n=0}^{N-1} |c_n|^2 E_n. \quad (3.200)$$

In order to have  $S(t) = 0$ , which implies  $\text{Re}(S) = \text{Im}(S) = 0$ , it is evident from the inequality (3.199) that we need the following condition:

$$t \geq \frac{\hbar}{4E}. \quad (3.201)$$

Therefore the minimum time at which  $S(t) = \langle \psi(0)|\psi(t)\rangle$  could reach the zero value is  $t_{\min} = \hbar/4E$ .

We note that the maximum speed is achievable if the spectrum of energies includes the energy  $2E$ , while it is approached if there is an eigenenergy that is close to this value. In that case, one can simply prepare the initial condition  $|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |2E\rangle)$ , which has average energy  $E$ , and is such that at time  $t_{\min}$  it evolves into the orthogonal state  $|\psi(t_{\min})\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |2E\rangle)$ .

### 3.14.2 \* Speed limit of single-qubit gates

To have an idea of the possible speed of a quantum gate, we follow Levitin *et al.* (2002) and provide a tangible example for a single-qubit operation, as it is not possible to give generic results, due to the large difference in the various implementations.

Let  $|\psi_1(0)\rangle = |0\rangle$  and  $|\psi_2(0)\rangle = |1\rangle$  be the two initial orthogonal stationary states of a qubit, where  $\{|0\rangle, |1\rangle\}$  is the computational basis. Consider a gate that complements the state of the qubit and then adds to it an arbitrary phase shift  $\theta$ . Namely, at the end of the operating time  $\tau$ , the two orthogonal states are swapped and a given phase shift  $\theta$  is added to the resulting state:

$$|\psi_1(\tau)\rangle = |\psi_2(0)\rangle e^{-i\theta}, \quad |\psi_2(\tau)\rangle = |\psi_1(0)\rangle e^{-i\theta}. \quad (3.202)$$

The evolution of the system is governed by the Hamiltonian  $H$ , such that  $|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle$ , with  $U(t) = e^{-\frac{i}{\hbar}Ht}$  being the unitary evolution operator. Note that, due to the linearity of quantum mechanics, Eq. (3.202) is a necessary and sufficient condition for an arbitrary state of the qubit  $a|\psi_1\rangle + b|\psi_2\rangle$  to be converted into the orthogonal state with a phase shift  $\theta$ , provided that  $\text{Re}(ab^*) = 0$  (the overall phase  $\theta$  is essential, since this qubit may be part of a many-qubit system).

The generic form of a two-qubit unitary time evolution in its diagonal basis is provided by

$$U_{\text{diag}}(t) = \begin{bmatrix} e^{-\frac{i}{\hbar}E_1 t} & 0 \\ 0 & e^{-\frac{i}{\hbar}E_2 t} \end{bmatrix}, \quad (3.203)$$

where  $E_1$  and  $E_2$  are the eigenvalues of  $H$ . This operator can be written in a generic basis by applying a change of basis:  $U(t) = WU_{\text{diag}}(t)W^\dagger$ , where  $W$  can be seen as an arbitrary rotation of the Bloch sphere. Parametrizing  $W$  as a rotation of a given angle  $-2\phi_1$  around the axis  $\mathbf{n} = (\cos\phi_2, -\sin\phi_2, 0)$ , we obtain from Eq. (3.60):

$$W(\phi_1, \phi_2) = \cos\phi_1 I + i\sin\phi_1(\mathbf{n} \cdot \boldsymbol{\sigma}) = \begin{bmatrix} \cos\phi_1 & i\sin\phi_1 e^{-i\phi_2} \\ i\sin\phi_2 e^{i\phi_2} & \cos\phi_1 \end{bmatrix} \quad (3.204)$$

from which it follows that

$$U(t) = e^{-\frac{i}{\hbar}Et} \begin{bmatrix} \cos^2\phi_1 e^{-\frac{i}{\hbar}\Delta_E t} + \sin^2\phi_1 e^{\frac{i}{\hbar}\Delta_E t} & -2ie^{i\phi_2} \sin\phi_1 \cos\phi_1 \sin\left(\frac{\Delta_E}{\hbar}t\right) \\ -2ie^{-i\phi_2} \sin\phi_1 \cos\phi_1 \sin\left(\frac{\Delta_E}{\hbar}t\right) & \cos^2\phi_1 e^{\frac{i}{\hbar}\Delta_E t} + \sin^2\phi_1 e^{-\frac{i}{\hbar}\Delta_E t} \end{bmatrix}, \quad (3.205)$$

where we defined

$$E = \frac{E_1 + E_2}{2} \quad \text{and} \quad \Delta_E = \frac{E_2 - E_1}{2}. \quad (3.206)$$

It can be readily shown that, to satisfy the conditions (3.202) and obtain the minimum time  $\tau$ , we need to choose  $\phi_1 = \pi/4$  and  $\phi_2 = 0$ . Therefore, in the original basis of the system with basis vectors  $|\psi_1(0)\rangle$  and  $|\psi_2(0)\rangle$ , the operator  $U(t)$  has the following form:

$$U(t) = e^{-\frac{i}{\hbar}Et} \begin{bmatrix} \cos\left(\frac{\Delta_E}{\hbar}t\right) & -i\sin\left(\frac{\Delta_E}{\hbar}t\right) \\ -i\sin\left(\frac{\Delta_E}{\hbar}t\right) & \cos\left(\frac{\Delta_E}{\hbar}t\right) \end{bmatrix}, \quad (3.207)$$

that is,  $U(t) = e^{-\frac{i}{\hbar}Et}e^{-\frac{i}{\hbar}\Delta_E t \sigma_x}$ , so that the corresponding Hamiltonian is

$$H = E + \Delta_E \sigma_x = \begin{bmatrix} E & \Delta_E \\ \Delta_E & E \end{bmatrix}. \quad (3.208)$$

Thus, in order to respect the conditions (3.202), we require

$$\frac{\Delta_E}{\hbar} \tau = \pm \frac{\pi}{2} \quad \text{and} \quad \frac{E}{\hbar} \tau = \theta \pm \frac{\pi}{2}. \quad (3.209)$$

From these we obtain that the minimum time of operation of the quantum gate is

$$\tau(\theta) = \frac{\hbar}{4E} \left( 1 + 2 \frac{\theta \bmod \pi}{\pi} \right). \quad (3.210)$$

At this point it is necessary to remark that the argument explained above relies on the hypothesis that the system must be autonomous. For driven systems, as those that are usually necessary to operate quantum gates, the results may be different. However, Svozil *et al.* (2005) have shown that the same expression (3.210) holds for the general case of time-dependent Hamiltonians.

### 3.15 \* Holonomic quantum computation

There is another quantum computation scheme that relies on the adiabatic theorem, and which is based on the possibility to perform quantum operations of a totally geometric origin. Namely, one can suitably modulate the system parameters in time so that, after an adiabatic cyclic evolution, the acquired geometric phase coincides with the required gate operation. This approach is called *holonomic quantum computation* (HQC) (Zanardi and Rasetti, 1999). It has been pointed out that a set of universal gates can be achieved by cleverly implementing geometric phases, and without any contribution from the dynamical phases. To reach this purpose, the computational space needs to be a degenerate eigenspace of the governing Hamiltonian, so that the acquired geometric phase will have a non-Abelian character (see Sec. 3.12). Despite the fact that the Hamiltonian in that subspace can be trivial and does not produce any dynamical evolution, the dependence on some adiabatically changing parameters may cause the computational space undergo a very complex and nontrivial evolution.

Here we would like to provide a hint on how this is possible, and focus on a specific physical implementation that has been first introduced by Duan *et al.* (2001). Namely, we consider a set of trapped ions, which can be manipulated by means of suitable laser configurations. Our aim is to realize the following set of universal quantum gates, which can be written in the computational basis as:

$$U_1^{(j)} = e^{i\varphi_1} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad U_2^{(j)} = e^{i\varphi_2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (\text{single-qubit gates}), \quad (3.211)$$

$$U_3^{(j,k)} = e^{i\varphi_3} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{two-qubit gate}). \quad (3.212)$$

Here  $\varphi_i$  ( $i = 1, 2, 3$ ) are arbitrary phases.

Let us start from single-qubit gates  $U_1^{(j)}$  and  $U_2^{(j)}$ , acting on the  $j$ -th qubit. We first identify a qubit system with two internal levels of a single ion in our computer. Here we are not going to present the details of the physical implementation; for our purposes, we just need to know that the relevant energy levels of the ion are three metastable states  $|0\rangle_j$ ,  $|1\rangle_j$ ,  $|a\rangle_j$  having roughly the same energy, and one excited state  $|e\rangle_j$ . Each of the three ground states is coupled to the excited state by a resonant laser, following the scheme depicted in Fig. 3.26. At the initial time, we suppose the logical states of the qubit to be encoded in  $|0\rangle_j$  and  $|1\rangle_j$ . It is possible to show that this system can be effectively described by the Hamiltonian

$$H_j = \begin{bmatrix} 0 & 0 & 0 & \Omega_0^*(t) \\ 0 & 0 & 0 & \Omega_1^*(t) \\ 0 & 0 & 0 & \Omega_a^*(t) \\ \Omega_0(t) & \Omega_1(t) & \Omega_a(t) & 0 \end{bmatrix}, \quad (3.213)$$

where  $\Omega_0(t)$ ,  $\Omega_1(t)$ ,  $\Omega_a(t)$  are the three Rabi frequencies associated with the respective transitions, and can be modulated in time.

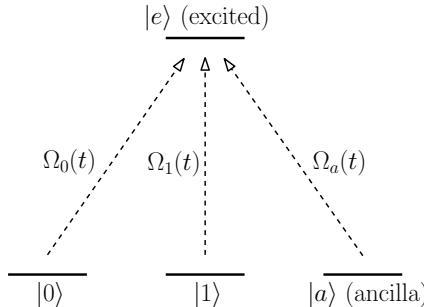


Fig. 3.26 Level structure and laser configuration for a so-called “tripod” system, used to implement single-qubit gates. The levels  $|0\rangle$  and  $|1\rangle$  denote the computational basis of the qubit, while  $|a\rangle$  is an ancillary state that is used for the intermediate steps of the transformation. The three degenerate levels are connected with an upper level  $|e\rangle$  by suitable time-dependent Rabi frequencies.

At time  $t = 0$  we suppose that  $\Omega_0 = \Omega_1 = 0$ , so that the computational space spanned by  $|0\rangle_j$  and  $|1\rangle_j$  constitutes an eigenspace of the Hamiltonian with zero eigenenergy. The three Rabi frequencies are then changed in time, such to make a cyclic evolution with a rate much smaller than the typical Rabi frequencies. The adiabatic theorem guarantees that the computational space remains the zero eigenspace of the Hamiltonian and no dynamical phase contribution emerges. We now explain the protocol for the realization of the universal single-qubit gates in (3.211).

For the gate  $U_1^{(j)}$ , we need to set:

$$\Omega_0(t) = 0, \quad \Omega_1(t) = -\Omega e^{i\phi(t)} \sin \frac{\theta(t)}{2}, \quad \Omega_a(t) = \Omega \cos \frac{\theta(t)}{2}, \quad (3.214)$$

where  $\Omega$ ,  $\theta(t)$ ,  $\phi(t)$  denote the control parameters. We consider a cyclic evolution of  $\theta(t)$  and  $\phi(t)$ , with the starting and the ending point being  $\theta = \phi = 0$ . The relevant states of the gate obviously are the zero-energy eigenstates (also called the “dark states”) of the Hamiltonian. From the eigenvalue equation for the indicated Hamiltonian, one finds the following two of such states:

$$|D_1\rangle_j = |0\rangle_j, \quad |D_2\rangle_j = \cos \frac{\theta}{2} |1\rangle_j + \sin \frac{\theta}{2} e^{i\phi} |a\rangle_j. \quad (3.215)$$

Note that, at the final point of the evolution,  $|D_2\rangle_j = |1\rangle_j$ . It is easy to calculate the connection form  $\mathbb{A} = (\mathbb{A}_\theta, \mathbb{A}_\phi)$  in this subspace, as defined in Eq. (3.171):

$$\begin{aligned} \mathbb{A}_\theta &= \begin{bmatrix} \langle D_1 | \partial_\theta | D_1 \rangle & \langle D_1 | \partial_\theta | D_2 \rangle \\ \langle D_2 | \partial_\theta | D_1 \rangle & \langle D_2 | \partial_\theta | D_2 \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathbb{A}_\phi &= \begin{bmatrix} \langle D_1 | \partial_\phi | D_1 \rangle & \langle D_1 | \partial_\phi | D_2 \rangle \\ \langle D_2 | \partial_\phi | D_1 \rangle & \langle D_2 | \partial_\phi | D_2 \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & i \sin^2 \frac{\theta}{2} \end{bmatrix}. \end{aligned} \quad (3.216)$$

This implies that, after an adiabatic loop in the parameter space, the gate  $U_1^{(j)} = e^{i\varphi_1 |1\rangle_j \langle 1|}$  is realized with the holonomic phase

$$\varphi_1 = \oint_C \sin^2 \frac{\theta}{2} d\phi = \frac{1}{2} \oint_C (1 - \cos \theta) d\phi. \quad (3.217)$$

This has a clear geometric interpretation, since the acquired geometric phase is exactly one half of the solid angle  $\Omega_S = \oint d\Omega \sin \theta d\theta d\phi$  spanned by the vector pointing to the  $(\theta, \phi)$  direction.

For the gate  $U_2^{(j)}$ , we need to take real-valued Rabi frequencies, according to:

$$\Omega_0(t) = \Omega \sin \theta(t) \cos \phi(t), \quad \Omega_1(t) = \Omega \sin \theta(t) \sin \phi(t), \quad \Omega_a(t) = \Omega \cos \theta(t).$$

As we did previously, we choose to fix  $\Omega$  and let the parameters  $\theta(t)$  and  $\phi(t)$  undergo an adiabatic cyclic evolution from the point  $\theta = \phi = 0$ . The two degenerate dark states of the Hamiltonian are

$$|D_1\rangle_j = \cos \theta (\cos \phi |0\rangle_j + \sin \phi |1\rangle_j) - \sin \theta |a\rangle_j, \quad |D_2\rangle_j = -\sin \phi |0\rangle_j + \cos \phi |1\rangle_j. \quad (3.218)$$

Note that, at the final point of the evolution,  $|D_1\rangle_j = |0\rangle_j$  and  $|D_2\rangle_j = |1\rangle_j$ . Using the same definition of Eq. (3.216), we find that the connection form is given by

$$\mathbb{A}_\theta = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbb{A}_\phi = \begin{bmatrix} 0 & \cos \theta \\ -\cos \theta & 0 \end{bmatrix}. \quad (3.219)$$

Therefore the cyclic evolution reproduces the gate  $U_2^{(j)} = e^{i\varphi_2 \sigma_j^y}$  with a phase

$$\gamma_2 = \oint_C \cos \theta d\phi. \quad (3.220)$$

Finally we need to construct the two-qubit gate  $U_3^{(j,k)}$ . In order to achieve this task, one has to exploit the fact that the different ions can interact through

a Coulomb force. Without going into the details, it is possible to show that the effective Hamiltonian describing the interaction can be well approximated by

$$H_{j,k} = \varepsilon \left[ -|\Omega_1|^2 \sigma_j^{(\mu, \varphi_1)} \sigma_k^{(1, \varphi_1)} + |\Omega_a|^2 \sigma_j^{(a, \varphi_a)} \sigma_k^{(a, \varphi_a)} \right], \quad (3.221)$$

where  $\sigma_j^{(\mu, \varphi_\mu)} = e^{i\varphi_\mu} |e\rangle_j \langle \mu| + \text{H.c.}$  ( $\mu = 1, a$ ). We now choose the parametrization

$$|\Omega_1|^2 / |\Omega_a|^2 = \tan \frac{\theta}{2}, \quad \varphi_1 - \varphi_a = \frac{\varphi}{2}, \quad (3.222)$$

and suppose that the control parameters  $(\theta, \varphi)$  follow an adiabatic cyclic evolution from  $\theta = 0$ . We can easily observe that during such evolution the computational basis elements for the two qubits  $|00\rangle_{jk}$ ,  $|01\rangle_{jk}$ ,  $|10\rangle_{jk}$  are decoupled, while the state  $|11\rangle_{jk}$  adiabatically evolves into  $|\psi\rangle_{jk} = \cos(\theta/2)|11\rangle_{jk} + \sin(\theta/2)e^{i\varphi}|aa\rangle_{jk}$ . Therefore, analogously to adiabatic evolution of the dark state  $|D_2\rangle_j$  that we used to achieve the single-qubit gate  $U_2^{(j)}$  (note that  $|\psi\rangle_{jk}$  has exactly the same structure), it is clear that the above transformation realizes the two-qubit gate  $U_3^{(j,k)}$  with the holonomic phase

$$\varphi_3 = \oint_C \sin^2 \frac{\theta}{2} d\varphi. \quad (3.223)$$

This coincides with one half of the solid angle spanned during the loop.

### 3.16 A guide to the bibliography

The quantum Turing machine is discussed, *e.g.*, in Galindo and Martin-Delgado (2002). A pioneering study of the circuit model of quantum computation was given by Deutsch (1989). A clear discussion of the role of entanglement in quantum computational speed-up is Jozsa and Linden (2003), see also Biham *et al.* (2004).

The universality of two-qubit quantum gates is discussed in Reck *et al.* (1994), DiVincenzo (1995), Barenco (1995), Deutsch *et al.* (1995) and Lloyd (1995). Many useful circuit constructions can be found in Barenco *et al.* (1995) and Song and Klappecker (2003). The decomposition of unitary matrices into matrices of smaller size is discussed by Tucci (1999).

A practical method of constructing quantum logic circuits is discussed in Lee *et al.* (1999). Quantum circuits implementing various arithmetic operations can be found in Vedral *et al.* (1996), Beckman *et al.* (1996), Miquel *et al.* (1996), Gossett (1998) and Draper (2000).

The importance of the Berry phase in the context of electric polarization is first discussed in Resta (1994), while a fairly exhaustive review on its effects in a variety of other solid-state applications is Xiao *et al.* (2010).

The adiabatic theorem is a fundamental concept of quantum mechanics which was put forward in Born and Fock (1928). However only several decades later it was realized in Berry (1984) that the presence of geometric phases may result in a highly non-trivial contribution to the dynamics. Now these concepts can be found in several textbooks as, for instance, Griffiths (2005).

The first fundamental papers dealing with adiabatic quantum computation are Fahri *et al.* (2000) and Fahri *et al.* (2001), while more advanced issues related to many-body systems are discussed in Santoro and Tosatti (2006). A more recent review with a fairly complete up-to-date survey has been written by Albash and Lidar (2018). The speed of quantum gates is discussed in Levitin *et al.* (2002) and Svozil *et al.* (2005). The idea of holonomic quantum computation was put forward in Zanardi and Rasetti (1999) and Pachos *et al.* (1999).

## Chapter 4

# Quantum algorithms

In this chapter, we focus on quantum algorithms and the techniques underlying their construction. These algorithms take advantage of the basic properties of quantum mechanics, from the superposition principle to entanglement and interference effects, to solve certain computational problems much more efficiently than a classical computer. This includes basic problems of computer science: from the search of a marked item in an unstructured database (Grover's algorithm) to integer factoring (Shor's algorithm). The latter case provides a striking exponential speedup over the best known classical algorithms. After that, we discuss a third relevant class of quantum algorithms, the simulation of physical systems.

### 4.1 Deutsch's algorithm

Deutsch's problem illustrates the computational power of quantum *interference*. We consider a *black box* (called the *oracle*) evaluating a one-bit Boolean function  $f : \{0, 1\} \rightarrow \{0, 1\}$ . There are 4 such functions, shown in Table 4.1. They differ in the following *global* property: two of them are *constant* ( $f_0$  and  $f_3$ ) and two *balanced* ( $f_1$  and  $f_2$ ). The problem is to decide whether a given function is constant or balanced. The solution to this problem necessarily requires two queries of the oracle in a classical computer. We shall show in this section that a quantum computer can solve the same problem with only one oracle query.

Table 4.1 One-bit logic functions.

$x$	$f_0$	$f_1$	$f_2$	$f_3$
0	0	0	1	1
1	0	1	0	1

The quantum circuit implementing Deutsch's algorithm is shown in Fig. 4.1. The function  $f(x)$  is evaluated in reversible computation using an ancillary qubit  $|y\rangle$ . The unitary transformation  $U_f$  transforms  $|x\rangle|y\rangle$  into  $|x\rangle|y \oplus f(x)\rangle$ ; that is, it flips the second qubit if and only if  $f(x) = 1$ . The initial state of the qubits is  $|x\rangle|y\rangle = |0\rangle|1\rangle$ . Then a Hadamard gate prepares the first qubit in the superposition

$(|0\rangle + |1\rangle)/\sqrt{2}$ . As will be shown below, this will allow the quantum computer to evaluate both  $f(0)$  and  $f(1)$  in a single run, a possibility that is beyond the reach of a classical computer. Another Hadamard gate prepares the ancillary qubit in  $(|0\rangle - |1\rangle)/\sqrt{2}$ . This is crucial since, for each  $x \in \{0, 1\}$ ,

$$U_f|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = (-1)^{f(x)}|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \quad (4.1)$$

namely, the phase factor  $(-1)^{f(x)}$  is propagated backwards (*kicked back*) in front of the first qubit. Thus, the state of the quantum computer after the function evaluation is

$$\frac{1}{\sqrt{2}} \left[ (-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle \right] \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (4.2)$$

The second qubit is no longer used and from now on we shall ignore it. The final Hadamard gate leaves the first qubit in the state

$$\frac{1}{2} \left\{ \left[ (-1)^{f(0)} + (-1)^{f(1)} \right] |0\rangle + \left[ (-1)^{f(0)} - (-1)^{f(1)} \right] |1\rangle \right\}. \quad (4.3)$$

If  $f(0) = f(1)$ , this state is equal to  $|0\rangle = |f(0) \oplus f(1)\rangle$ . If instead  $f(0) \neq f(1)$ , this state is  $|1\rangle = |f(0) \oplus f(1)\rangle$  (in both cases up to an overall phase factor of no physical significance). In any case, we can write the final state of the first qubit as

$$|f(0) \oplus f(1)\rangle. \quad (4.4)$$

Then a measurement of the first qubit gives with unit probability the outcome 0 if the function is constant and the outcome 1 if the function is balanced. Therefore, a global property of the function  $f(x)$  has been encoded in a single qubit after a single call of  $f$ . This is because a quantum computer can evaluate both  $f(0)$  and  $f(1)$  simultaneously. The main point is that these two alternatives “paths” are combined by the final Hadamard gate, giving the desired interference pattern. The interference is constructive for the outcome  $f(0) \oplus f(1)$  and destructive for the alternative outcome.

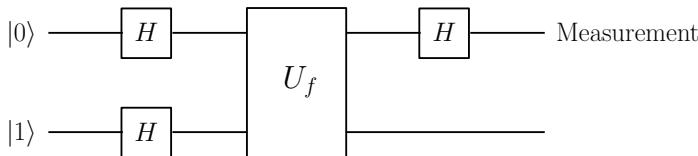


Fig. 4.1 A quantum circuit implementing Deutsch’s algorithm.

#### 4.1.1 The Deutsch–Jozsa problem

Now we shall consider some generalizations of Deutsch’s problem. The Deutsch–Jozsa algorithm solves the following problem in a single oracle query: we have an  $n$ -bit binary function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ , which is “promised” to be constant or balanced (*i.e.*, it has an equal number of output 0’s and 1’s), and we wish to

determine if it is constant or balanced. The quantum circuit that solves this problem is the same as for Deutsch's algorithm (Fig. 4.1), but with  $n$  qubits to store the input  $x = (x_{n-1}, x_{n-2}, \dots, x_0)$ . The Hadamard gates are now applied in parallel to all  $n$  qubits,

$$H^{\otimes n} = H \otimes H \otimes \cdots \otimes H. \quad (4.5)$$

It is easy to check that the action of  $H^{\otimes n}$  on a state  $|x\rangle$  of the computational basis gives

$$H^{\otimes n}|x\rangle = \prod_{i=0}^{n-1} \left( \frac{1}{\sqrt{2}} \sum_{y_i=0}^1 (-1)^{x_i y_i} |y_i\rangle \right) = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle, \quad (4.6)$$

where  $x \cdot y$  denotes the inner product of  $x$  and  $y$ , modulo 2:

$$x \cdot y = x_{n-1}y_{n-1} \oplus x_{n-2}y_{n-2} \oplus \cdots \oplus x_0y_0. \quad (4.7)$$

The circuit in Fig. 4.1, generalized to  $n$ -bit input, applies the transformation

$$(H^{\otimes n} \otimes I) U_f (H^{\otimes n} \otimes H) \quad (4.8)$$

to the input  $|00\dots0\rangle|1\rangle$  and generates the output

$$\left( \frac{1}{2^n} \sum_{x,y=0}^{2^n-1} (-1)^{f(x)+x \cdot y} |y\rangle \right) \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \quad (4.9)$$

This implies that, at the end of the circuit, a measurement of the  $n$  qubits in the computational basis gives the state  $|00\dots0\rangle$  with unit probability if  $f$  is constant and with zero probability if  $f$  is balanced. Therefore, a single run of the algorithm, with a single query of the function  $f(x)$ , determines with certainty if  $f$  is constant or balanced. This is an impressive result, since in classical computation one can be sure that  $f(x)$  is balanced only after  $2^n/2 + 1$  queries of the function. However, it would be rather unfair to claim an exponential gain of the quantum computer in this algorithm. Indeed, if  $f$  is balanced, the probability of obtaining the same response every time is  $1/2^{k-1}$ , with  $k$  number of function queries. Since this probability drops exponentially with  $k$ , in this case one can guess that  $f$  is constant with a probability of giving the wrong answer that drops exponentially with  $k$ . Therefore, as we have discussed in Sec. 1.3, a classical algorithm can reduce the probability of error below a level  $\epsilon$  after a number of queries  $k = O(\log(1/\epsilon))$ . We also note that on physical grounds it is not reasonable to ask the quantum computer to give an answer with absolute certainty, since some amount of error will be unavoidably present (see Sec. 3.8 and Chap. 7).

#### 4.1.2 \* An extension of Deutsch's algorithm

As an exercise, we now consider a further generalization of the algorithm, which shows that other global properties of a given function can be determined using a quantum computer. Let us consider, for the sake of simplicity, the case of  $n = 2$

Table 4.2 Two-qubit states at the end of the Deutsch–Jozsa circuit.

	$f_0$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$f_7$	$f_8$	$f_9$	$f_{10}$	$f_{11}$	$f_{12}$	$f_{13}$	$f_{14}$	$f_{15}$
$c_0$	4	2	2	0	2	0	0	-2	2	0	0	-2	0	-2	-2	-4
$c_1$	0	2	-2	0	2	4	0	2	-2	0	-4	-2	0	2	-2	0
$c_2$	0	2	2	4	-2	0	0	2	-2	0	0	2	-4	-2	-2	0
$c_3$	0	-2	2	0	2	0	4	2	-2	-4	0	-2	0	-2	2	0

qubits. The state of the 2 qubits at the end of the Deutsch–Jozsa circuit can be written in the computational basis as  $|\psi_f\rangle = c_0|00\rangle + c_1|01\rangle + c_2|10\rangle + c_3|11\rangle$ , with the amplitudes  $c_i$  shown in Table 4.2 for the 16 two-qubit logic functions of Table 3.1 (we omit the normalization factor to simplify writing).

The measurement of the two qubits gives the outcome 00 with unit probability for the constant functions  $f_0$  and  $f_{15}$ , and with zero probability for the balanced functions  $f_3$ ,  $f_5$ ,  $f_6$ ,  $f_9$ ,  $f_{10}$  and  $f_{12}$ . Furthermore, if the function is balanced, the measurement allows one to distinguish between three subclasses, that is,  $f_5$  and  $f_{10}$  (if the outcome is 01),  $f_3$  and  $f_{12}$  (outcome 10) and  $f_6$  and  $f_9$  (outcome 11). We note that it is not possible to distinguish between  $f_i$  and  $f_{15-i}$ , since the global sign of the wave function is not observable. With the Deutsch–Jozsa circuit, it is not possible to learn anything about the remaining 8 two-bit logic functions ( $f_1$ ,  $f_2$ ,  $f_4$ ,  $f_7$ ,  $f_8$ ,  $f_{11}$ ,  $f_{13}$  and  $f_{14}$ ) since the 4 possible outcomes of the two-qubit measurement have equal probability.

It is also possible to discriminate between the other global properties of the two-qubit binary functions by performing a permutation of the two-qubit binary functions  $f_i$ . These permutations can be implemented conveniently if the transformation (4.8) is slightly modified and becomes:

$$(H^{\otimes 2} A_j \otimes I) U_f (H^{\otimes 2} \otimes H), \quad (4.10)$$

where the  $A_j$  are unitary diagonal matrices, with diagonal elements equal to  $\pm 1$ . There are  $2^{2^n} = 16$  such matrices, which can be coded similarly to the Boolean function  $f_i$ :

$$A_0 \equiv \text{diag}\{1, 1, 1, 1\}, \quad A_1 \equiv \text{diag}\{1, 1, 1, -1\}, \quad \dots \quad (4.11)$$

It is easy to check that each unitary transformation  $A_j$  changes  $f(x)$  in the output wave function (4.9) as follows:  $f(x) \rightarrow f(x) \oplus (A_j)_{xx}$ , where  $(A_j)_{xx} = \langle x|A_j|x\rangle$  denotes a diagonal matrix element of  $A_j$ . Thus, each unitary transformation  $A_j$  induces a permutation of the functions  $f_i$ . Therefore, the standard Deutsch–Jozsa circuit (*i.e.*,  $A_j = I = A_0$ ) allows us to discriminate with certainty between

$$\{(f_0, f_{15}), (f_3, f_{12}), (f_5, f_{10}), (f_6, f_9)\}, \quad (4.12)$$

while by using  $A_1$  we can discriminate between

$$\{(f_1, f_{14}), (f_2, f_{13}), (f_4, f_{11}), (f_7, f_8)\}, \quad (4.13)$$

and so on.

## 4.2 Quantum search

In this section, we show that a quantum computer can usefully face the following problem: searching for one marked item inside an *unstructured* database of  $N = 2^n$  items. Let us give a simple example: we have a phone book and a given number and we wish to find the corresponding name. The best we can do with a classical computer is to go through the phone book, until we find the solution (the name). This means that it is easy to recognize the solution, but difficult to find it. This is a characteristic of the problems in the computational class **NP** and in many cases to solve these problems there is no better classical algorithm than exhaustive search through all possible solutions. In the following, we shall show that a quantum computer considerably speeds up this search.

The search problem can be rephrased as an *oracle problem*: we label the items of the database as  $\{0, 1, \dots, N-1\}$  and  $x_0$  is the unknown marked item. The oracle computes the  $n$ -bit binary function

$$f : \{0, 1\}^n \rightarrow \{0, 1\}, \quad (4.14)$$

defined as

$$f(x) = \begin{cases} 1 & \text{if } x = x_0, \\ 0 & \text{otherwise.} \end{cases} \quad (4.15)$$

The problem is to find  $x_0$  with the minimum number of queries of the oracle  $f$ . Elementary probability theory tells us that, if one enters  $k$  items into the black-box function  $f$ , the probability of finding  $x_0$  is  $k/N$ . Therefore, in order to find  $x_0$  with success probability  $p$ , each classical algorithm requires  $pN = O(N)$  oracle queries. Grover showed that the same problem can be solved by a quantum computer in  $O(\sqrt{N})$  queries. Thus, the quantum computer allows a quadratic speed up. The gain is not exponential and therefore does not change the computational class of the problem, but still there is a significant improvement. To give an idea of the importance of a quadratic speed up, it is sufficient to consider an example from classical computation, the fast Fourier transform algorithm. Indeed, the gain of the fast Fourier transform over the standard Fourier transform is quadratic and the fast Fourier transform has had an enormous impact on signal analysis and countless other applications.

### 4.2.1 Searching one item out of four

It is instructive to examine first the simple case of finding one item out of  $N = 4$  items (two qubits). The quantum circuit corresponding to this simple instance of Grover's algorithm is shown in Fig. 4.2. Initially, the two qubits are prepared in the state  $|00\rangle$ , while the required ancillary qubit is in the state  $|y\rangle = |1\rangle$ . The first Hadamard gates prepare the two qubits in the equal superposition state and the ancillary in the state  $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ . Thus, the quantum-computer wave function becomes

$$\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (4.16)$$

Then we query the oracle and evaluate the function  $f(x)$ . Indeed, after the oracle query we have  $|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle$ . This means that the value  $f(x)$  provided by the oracle is loaded into the ancillary qubit. Since the ancillary has been prepared in the state  $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ , it stays the same when  $f(x) = 0$  and changes its sign when  $f(x) = 1$ . In the following we consider the special case in which the oracle produces  $f(x) = 1$  only when  $x = x_0 = (1, 0)$ . There is no loss of generality in this since the other possible  $x_0$  values are treated equivalently: the circuit drawn in Fig. 4.2 solves the searching problem for the four possible values of  $x_0$ . The quantum-computer wave function after the oracle query is given by

$$\frac{1}{2}(|00\rangle + |01\rangle - |10\rangle + |11\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \quad (4.17)$$

which differs from the state (4.16) in the sign of the coefficient in front of the marked state. Note that, similarly to Deutsch's algorithm, the sign has been kicked back in front of the register  $|x\rangle$ . The second register is unchanged and we do not consider it any more.

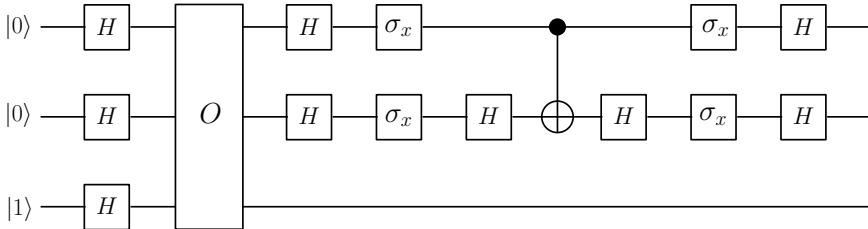


Fig. 4.2 A quantum circuit implementing Grover's algorithm for  $N = 4$  items. The Pauli matrix  $\sigma_x$  performs the NOT gate. The rectangle with a letter  $O$  inside denotes the oracle query.

After the oracle query, a measurement of the  $|x\rangle$  register would not distinguish between the different states of the computational basis, since the amplitudes of the coefficients in front of them are all the same. The key point of Grover's algorithm is to transform the phase difference, which appears in front of the component  $|10\rangle$  in (4.17), into an amplitude difference. This is achieved by means of the following unitary transformation:

$$D_{ij} = -\delta_{ij} + \frac{2}{2^n}. \quad (4.18)$$

In the present case with  $n = 2$ , the matrix  $D$  is given by

$$D = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}. \quad (4.19)$$

The transformation  $D$  can be implemented by means of that part of the circuit in Fig. 4.2 that follows the oracle query. Indeed, this transformation can be decomposed as follows:

$$D = H^{\otimes 2} D' H^{\otimes 2}, \quad (4.20)$$

where the diagonal matrix

$$D' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (4.21)$$

gives a controlled phase shift through an angle  $\pi$  (a minus sign) in the coefficient in front of the basis element  $|00\rangle$ . The matrix  $D'$  can be decomposed (up to an overall phase) in the following way:

$$D' = \sigma_x^{\otimes 2} (I \otimes H) \text{CNOT} (I \otimes H) \sigma_x^{\otimes 2}, \quad (4.22)$$

where  $\sigma_x^{\otimes 2} = \sigma_x \otimes \sigma_x$ . Indeed,  $(I \otimes H) \text{CNOT} (I \otimes H) = \text{CMINUS}$  (see exercise 3.13). In addition, the NOT gates  $\sigma_x^{\otimes 2}$  allow the phase factor to be placed in front of the state  $|00\rangle$  instead of the state  $|11\rangle$ , as it would in a standard controlled phase-shift gate.

The application of the transformation  $D$  to the state of Eq. (4.17) gives

$$\mathbf{D}^{\frac{1}{2}} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}. \quad (4.23)$$

Thus, a standard measurement of the two qubits now gives outcome 10 with certainty. Therefore, the problem has been solved with a single query of the function  $f$ , while a classical computer requires, on average,  $N_c = 2.25$  queries (the marked item can be found after the first, second or third query, with probability  $\frac{1}{4}$  each time; after the third function evaluation the fourth is in any case useless since we know that there is only one marked item and therefore  $N_c = \frac{1}{4} \times 1 + \frac{1}{4} \times 2 + \frac{1}{2} \times 3 = 2.25$ ).

#### 4.2.2 Searching one item out of $N$

Now we describe Grover's algorithm for a generic number of items  $N = 2^n$ . Again we need a single ancillary qubit  $|y\rangle$  and we prepare the quantum-computer wave function in the state  $|x\rangle|y\rangle = |00\dots 0\rangle|1\rangle$ . Then we apply  $n+1$  Hadamard gates, one for each qubit, in order to obtain the equal superposition of all basis states for the  $|x\rangle$  register and the state  $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  for the ancillary qubit. Then we evaluate the oracle function,  $|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle$ . As for the case of  $n=2$  qubits, this function evaluation kicks the sign back in front of the  $|x\rangle$  register:

$$\begin{aligned} \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) &\rightarrow \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \frac{1}{\sqrt{2}} (|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle) \\ &= \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \end{aligned} \quad (4.24)$$

However, we point out that, unlike the previous elementary example with  $n=2$  qubits, a single evaluation of the function  $f(x)$  is not sufficient to find the marked

item  $x_0$ . We must iterate the oracle query, which evaluates the function  $f(x)$ , several times. More precisely, we have to iterate several times the so-called Grover iteration  $G$ , defined as  $G = DO$ , where  $O$  denotes the oracle query and

$$D = H^{\otimes n} (-I + 2|0\rangle\langle 0|) H^{\otimes n}. \quad (4.25)$$

The transformation in between the two  $n$ -qubit Hadamard gates in (4.25) is a conditional phase shift, that puts a phase shift of  $-1$  in front of all the states of the computational basis, except for the state  $|00\dots 0\rangle$ . The quantum search algorithm is performed by repeatedly applying  $G$  until the state of the register  $|x\rangle$  is such that a standard measurement gives the outcome  $x_0$  with high probability.

#### 4.2.3 Geometric visualization

A simple geometric visualization helps in understanding the number of times that Grover's iteration  $G$  has to be applied. To this end we can ignore the ancillary qubit, whose state is always factorized and never changes.

First of all, it is clear from Eq. (4.24) that the action of the oracle  $O$  on an  $n$ -qubit state vector  $|x\rangle$  is given by

$$O : |x\rangle \rightarrow (-)^{f(x)} |x\rangle. \quad (4.26)$$

Therefore, we can write

$$O = I - 2|x_0\rangle\langle x_0| \equiv R_{|0\rangle}, \quad (4.27)$$

which corresponds to a reflection about the hyperplane perpendicular to  $|x_0\rangle$ . For example, if we consider a bidimensional space spanned by the vectors  $\{|x_0\rangle, |x_0^\perp\rangle\}$  and a generic vector  $|\psi\rangle = \alpha|x_0\rangle + \beta|x_0^\perp\rangle$ , we have  $O|\psi\rangle = -\alpha|x_0\rangle + \beta|x_0^\perp\rangle$ . Thus, as appears clearly from Fig. 4.3,  $O$  induces a reflection of the vector  $|\psi\rangle$  about the axis  $|x_0^\perp\rangle$ , that is, about the hyperplane perpendicular to  $|x_0\rangle$ .

Next we consider the transformation  $D$ . Note that

$$(H^{\otimes n})^\dagger (-I + 2|0\rangle\langle 0|) H^{\otimes n} = -I + 2|S\rangle\langle S|, \quad (4.28)$$

where we have exploited the fact that  $(H^{\otimes n})^\dagger = H^{\otimes n}$  and where  $|S\rangle$  is the equal superposition state:

$$|S\rangle \equiv H^{\otimes n} |0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle. \quad (4.29)$$

Thus, we can write Eq. (4.25) as minus the reflection about the hyperplane orthogonal to  $|S\rangle$ :

$$D = -(I - 2|S\rangle\langle S|) \equiv -R_{|S\rangle}. \quad (4.30)$$

Therefore,

$$G = DO = -R_{|S\rangle} R_{|0\rangle}. \quad (4.31)$$

We now consider the two-dimensional plane spanned by  $\{|x_0\rangle, |S\rangle\}$  and in this plane we draw the corresponding perpendicular unit vectors  $\{|x_0^\perp\rangle, |S^\perp\rangle\}$  (see Fig. 4.3). It is easy to prove<sup>1</sup> that

$$G = -R_{|S\rangle}R_{|0\rangle} = R_{|S^\perp\rangle}R_{|0\rangle}. \quad (4.32)$$

Thus, if  $\theta$  denotes the angle between the vectors  $|x_0^\perp\rangle$  and  $|S\rangle$ ,  $G$  rotates a generic vector  $|\psi\rangle$  in this plane by an angle  $2\theta$  (see Fig. 4.3). Since prior to the first oracle query the  $n$ -qubit state is the equal superposition state

$$|\psi_0\rangle \equiv |S\rangle = \sin \theta |x_0\rangle + \cos \theta |x_0^\perp\rangle, \quad (4.33)$$

and  $G$  rotates any vector in the plane spanned by  $\{|x_0\rangle, |S\rangle\}$  by an angle  $2\theta$ , then, after  $j$  steps of Grover's iteration, the  $n$ -qubit state is given by

$$|\psi_j\rangle \equiv G^j |\psi_0\rangle = \sin((2j+1)\theta) |x_0\rangle + \cos((2j+1)\theta) |x_0^\perp\rangle. \quad (4.34)$$

We note that this state always belongs to the plane of Fig. 4.3. The process must stop after  $k$  steps, where  $k$  is such that  $|\psi_k\rangle$  is very close to the marked state  $|x_0\rangle$ . This takes place when  $|\sin(2k+1)\theta| \approx 1$ . The smallest integer  $k$  that fulfils this condition is determined by the following relation:

$$(2k+1)\theta \approx \frac{\pi}{2}, \quad (4.35)$$

which implies

$$k = \text{round}\left(\frac{\pi}{4\theta} - \frac{1}{2}\right), \quad (4.36)$$

where round signifies the nearest integer. Since one starts from the equal superposition state,

$$\sin \theta = \langle x_0 | \psi_0 \rangle = \frac{1}{\sqrt{N}}, \quad (4.37)$$

and therefore, for large  $N$ ,  $\theta \approx \frac{1}{\sqrt{N}}$ . Thus, the number of required iterations in Grover's algorithm is

$$k = \text{round}\left(\frac{\pi}{4}\sqrt{N} - \frac{1}{2}\right) = O(\sqrt{N}). \quad (4.38)$$

The algorithm ends with a standard measurement in the computational basis, giving outcome  $x = \bar{x}$ . Then we proceed as in any probabilistic algorithm: we check by means of the oracle if the solution is correct. This is the case if  $f(\bar{x}) = 1$ , so that  $\bar{x} = x_0$ . If this is not the case, that is,  $f(\bar{x}) = 0$  implying  $\bar{x} \neq x_0$ , we repeat the quantum computation from the beginning. Therefore, unlike the special case with  $N = 4$ , the Grover algorithm succeeds with a probability that is not equal to one. However, the success probability is very close to one (see exercise 4.1).

We note that the geometric interpretation of Grover's algorithm is consistent with the particular case  $N = 4$ , in which  $\theta = \frac{\pi}{6}$ , and therefore condition (4.35) is fulfilled after  $k = 1$  step.

**Exercise 4.1** Show that the probability that Grover's algorithm fails (*i.e.*, the measurement does not give the marked item  $x_0$ ) drops like  $1/N$ .

<sup>1</sup>In order to prove that  $R_{|S\rangle} = -R_{|S^\perp\rangle}$ , we consider a generic vector  $|u\rangle = \mu|S\rangle + \nu|S^\perp\rangle$  residing in the plane spanned by  $\{|S\rangle, |S^\perp\rangle\}$ . We have  $R_{|S\rangle}|u\rangle = -\mu|S\rangle + \nu|S^\perp\rangle$  while  $R_{|S^\perp\rangle}|u\rangle = \mu|S\rangle - \nu|S^\perp\rangle = -R_{|S\rangle}|u\rangle$ .

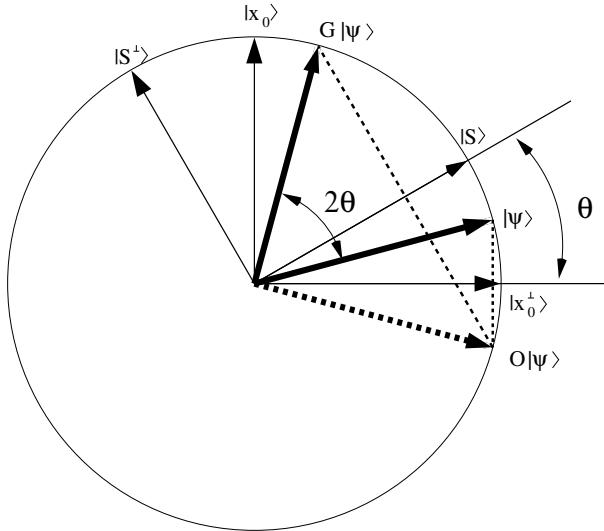


Fig. 4.3 Geometric visualization of Grover's iteration.

**Exercise 4.2** Estimate the number of elementary gates required to compute one step of Grover's iteration.

We point out that Grover's algorithm has been extended to the search for multiple items (also in the case in which the number of marked items is not known in advance) and many other situations. Unfortunately, it is also possible to prove that Grover's algorithm is optimal; that is, any other quantum algorithm for searching in an unstructured database would require at least  $O(\sqrt{N})$  oracle queries.

#### 4.2.4 Searching by adiabatic quantum evolution

We now describe how it is possible to implement Grover's algorithm with an AQC scheme, discussed in Sec. 3.13. We consider as an initial Hamiltonian

$$H_I = I - |S\rangle\langle S|, \quad (4.39)$$

whose ground state is the equal superposition state  $|S\rangle$  of Eq. (4.29). Let us suppose we are also able to apply to our system the Hamiltonian

$$H_F = I - |x_0\rangle\langle x_0|, \quad (4.40)$$

such that its ground state corresponds to the marked item  $|x_0\rangle$ . Note that, since  $H_F$  cannot be constructed without explicitly knowing  $x_0$ , this corresponds to applying the oracle (more precisely, evolving the system with  $H_F$  during a certain time interval is roughly equivalent to applying the quantum oracle). As originally proposed by Fahri *et al.* (2000), the time-dependent AQC Hamiltonian  $H(t)$  can be simply taken as a linear interpolation between these two Hamiltonians, as in Eq. (3.182).

Thus one would only need to prepare the system in the state  $|\psi(0)\rangle = |S\rangle$ , and then apply the Hamiltonian  $H(t)$  for a time  $T$ .

In order understand when the adiabatic condition in Eq. (3.183) holds, we first note that the matrix element contained in it is bounded from above by 1. The successive step is to diagonalize the Hamiltonian  $H(t)$  and find the instantaneous ground-state energy gap  $\Delta E_{10}(t)$ . It turns out that the two lowest eigenvalues  $E_0$  and  $E_1$  are non degenerate, and are separated by a gap

$$\Delta E_{10}(s) = \sqrt{1 - 4 \left(1 - \frac{1}{N}\right) s(1-s)}, \quad (4.41)$$

where  $s = t/T$ . Note that the minimum gap  $\Delta E_{\min} = \sqrt{1/N}$  is attained in the middle of the evolution, at  $s = 1/2$  (see the upper panel of Fig. 4.4). From this we conclude that the time required to successfully perform the algorithm has to be larger than  $N$ , meaning that, with this approach, there would be no advantage as compared to a classical search.

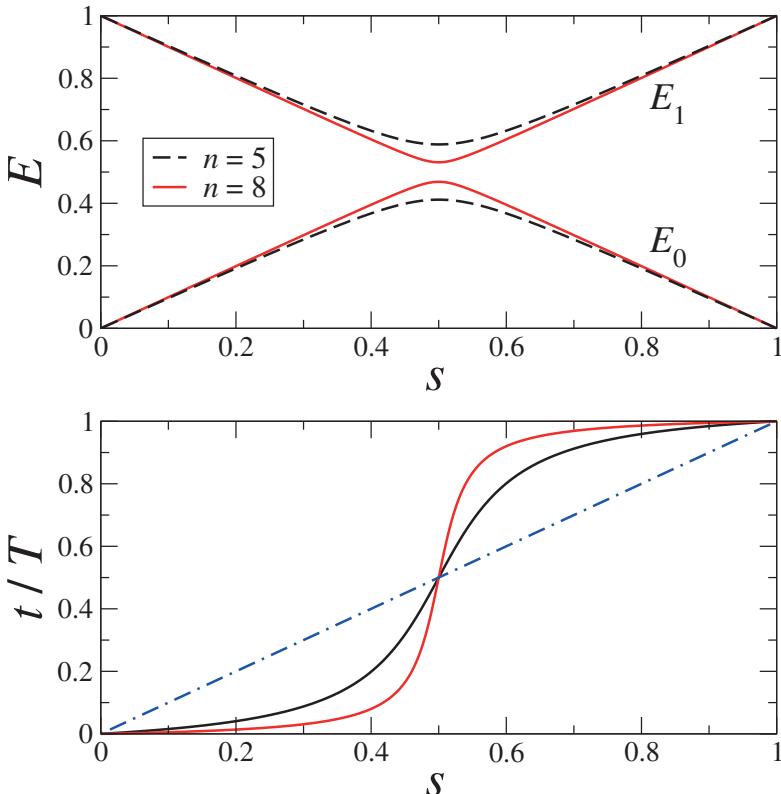


Fig. 4.4 Upper panel: the two lowest eigenvalues of the time-dependent Hamiltonian  $H(t)$ , which interpolated  $H_I$  and  $H_F$  of Eqs. (4.39) and (4.40) respectively, as a function of the reduced time  $s(t)$ , for  $n = 5$  (dashed black lines) and  $n = 8$  (continuous red lines) qubits. Lower panel: behaviour of the real time  $t$  as a function of the reduced time  $s$ , for the same values of  $n$ . Note that the system evolves slower in proximity of the point at  $s = 1/2$ , where the gap closes. The dotted-dashed blue line represents the plain case of a linear interpolation scheme, where  $s = t/T$ .

A few years later, Roland and Cerf (2002) pointed out that this adiabatic evolution method can be significantly improved by adapting the evolution rate  $ds/dt$  to the local adiabaticity condition, thus changing the linear evolution function  $s(t) = t/T$  into a nonlinear function such that  $s(0) = 0$  and  $s(T) = 1$ . The idea is to divide the total evolution time  $T$  into infinitesimal time intervals, and vary the evolution rate continuously in time, by applying the adiabaticity condition (3.183) locally to each of these infinitesimal time intervals  $dt$ :

$$\left| \frac{ds}{dt} \right| \ll \Delta E_{10}(s) \times \left| \langle \psi_1(s) | \frac{dH(s)}{ds} | \psi_0(s) \rangle \right|^{-1}. \quad (4.42)$$

It is thus sufficient to evolve the Hamiltonian  $H(t)$  at a rate which is a solution of

$$\frac{ds}{dt} = \epsilon \Delta E_{12}^2(s) = \epsilon \left[ 1 - 4 \left( 1 - \frac{1}{N} \right) s(1-s) \right], \quad (4.43)$$

where  $\epsilon \ll 1$ . Integrating it, we find

$$t = \frac{1}{2\epsilon} \frac{N}{\sqrt{N-1}} \left\{ \arctan [\sqrt{N-1}(2s-1)] + \arctan \sqrt{N-1} \right\}, \quad (4.44)$$

whose behaviour is plotted in the lower panel of Fig. 4.4: a gradual change in the switching between  $H_I$  and  $H_F$  is required, with  $H(t)$  changing faster when the gap is large (and slower when the gap is close to its minimum value for  $s \approx 1/2$ ). Finally, the computation time of this modified AQC algorithm can be evaluated by taking  $s = 1$  in Eq. (4.44):

$$T = \frac{1}{\epsilon} \frac{N}{\sqrt{N-1}} \arctan \sqrt{N-1} \xrightarrow{N \gg 1} \frac{\pi}{2\epsilon} \sqrt{N}. \quad (4.45)$$

The net result is a quadratic speed-up with respect to the classical search, in accordance with the performance of the standard Grover's algorithm, as discussed in the previous subsection.

### 4.3 The quantum Fourier transform

The *discrete* Fourier transform of a vector with complex components  $\{f(0), f(1), \dots, f(N-1)\}$  is a new complex vector  $\{\tilde{f}(0), \tilde{f}(1), \dots, \tilde{f}(N-1)\}$ , defined as

$$\tilde{f}(k) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{2\pi i \frac{j}{N} k} f(j). \quad (4.46)$$

The quantum Fourier transform does exactly the same. It is defined on a quantum register of  $n$  qubits ( $N = 2^n$ ) as the unitary operator  $F$  whose action on the states of the computational basis is given by:

$$F(|j\rangle) = \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{2\pi i \frac{j}{2^n} k} |k\rangle. \quad (4.47)$$

As a consequence, an arbitrary state  $|\psi\rangle = \sum_j f(j)|j\rangle$  is transformed into

$$|\tilde{\psi}\rangle = F|\psi\rangle = \sum_{k=0}^{2^n-1} \tilde{f}(k)|k\rangle, \quad (4.48)$$

where the coefficients  $\{\tilde{f}(k)\}$  are the discrete Fourier transform of the coefficients  $\{f(j)\}$ , according to the relation (4.46).

Now we construct the quantum circuit for computing the quantum Fourier transform. It is useful to introduce the following notations for the binary representation of  $j$ :

$$j = j_{n-1} j_{n-2} \dots j_0 = j_{n-1} 2^{n-1} + j_{n-2} 2^{n-2} + \dots + j_0 2^0, \quad (4.49)$$

and for a binary fraction:

$$0.j_l j_{l+1} \dots j_m = \frac{1}{2} j_l + \frac{1}{4} j_{l+1} + \dots + \frac{1}{2^{m-l+1}} j_m. \quad (4.50)$$

In a few simple steps we obtain the *product representation* of the Fourier transform:

$$\begin{aligned} F(|j\rangle) &= \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} \exp\left(\frac{2\pi ijk}{2^n}\right) |k\rangle \\ &= \frac{1}{\sqrt{2^n}} \sum_{k_{n-1}=0}^1 \dots \sum_{k_0=0}^1 \exp\left(2\pi ij \sum_{l=1}^n \frac{k_{n-l}}{2^l}\right) |k_{n-1} \dots k_0\rangle \\ &= \frac{1}{\sqrt{2^n}} \sum_{k_{n-1}=0}^1 \dots \sum_{k_0=0}^1 \otimes_{l=1}^n \exp\left(2\pi ij \frac{k_{n-l}}{2^l}\right) |k_{n-l}\rangle \\ &= \frac{1}{\sqrt{2^n}} \otimes_{l=1}^n \left[ \sum_{k_{n-l}=0}^1 \exp\left(2\pi ij \frac{k_{n-l}}{2^l}\right) |k_{n-l}\rangle \right] \\ &= \frac{1}{\sqrt{2^n}} \otimes_{l=1}^n \left[ |0\rangle + \exp\left(2\pi ij \frac{1}{2^l}\right) |1\rangle \right] \\ &= \frac{1}{\sqrt{2^n}} (|0\rangle + e^{2\pi i 0.j_0}|1\rangle) (|0\rangle + e^{2\pi i 0.j_1 j_0}|1\rangle) \dots (|0\rangle + e^{2\pi i 0.j_{n-1} j_{n-2} \dots j_0}|1\rangle). \end{aligned} \quad (4.51)$$

It is interesting to note that this product representation is factorized; this shows that the corresponding quantum state is *not entangled*.

The product representation (4.51) makes it easy to construct a quantum circuit that computes the quantum Fourier transform *efficiently*. We show such a circuit in Fig. 4.5, where  $R_k$  denotes the operator

$$R_k = \begin{bmatrix} 1 & 0 \\ 0 & \exp\left(\frac{2\pi i}{2^k}\right) \end{bmatrix}. \quad (4.52)$$

We consider the action of the circuit on a state  $|j\rangle = |j_{n-1} j_{n-2} \dots j_0\rangle$  of the computational basis (for a generic state  $|\psi\rangle = \sum_j c_j |j\rangle$  it is sufficient to remember that

the Fourier transform is a linear operator). The first Hadamard gate acts on the most significant qubit and generates the state

$$\frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i 0 \cdot j_{n-1}} |1\rangle) |j_{n-2} \dots j_0\rangle. \quad (4.53)$$

The subsequent controlled phase rotations, controlled- $R_2$  to controlled- $R_n$ , add phases from  $\pi/2$  to  $\pi/2^{n-1}$  if the corresponding control qubit is set to one. After these  $n - 1$  two-qubit gates, the quantum-computer wave function is left in the state

$$\frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i 0 \cdot j_{n-1} j_{n-2} \dots j_0} |1\rangle) |j_{n-2} \dots j_0\rangle. \quad (4.54)$$

A similar procedure is then repeated for the other qubits and therefore the quantum circuit in Fig. 4.5 generates the output

$$\frac{1}{\sqrt{2^n}} (|0\rangle + e^{2\pi i 0 \cdot j_{n-1} j_{n-2} \dots j_0} |1\rangle) (|0\rangle + e^{2\pi i 0 \cdot j_{n-2} \dots j_0} |1\rangle) \cdots (|0\rangle + e^{2\pi i 0 \cdot j_0} |1\rangle). \quad (4.55)$$

This state coincides with the product representation (4.51), except for the fact that the order of the qubits is reversed. The correct order can be obtained by means of  $O(n)$  SWAP gates (see Fig. 3.5); alternatively, one can simply relabel qubits. The circuit of Fig. 4.5 shows that the discrete Fourier transform of a complex vector of size  $N = 2^n$  can be implemented efficiently on a quantum register of  $n$  qubits using  $n$  Hadamard and  $n(n-1)/2$  controlled phase-shift gates. Therefore, the computation of the quantum Fourier transform requires  $O(n^2)$  elementary quantum gates, whereas the most efficient classical algorithm, the *fast Fourier transform*, computes the discrete Fourier transform in  $O(2^n n)$  elementary operations.

However, we emphasize that we cannot really talk about an exponential speed up in the computation of the Fourier transform, since a generic state  $|\psi\rangle = \sum_j f(j)|j\rangle$  cannot be prepared efficiently (see the discussion at the end of Sec. 3.7) and the final state  $|\tilde{\psi}\rangle = \sum_k \tilde{f}(k)|k\rangle$  is not readily accessible. Indeed, a standard measurement simply gives an outcome  $\bar{k}$  with probability  $|\tilde{f}(\bar{k})|^2$ . The problem is that the quantum Fourier transform is performed on the amplitudes of the wave function, which are *not directly accessible*. They can only be reconstructed with finite accuracy after many runs (each run computes the Fourier transform of the state  $|\psi\rangle$  and ends up with a standard projective measurement). If  $N$  denotes the number of runs, we can estimate  $\tilde{f}(k)$  as  $N_k/N$ , where  $N_k$  is the number of times that the measurement gives outcome  $k$ . We already encountered this problem in relation to function evaluation in Sec. 3.9 and it is actually a typical difficulty of quantum computation. Quantum algorithms find a way to *extract* efficiently useful information from the quantum-computer wave function. As we shall see in the following sections, the quantum Fourier transform is a key ingredient of exponentially efficient quantum algorithms.

**Exercise 4.3** Estimate the effect of unitary errors on the stability of the quantum Fourier transform.

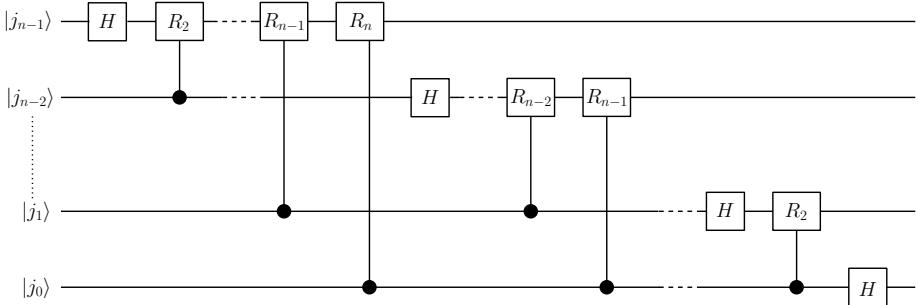


Fig. 4.5 A circuit implementing the quantum Fourier transform. The output is given by Eq. (4.55).

**Exercise 4.4** Construct a quantum circuit to compute the inverse Fourier transform

$$F^{-1}(|j\rangle) = \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{-2\pi i \frac{jk}{2^n}} |k\rangle. \quad (4.56)$$

#### 4.4 Quantum phase estimation

As a first application of the quantum Fourier transform we consider the following problem: a unitary operator  $U$  has an eigenvector  $|u\rangle$  with eigenvalue  $e^{i\phi}$  ( $0 \leq \phi < 2\pi$ ). Assume that we are able to prepare the state  $|u\rangle$  and there is a black box routine capable of performing the operations controlled- $U^{2^j}$ , where  $j$  is a non-negative integer. We wish to obtain the best  $n$ -bit estimate of  $\phi$ .

The quantum circuit solving this problem is shown in Fig. 4.6. The first register contains  $n$  qubits,  $n$  depending on the required accuracy for  $\phi$ . The second register contains the  $m$  qubits necessary to store  $|u\rangle$ . The action of the gate  $C - U^{2^j}$  on a state  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|u\rangle$  is given by

$$C - U^{2^j} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|u\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + |1\rangle U^{2^j}|u\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + e^{i2^j\phi}|1\rangle)|u\rangle. \quad (4.57)$$

Taking into account this result, it is easy to check that the output of the circuit (4.6) is given by

$$\frac{1}{\sqrt{2^n}}(|0\rangle + e^{i2^{n-1}\phi}|1\rangle) \cdots (|0\rangle + e^{i2\phi}|1\rangle)(|0\rangle + e^{i\phi}|1\rangle)|u\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{iy\phi}|y\rangle|u\rangle. \quad (4.58)$$

As in Deutsch's and Grover's algorithms, the key point is that the quantum register that stores  $|u\rangle$  is prepared in an eigenstate of the operators  $U, U^2, U^4, \dots$ . As a consequence, the state of this register never changes and the phase factors  $e^{i\phi}, e^{2i\phi}, e^{4i\phi}, \dots$  are propagated backwards in the control register.

From now on we consider only the control register and show that its quantum Fourier transform gives the best  $n$ -bit estimate of  $\phi$  with high probability. It is

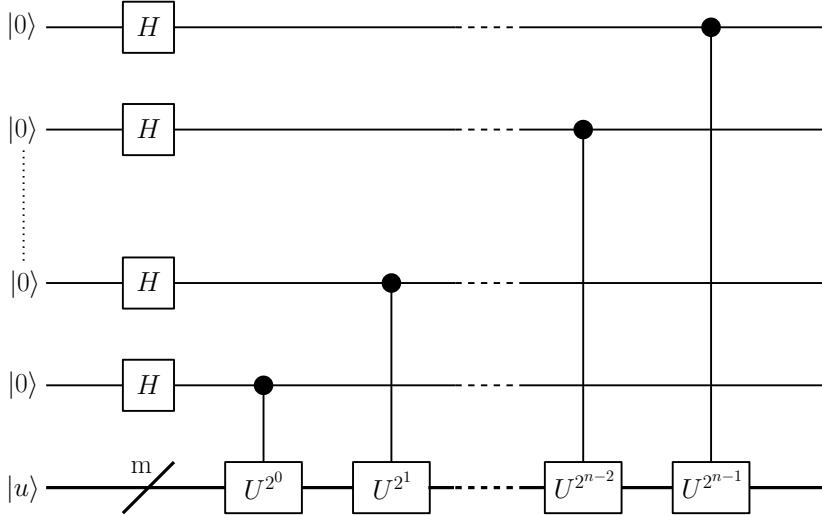


Fig. 4.6 A quantum circuit used to obtain the state (4.58) in the phase estimation problem. The wire with a dash represents a set of  $m$  qubits.

convenient to write

$$\phi = 2\pi \left( \frac{a}{2^n} + \delta \right), \quad (4.59)$$

where  $a = a_{n-1}a_{n-2}\cdots a_1a_0$  (in binary notation),  $2\pi a/2^n$  is the best  $n$ -bit approximation of  $\phi$  and therefore  $0 \leq |\delta| \leq 1/2^{n+1}$ . One can check that the inverse quantum Fourier transform (defined by Eq. (4.56)) of the first register, applied to the state (4.58), gives

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} e^{-2\pi i \frac{xy}{2^n}} e^{2\pi i(\frac{a}{2^n} + \delta)y} |x\rangle = \frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} e^{2\pi i(a-x)y/2^n} e^{2\pi i\delta y} |x\rangle. \quad (4.60)$$

We now perform a standard measurement of the first ( $|x\rangle$ ) register. If  $\delta = 0$ , the wave vector (4.60) reduces to  $|a\rangle$ . Therefore, in this case a standard measurement of the first register gives outcome  $x = a$  with certainty and the phase  $\phi$  is determined exactly. In the general case  $\delta \neq 0$ , the best  $n$ -bit estimate of  $\phi$  is given by  $a$  and is obtained from a standard measurement of the first register with probability  $p_a = |c_a|^2$ . Here  $c_a$  denotes the projection of the wave function (4.60) over the state  $|a\rangle$  and is given by

$$c_a = \frac{1}{2^n} \sum_{y=0}^{2^n-1} (e^{2\pi i\delta})^y = \frac{1}{2^n} \sum_{y=0}^{2^n-1} \alpha^y \quad (\alpha \equiv e^{2\pi i\delta}); \quad (4.61)$$

this finite geometric series can be added, giving

$$c_a = \frac{1}{2^n} \left[ \frac{1 - \alpha^{2^n}}{1 - \alpha} \right]. \quad (4.62)$$

Since for any  $z \in [0, \frac{1}{2}]$  we have  $2z \leq \sin(\pi z) \leq \pi z$ , we obtain

$$|1 - e^{2\pi i \delta 2^n}| = 2 |\sin(\pi \delta 2^n)| \geq 4 |\delta| 2^n, \quad |1 - e^{2\pi i \delta}| = 2 |\sin(\pi \delta)| \leq 2\pi |\delta|. \quad (4.63)$$

We can insert these two inequalities into (4.62) and obtain

$$|c_a|^2 \geq \frac{4}{\pi^2} \approx 0.405. \quad (4.64)$$

Therefore, the best  $n$ -bit estimate of  $\phi$  is obtained with high probability  $|c_a|^2$ .

We point out that it is possible to obtain the best  $l$ -bit approximation of the phase  $\phi$  with probability arbitrarily close to 1, provided that the number  $n$  of qubits is large enough. More precisely, the best  $l$ -bit approximation of  $\phi$  is obtained with probability  $> 1 - \epsilon$  if the circuit of Fig. 4.6 contains  $n = l + O(\log(1/\epsilon))$  qubits in the first register and the obtained result is rounded off to its most significant  $l$  bits (see Cleve *et al.*, 1998). Thus, by increasing the number  $n$  we raise not only the accuracy of our phase estimation but also the probability that our algorithm succeeds.

In summary, the quantum-phase estimation algorithm exploits the ability of the inverse quantum Fourier transform to perform the transformation from the state (4.58) to the state (4.60). As we have seen above, this latter state, when measured, gives with high probability a good estimate of the phase  $\phi$ . We stress that this algorithm is exponentially efficient with respect to any known classical algorithm solving the phase estimation problem, provided that the unitary operator  $U$  can be decomposed efficiently into elementary gates on a quantum computer (that is,  $U|u\rangle$  can be computed in a number of elementary quantum gates polynomial in the number  $m$  of qubits necessary to store  $|u\rangle$ ). Moreover, we assume that we are capable of preparing the eigenvector  $|u\rangle$ . We shall discuss an application of the quantum phase algorithm in Sec. 4.5.

## 4.5 \* Finding eigenvalues and eigenvectors

In this section, we describe a quantum algorithm that computes eigenvalues and eigenvectors of a given unitary operator  $U$ . To be concrete, we consider the case in which

$$U(\bar{t}) = \exp(-iH\bar{t}/\hbar) \quad (4.65)$$

is the evolution operator up to time  $\bar{t}$ , associated with a time-independent Hamiltonian  $H$ . The corresponding Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = H(x) \psi(x, t), \quad (4.66)$$

where, for the sake of simplicity, we have considered the one-dimensional case (the extension of what follows to higher dimensions is trivial). We note that an eigenvector  $\phi_\alpha(x)$  of  $H$  with eigenvalue  $E_\alpha$  (*i.e.*,  $H\phi_\alpha(x) = E_\alpha\phi_\alpha(x)$ ) is also an eigenvector of  $U(\bar{t})$  with eigenvalue  $e^{-iE_\alpha\bar{t}/\hbar}$ .

The eigenvalues and eigenvectors of the Hamiltonian operator  $H$  can be computed using a classical computer as follows. We evolve some initial state

$$\psi_0(x) \equiv \psi(x, t=0), \quad (4.67)$$

obtaining  $\psi(x, \bar{t}) = U(x, \bar{t})\psi_0(x)$ . If we expand  $\psi_0(x)$  over the basis of eigenfunctions of  $H$ ,

$$\psi_0(x) = \sum_{\alpha} a_{\alpha} \phi_{\alpha}(x), \quad (4.68)$$

we can write

$$\psi(x, t) = \sum_{\alpha} a_{\alpha} e^{-i\omega_{\alpha} t} \phi_{\alpha}(x) \quad (\omega_{\alpha} \equiv E_{\alpha}/\hbar). \quad (4.69)$$

Then we compute the Fourier transform

$$\tilde{\psi}(x_0, \omega) = F[\psi(x_0, t)], \quad (4.70)$$

for a given  $x = x_0$ . It is evident from (4.69) that the Fourier transform  $\tilde{\psi}$  exhibits peaks corresponding to the frequencies of motion  $\omega_{\alpha}$ . A given peak is resolved if the time evolution is computed up to time  $\bar{t}$  much longer than the inverse of the frequency of interest. If the Fourier transform  $\tilde{\psi}$  is repeated for different  $x_0$  values, one obtains the eigenfunctions  $\phi_{\alpha}(x)$  from the relative amplitudes of the peaks corresponding to the frequencies  $\omega_{\alpha}$ . This latter point is evident from the expansion (4.69), which implies

$$\frac{\tilde{\psi}(x_1, \omega_{\alpha})}{\tilde{\psi}(x_2, \omega_{\alpha})} = \frac{\phi_{\alpha}(x_1)}{\phi_{\alpha}(x_2)}. \quad (4.71)$$

Let us now repeat the same procedure on a quantum computer. The wave function is coded inside some interval  $x \in [-L, L]$  of interest for the physical motion, using a grid of  $2^n$  points separated by an interval  $\Delta x = 2L/(2^n - 1)$ :

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \psi(i) |i\rangle, \quad (4.72)$$

where  $\psi(i) = \psi(-L + i\Delta x)$ . We note that the quantum computer has an exponential advantage in memory requirements, since one needs  $n$  qubits to store the  $2^n$  complex numbers of the wave function (4.72). The evolution in a time step  $\Delta t$  is given by the unitary operator

$$U = \exp(-i H \Delta t/\hbar), \quad (4.73)$$

which, for a large class of physically significant Hamiltonians, can be simulated efficiently on a quantum computer (see Lloyd, 1996). The appropriate circuit for computing the eigenvalues and eigenvectors of the operator  $U$  is that introduced in Fig. 4.6 for the phase estimation problem. The second (target) register ( $n$  qubits) is used to store the wave function, initially prepared in the state  $|\psi_0\rangle = \sum_i \psi_0(i) |i\rangle$ . Such a preparation in special cases can be performed efficiently, for example when

the wave function is localized at a point,  $|\psi_0\rangle = |\bar{i}\rangle$ . We note that this choice is good enough to search for all eigenstates  $|\phi_\alpha\rangle = \sum_i \phi_\alpha(i)|i\rangle$  (and corresponding eigenvalues) having non-zero projections over  $|\psi_0\rangle$ , namely, a non-zero component  $\phi_\alpha(\bar{i})$ . The control register ( $l$  qubits) is prepared in the uniform superposition state  $\sum_j |j\rangle/\sqrt{2^l}$  and is necessary to simultaneously store the wave function at different times  $0, \Delta t, 2\Delta t, 3\Delta t, \dots, 2^{l-1}\Delta t$ . After execution of the gates controlled- $U^{2^0}$ , controlled- $U^{2^1}$ ,  $\dots$ , controlled- $U^{2^{l-1}}$  (see Fig. 4.6) the state of the quantum computer is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2^l}} \sum_{j=0}^{2^l-1} |j\rangle U^j |\psi_0\rangle, \quad (4.74)$$

where we have

$$|\psi(j\Delta t)\rangle = U^j |\psi_0\rangle. \quad (4.75)$$

Using the expansion (4.68), we can write

$$|\Psi\rangle = \frac{1}{\sqrt{2^l}} \sum_j |j\rangle \sum_{\alpha=0}^{2^n-1} a_\alpha e^{-i\omega_\alpha j\Delta t} |\phi_\alpha\rangle. \quad (4.76)$$

It can be shown that the Fourier transform of the state  $|\Psi\rangle$  with respect to the  $|j\rangle$  register has peaks corresponding to the frequencies  $\omega_\alpha$ . A sufficiently large number of computer runs followed by a standard measurement of the  $|j\rangle$  register would allow one to obtain the frequency spectrum of the Hamiltonian operator. It can be checked that, after any such measurement, the other quantum register collapses onto an eigenstate  $|\phi_\alpha\rangle$ . The eigenstates can in principle be reconstructed by means of standard measurements of this register.

The limitation of the method is that each computer run singles out different eigenvalues  $\omega_\alpha$ , with probability  $|a_\alpha|^2$ . With respect to the classical algorithm described above, there is an exponential advantage if the desired eigenvalues and eigenvectors are obtained after a polynomial number of trials. This is, for instance, the case if the initial state  $|\psi_0\rangle$  has a not exponentially small projection over the desired eigenstate, typically the ground state of a complex system (see Abrams and Lloyd, 1999).

## 4.6 Period finding and Shor's algorithm

One of the most important quantum algorithms is that for finding the period of a function. Suppose a function  $f(x)$  is periodic, namely  $f(x+r) = f(x)$ , with the period  $r < N$ . It is in general hard to find the period  $r$  on a classical computer. We could solve this problem by computing  $f(x)$  for subsequent inputs until the function repeats. This would take order of  $N/2$  values of  $x$ . For “well-behaved” functions better strategies are possible, but always inefficient, that is, exponential in the input size  $\log N$  required to specify  $N$ . On the other hand a quantum computer can find

the period efficiently, provided that the function  $f(x)$  can be computed efficiently from  $x$ .

To simplify the discussion, we shall consider the particular case in which  $r$  exactly divides the number of points  $N = 2^n$  over which the function  $f(x)$  is evaluated, that is  $N/r = m$ , with  $m$  integer. Moreover, we assume  $f(x) = f(y)$  if and only if  $x \equiv y \pmod{r}$ . The general case adds some complications<sup>2</sup> but does not change the general ideas discussed below. We need two registers, the first is prepared in the equal superposition of all states  $|x\rangle$  of the computational basis and the second stores the function  $f(x)$ , building the total state

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle |f(x)\rangle. \quad (4.77)$$

We stress that the *quantum parallelism* allows the function  $f(x)$  to be computed for all  $x$  in a single run. We also note that after function evaluation the two registers are *entangled*. Unfortunately, we cannot gain direct access to all values  $f(x)$ . On the contrary, after a measurement of the second register, it collapses onto a particular state  $|f(x_0)\rangle$ . Thus, the quantum-computer wave function becomes

$$\frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle |f(x_0)\rangle \quad (0 \leq x_0 < r-1), \quad (4.78)$$

where  $m = N/r$  is the number of  $x$  values such that  $f(x) = f(x_0)$ . Since  $r$  is the period of  $f(x)$ , we have  $f(x_0) = f(x_0 + r) = f(x_0 + 2r) = \dots = f(x_0 + (m-1)r)$ . Note that the offset  $x_0$  depends on the value obtained in the measurement of the second register. This means that we cannot extract the period of the function by simply measuring the first register, since each time we run the algorithm we would measure a different value of  $f(x)$  and therefore obtain a different offset. On the other hand, due to the shift property of the Fourier transform  $\tilde{f}$ , the offset  $x_0$  affects only the phases of  $\tilde{f}$  in a way, as we shall see below, unimportant for our purpose. Moreover, if  $f$  has period  $r$ ,  $\tilde{f}$  has period  $N/r$ .

We can now neglect the second register, which is factorized and does not concern us any more. A quantum Fourier transform of the first register gives the state

$$\frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{2\pi i(x_0 + jr)k/2^n} |k\rangle = \frac{1}{\sqrt{m}2^n} \sum_{k=0}^{2^n-1} e^{2\pi i x_0 k/2^n} \sum_{j=0}^{m-1} e^{2\pi i jrk/2^n} |k\rangle. \quad (4.79)$$

We must distinguish two instances. If  $k$  is a multiple of  $N/r$ , then  $e^{2\pi i jrk/2^n} = 1$  and therefore the coefficient in front of ket  $|k\rangle$  is equal to

$$\frac{1}{\sqrt{m}2^n} m e^{2\pi i x_0 k/2^n} = \frac{1}{\sqrt{r}} e^{2\pi i x_0 k/2^n}. \quad (4.80)$$

There is *constructive interference* in the summation of  $m$  terms in (4.79). On the other hand, if  $k$  is not a multiple of  $N/r$ , then we have *destructive interference*:

$$\sum_{j=0}^{m-1} w^j = \frac{1 - w^m}{1 - w} = 0, \quad (4.81)$$

---

<sup>2</sup>Elements of number theory are needed in this case, see for instance Ekert and Jozsa (1996).

where to simplify writing we have defined  $w = e^{2\pi ijr/2^n}$ . Therefore, quantum interference has selected a few specific frequencies. Actually, a quantum measurement of the wave function (4.79) will give one of the  $r$  outcomes  $lN/r$  ( $l = 0, 1, \dots, r-1$ ) with equal probability (the phase factor  $e^{2\pi ix_0k/2^n}$  being irrelevant). Thus, if we denote by  $c$  the measured value of  $k$ , we have  $c/N = \lambda/r$ , with  $\lambda$  an unknown integer. If  $\lambda$  and  $r$  have no common factors, then we reduce  $c/N$  to an irreducible fraction and thus obtain both  $\lambda$  and  $r$ . Number theory tells us that this takes place with probability at least  $1/\log \log r$ . Otherwise, the algorithm fails and must be repeated. It can be shown that the algorithm succeeds with probability arbitrarily close to one after a number of runs  $O(\log \log r)$ , see Shor (1997).

As a simple example, let us attempt to find the period of the function  $f(x) = \frac{1}{2}(\cos(\pi x) + 1)$ , assuming that  $x$  is loaded in a 3-qubit register, that is,  $N = 2^3 = 8$ . The function  $f(x)$  may be equal to 0 or 1 and therefore can be stored in a single-qubit register. Unknown to us, the function has a period  $r = 2$ , since  $f(x) = 1$  for even  $x$  and  $f(x) = 0$  for odd  $x$ . We wish to determine this period using the period-finding algorithm. For this purpose, we start from the fiducial state  $|000\rangle|0\rangle$  and build the state (4.77), which reads

$$\begin{aligned} & \frac{1}{\sqrt{8}} (|0\rangle|f(0)\rangle + |1\rangle|f(1)\rangle + |2\rangle|f(2)\rangle + |3\rangle|f(3)\rangle \\ & + |4\rangle|f(4)\rangle + |5\rangle|f(5)\rangle + |6\rangle|f(6)\rangle + |7\rangle|f(7)\rangle). \end{aligned} \quad (4.82)$$

Then we measure the second register obtaining, for example, the outcome 0. Thus, the total wave function collapses onto the state

$$\frac{1}{2} (|1\rangle + |3\rangle + |5\rangle + |7\rangle)|0\rangle, \quad (4.83)$$

where  $x = 1, 3, 5, 7$  are the values such that  $f(x) = 0$ . From now on, we neglect the second register, which is no longer used. Next, we apply the quantum Fourier transform to the first register, leading to the following transformation:

$$|x\rangle \rightarrow \frac{1}{\sqrt{8}} \sum_{k=0}^7 e^{2\pi i k x / 8} |k\rangle. \quad (4.84)$$

As a result, we obtain the wave function

$$\begin{aligned} |\psi\rangle = & \frac{1}{2\sqrt{8}} (|0\rangle + e^{i\pi/4}|1\rangle + e^{i2\pi/4}|2\rangle + \dots + e^{i7\pi/4}|7\rangle) \\ & + \frac{1}{2\sqrt{8}} (|0\rangle + e^{i3\pi/4}|1\rangle + e^{i6\pi/4}|2\rangle + \dots + e^{i21\pi/4}|7\rangle) \\ & + \frac{1}{2\sqrt{8}} (|0\rangle + e^{i5\pi/4}|1\rangle + e^{i10\pi/4}|2\rangle + \dots + e^{i35\pi/4}|7\rangle) \\ & + \frac{1}{2\sqrt{8}} (|0\rangle + e^{i7\pi/4}|1\rangle + e^{i14\pi/4}|2\rangle + \dots + e^{i49\pi/4}|7\rangle). \end{aligned} \quad (4.85)$$

It is easy to check that the complex amplitudes in front of the states  $|1\rangle, |2\rangle, |3\rangle, |5\rangle, |6\rangle$  and  $|7\rangle$  cancel each other. The interference is constructive only for the terms in front of the states  $|0\rangle$  and  $|4\rangle$ . Thus, we can write the wave function  $|\psi\rangle$  as follows:

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |4\rangle). \quad (4.86)$$

This expression is in agreement with the general formula (4.79). Finally, from the measurement of the first register we obtain outcomes 0 or 4 with equal probability. In the first case, we cannot find the period  $r$  of the function  $f(x)$  and the algorithm must be repeated. In the latter case, we write  $c/N = \lambda/r$ , with  $c = 4$  the measured value. Eventually, we reduce  $c/N = 4/8$  to an irreducible fraction  $c/N = \lambda/r = 1/2$ , thus obtaining the period  $r = 2$ .

The period finding algorithm is the essential ingredient in the most spectacular discovery of quantum computation, Shor's algorithm. Such quantum algorithm efficiently solves the *prime factorization* problem: given a composite odd positive integer  $N$ , find its prime factors. This is a central problem in computer science and it is conjectured, though not proven, that in classical computation this problem is in the class **NP** but not **P**: given the factors, it is easy to check if they solve the problem, but it is hard to find them. There are cryptographic systems (such as RSA, see Chap. 5) that are extensively used today and that are based on this conjecture. Indeed, in spite of the efforts over many centuries to find a polynomial time factoring algorithm, at present the best classical algorithm, the number field sieve, requires  $\exp(O(n^{1/3}(\log n)^{2/3}))$  operations; that is, it is super-polynomial in the input size  $n = \log N$ . Shor discovered a quantum algorithm that, using the above described period finding algorithm, accomplishes the same task in  $O(n^2 \log n \log \log n)$  elementary gates. Therefore, this algorithm is polynomial in the input size and there is an exponential speed up with respect to any known classical algorithm.

The factorization problem can be reduced to the problem of finding the period of the function  $f(x) = a^x \bmod N$ , where  $N$  is the number to be factorized and  $a < N$  is chosen randomly. Note that the modular exponentiation function can be computed efficiently on a quantum as well as on a classical computer (for the quantum case, see Barenco *et al.*, 1995; Miquel *et al.*, 1996). The advantage in the quantum case resides in the quantum parallelism which allows the function  $a^x$  to be computed for all  $x$  in a single run. The steps to operate the reduction from factorization to period finding can be performed efficiently on a classical computer. The period  $r$  of  $a^x \bmod N$  is the *order* of  $a$ , that is, the least integer such that  $a^r \equiv 1 \pmod{N}$ . We then compute  $\gcd(a^{r/2} \pm 1, N)$ , where  $\gcd(p, q)$  is the greatest common divisor of  $p$  and  $q$ , namely the largest integer that divides both  $p$  and  $q$ .<sup>3</sup> Since  $(a^{r/2} - 1)(a^{r/2} + 1) = a^r - 1 = 0 \bmod N$ , the numbers  $\gcd(a^{r/2} - 1, N)$  and  $\gcd(a^{r/2} + 1, N)$  are two non-trivial factors of  $N$ . Indeed in this case  $(a^{r/2} - 1)(a^{r/2} + 1) = a^r - 1 = lN$ , with  $l$  integer. Thus  $(a^{r/2} - 1)$  must contain one factor of  $N$ , and  $(a^{r/2} + 1)$  another. The procedure fails if  $r$  is odd or  $a^{r/2} \pm 1 = 0 \bmod N$ : in the first case  $r/2$  is not an integer, in the latter we obtain the trivial factors 1 and  $N$ . When the procedure fails,<sup>4</sup> it must be repeated with a different random number  $a < N$ . Therefore Shor's algorithm belongs to the class of probabilistic algorithms. More precisely,

<sup>3</sup>The Euclidean algorithm (see Knuth, 1997–1998) can be used to compute efficiently  $\gcd(p, q)$ .

<sup>4</sup>It can be shown (see Shor, 1997) that, by choosing a random integer  $a < N$ , the probability that  $r$  is even and  $a^{r/2} \neq \pm 1$  is at least  $1 - 1/2^{k-1}$ , where  $k$  is the number of distinct prime factors of  $N$ .

since the period of the modular exponentiation function can be computed efficiently on a quantum computer, we have a **BQP** (bounded-error quantum probabilistic polynomial) algorithm, see Sec. 1.3.2.

**Exercise 4.5** Find the order on  $a = 3 \bmod N = 91$  and use it to factor  $N$ .

Turning to the power of quantum computation, the following question naturally arises: Why is the speedup of quantum computation (with respect to any known classical algorithm) exponential for the factoring problem and only quadratic for the searching problem? Let us give an intuitive argument to help understanding this fact. The searching problem is a typical structureless problem. Indeed, we search for an item in a database without any structure. Fortunately, quantum mechanics helps this search giving a significant quadratic speedup. Unfortunately, it turns out that this gain is the maximum possible. On the other hand, Shor's algorithm exploits a structure hidden in the factoring problem. This structure allows the reduction of the integer-factoring problem to the problem of finding the period of a particular function. The natural way to find the period of a function is to compute its Fourier transform and, as we have seen, the Fourier transform can be implemented efficiently on a quantum computer. This being said, we do not know the answer to the following fundamental question: What class of problems can be simulated efficiently on a quantum computer? Are there other problems for which the quantum computer gives an exponential gain, beyond those based on the quantum Fourier transform?

## 4.7 Quantum computation of dynamical systems

In this section, we show that a quantum computer would be useful to simulate the dynamical evolution of quantum systems. The simulation of quantum many-body problems on a classical computer is a difficult task as the size of the Hilbert space grows exponentially with the number of particles. For instance, if we wish to simulate a chain of  $n$  spin- $\frac{1}{2}$  particles, the size of the Hilbert space is  $2^n$ . Namely, the state of this system is determined by  $2^n$  complex numbers. As observed by Feynman (1982), the growth in memory requirement is only linear on a quantum computer, which is itself a many-body quantum system. For example, to simulate  $n$  spin- $\frac{1}{2}$  particles we only need  $n$  qubits. Therefore, a quantum computer operating with only a few tens of qubits can outperform a classical computer. Of course, this is only true if we can find an efficient quantum algorithm and if we can efficiently extract useful information from the quantum computer.

### 4.7.1 Quantum simulation of the Schrödinger equation

To be concrete, let us consider the quantum-mechanical motion of a particle in one dimension (the extension to higher dimensions is straightforward). It is governed

by the Schrödinger equation

$$i\hbar \frac{d}{dt} \psi(x, t) = H \psi(x, t), \quad (4.87)$$

where the Hamiltonian  $H$  is given by

$$H = H_0 + V(x) = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x). \quad (4.88)$$

The Hamiltonian  $H_0 = -(\hbar^2/2m)d^2/dx^2$  governs the free motion of the particle, while  $V(x)$  is a one-dimensional potential. To solve Eq. (4.87) on a quantum computer with finite resources (a finite number of qubits and a finite sequence of quantum gates), we must first of all discretize the continuous variables  $x$  and  $t$ . If the motion essentially takes place inside a finite region, say  $-d \leq x \leq d$ , we decompose this region into  $2^n$  intervals of length  $\Delta = 2d/2^n$  and represent these intervals by means of the Hilbert space of an  $n$ -qubit quantum register (this means that the discretization step drops exponentially with the number of qubits). Hence, the wave function  $|\psi(t)\rangle$  is approximated as follows:

$$|\tilde{\psi}(t)\rangle = \sum_{i=0}^{2^n-1} c_i(t) |i\rangle = \frac{1}{\mathcal{N}} \sum_{i=0}^{2^n-1} \psi(x_i, t) |i\rangle, \quad (4.89)$$

where

$$x_i \equiv -d + \left(i + \frac{1}{2}\right) \Delta, \quad (4.90)$$

$|i\rangle = |i_{n-1}\rangle \otimes |i_{n-2}\rangle \otimes \cdots \otimes |i_0\rangle$  is a state of the computational basis of the  $n$ -qubit quantum register and

$$\mathcal{N} \equiv \sqrt{\sum_{i=0}^{2^n-1} |\psi(x_i, t)|^2} \quad (4.91)$$

is a factor that ensures correct normalization of the wave function. It is intuitive that  $|\tilde{\psi}\rangle$  provides a good approximation to  $|\psi\rangle$  when the discretization step  $\Delta$  is much smaller than the shortest length scale relevant for the motion of the system.

As we have seen in Sec. 2.3, the Schrödinger equation (4.87) may be integrated formally by propagating the initial wave function  $\psi(x, 0)$  for each time-step  $\epsilon$  as follows:

$$\psi(x, t + \epsilon) = e^{-\frac{i}{\hbar}[H_0 + V(x)]\epsilon} \psi(x, t). \quad (4.92)$$

If the time-step  $\epsilon$  is small enough, so that terms of order  $\epsilon^2$  may be neglected, it is possible to write

$$e^{-\frac{i}{\hbar}[H_0 + V(x)]\epsilon} \approx e^{-\frac{i}{\hbar}H_0\epsilon} e^{-\frac{i}{\hbar}V(x)\epsilon}. \quad (4.93)$$

Note that this equation, known as the Trotter decomposition, is only exact up to terms of order  $\epsilon^2$  since the operators  $H_0$  and  $V$  do not commute. The operator on the right-hand side of Eq. (4.93) is still unitary, simpler than that on the left-hand side, and, in many interesting physical problems, can be efficiently implemented

on a quantum computer. We take advantage of the fact that, as we have seen in Sec. 4.3, the Fourier transform can be efficiently preformed by a quantum computer. We call  $k$  the variable conjugate to  $x$ , that is,  $-i(d/dx) = F^{-1}kF$ , where  $F$  is the Fourier transform. Therefore, we can write the first operator in the right-hand side of (4.93) as

$$e^{-\frac{i}{\hbar}H_0\epsilon} = F^{-1} e^{+\frac{i}{\hbar}\left(\frac{\hbar^2 k^2}{2m}\right)\epsilon} F. \quad (4.94)$$

In this expression, we pass, by means of the Fourier transform  $F$ , from the  $x$ -representation to the  $k$ -representation, in which this operator is diagonal. Then, using the inverse Fourier transform  $F^{-1}$ , we return to the  $x$ -representation, in which the operator  $\exp(-iV(x)\epsilon/\hbar)$  is diagonal. The wave function  $\psi(x, t)$  at time  $t = l\epsilon$  is obtained from the initial wave function  $\psi(x, 0)$  by applying  $l$  times the unitary operator

$$F^{-1} e^{+\frac{i}{\hbar}\left(\frac{\hbar^2 k^2}{2m}\right)\epsilon} F e^{-\frac{i}{\hbar}V(x)\epsilon}. \quad (4.95)$$

Therefore, simulation of the Schrödinger equation is now reduced to the implementation of the Fourier transform plus diagonal operators of the form

$$|x\rangle \rightarrow e^{icf(x)} |x\rangle, \quad (4.96)$$

where  $c$  is some real constant. Note that an operator of the form (4.96) appears both in the computation of  $\exp(-iV(x)\epsilon/\hbar)$  and of  $\exp(-iH_0\epsilon/\hbar)$ , when this latter operator is written in the  $k$ -representation. The construction of the Fourier transform was discussed in Sec. 4.3 and requires  $O(n^2)$  elementary quantum gates (Hadamard and controlled phase-shift gates). The quantum computation of (4.96) is possible, using an ancillary quantum register  $|y\rangle_a$ , by means of the following steps:

$$\begin{aligned} |0\rangle_a \otimes |x\rangle &\rightarrow |f(x)\rangle_a \otimes |x\rangle \rightarrow e^{icf(x)} |f(x)\rangle_a \otimes |x\rangle \\ &\rightarrow e^{icf(x)} |0\rangle_a \otimes |x\rangle = |0\rangle_a \otimes e^{icf(x)} |x\rangle. \end{aligned} \quad (4.97)$$

The first step is a standard function evaluation and, as we have discussed in Sec. 3.9, may be implemented by means of  $O(2^n)$  generalized  $C^n$ -NOT gates. Of course, more efficient implementations (polynomial in  $n$ ) are possible when the function  $f(x)$  has some structure. This is the case for the potentials  $V(x)$  usually considered in quantum-mechanical problems. The second step in (4.97) is the transformation  $|y\rangle_a \rightarrow e^{icy}|y\rangle_a$  and can be performed in  $m$  single-qubit phase-shift gates,  $m$  being the number of qubits in the ancillary register. Indeed, we may write the binary decomposition of an integer  $y \in [0, 2^m - 1]$  as  $y = \sum_{j=0}^{m-1} y_j 2^j$ , with  $y_j \in \{0, 1\}$ . Therefore,

$$\exp(iy) = \exp\left(\sum_{j=0}^{m-1} icy_j 2^j\right) = \prod_{j=0}^{m-1} \exp(icy_j 2^j), \quad (4.98)$$

which is the product of  $m$  single-qubit gates, each acting non-trivially (differently from the identity) only on a single qubit. The  $j$ -th gate operates the transformation

$|y_j\rangle_a \rightarrow \exp(icy_j 2^j)|y_j\rangle_a$ , with  $|y_j\rangle_a \in \{|0\rangle, |1\rangle\}$  vectors of the computational basis for the  $j$ -th ancillary qubit. The matrix representation of this phase shift gate is given by

$$R_z(c 2^j) = \begin{bmatrix} 1 & 0 \\ 0 & \exp(ic 2^j) \end{bmatrix}. \quad (4.99)$$

The third step in (4.97) is just the reverse of the first and may be implemented by the same array of gates as the first but applied in the reverse order. After this step the ancillary qubits are returned to their standard configuration  $|0\rangle_a$  and it is therefore possible to use the same ancillary qubits for every time-step. Note that the number of ancillary qubits  $m$  determines the resolution in the computation of the diagonal operator (4.96). Indeed, the function  $f(x)$  appearing in (4.96) is discretized and can take  $2^m$  different values.

An alternative method for the quantum computation of (4.96), which does not uses ancillary qubits, is provided by the quantum circuit setting the phases of a state, described in Sec. 3.7.1. Indeed, any operator of the form in Eq. (4.96) can be implemented by means of  $2^n/2$  generalized controlled-phase shift gates, which apply the transformation  $\Gamma_k$  to the target qubit only when the  $n - 1$  (control) qubits are in the state  $|k\rangle$ . Here  $\Gamma_k$  is a single-qubit gate, which maps  $|0\rangle$  into  $e^{if(2k)}|0\rangle$  and  $|1\rangle$  into  $e^{if(2k+1)}|1\rangle$ . It is easy to check this construction for the three-qubit case, by means of the quantum circuit drawn in Fig. 3.17.  $\Gamma_0$  acts only when the first two qubits are in the state  $|00\rangle$ , and therefore it sets the phases  $e^{if(0)}$  and  $e^{if(1)}$  in front of the basis vectors  $|000\rangle$  and  $|001\rangle$ , respectively. Similarly,  $\Gamma_1$  acts only when the first two qubits are in the state  $|01\rangle$ , and therefore it sets the phase  $e^{if(2)}$  and  $e^{if(3)}$  in front of the basis vectors  $|010\rangle$  and  $|011\rangle$  and so on.

The implementation described in Fig. 3.17 is inefficient, because it scales exponentially with the number of qubits. However, efficient (that its, polynomial in  $n$ ) implementations are possible for most of the cases of physical interest. For instance,  $n^2$  two-qubit phase-shift gates are sufficient for the harmonic oscillator potential  $V(x) = \frac{1}{2}m\omega^2x^2$ . If we use Eqs. (4.90), we can write the discretized variable  $x$  as  $\alpha \sum_{j=0}^{n-1} (i_j 2^j + \beta)$ , with  $i_j \in \{0, 1\}$  ( $|i\rangle = |i_{n-1}, \dots, i_0\rangle$ ), the constants  $\alpha = \Delta$  and  $\beta = \frac{-d+\Delta/2}{\alpha n}$ . Therefore,  $x^2 = \alpha^2 \sum_{j,l=0}^{n-1} (i_j 2^j + \beta)(i_l 2^l + \beta)$ , and

$$e^{-iV(x)\epsilon/\hbar} = \prod_{j,l=0}^{n-1} e^{-i\gamma(i_j 2^j + \beta)(i_l 2^l + \beta)}, \quad (4.100)$$

where  $\gamma = m\omega^2\alpha^2\epsilon/(2\hbar)$ . The right-hand side of Eq. (4.100) is the product of  $n^2$  phase-shift gates, each acting non-trivially (differently from the identity) only on the qubits  $j$  and  $l$ . Because the kinetic energy  $H_0$  is proportional to  $p^2 \equiv (\hbar k)^2$ , with  $p$  momentum of the particle, an analogous decomposition is readily obtained in the momentum eigenbasis for  $\exp(-iH_0\epsilon/\hbar)$ . Efficient implementations are possible for piecewise analytic potentials  $V(x)$  but require, in general, the use of ancillary qubits.

There is an exponential advantage in memory requirements with respect to classical computation. A classical computer needs  $O(N = 2^n)$  bits to load the state vector of a system of size  $N$  (that is, the coefficients  $\psi(x_i, t)$  of its expansion over the computational basis). In contrast, a quantum computer accomplishes the same task with just  $n = \log_2 N$  qubits, namely, with memory resources only logarithmic in the system size.

It turns out that standard quantum-mechanical problems can be simulated with sufficient resolution using only a small number of qubits ( $n \approx 10$ ) to discretize the coordinate  $x$ . To illustrate this point, we simulate the working of a quantum computer on a classical computer, with an exponential slowing down with respect to a true quantum computation. The inefficiency of classical simulation is due to the fact that, to simulate the action of an elementary quantum gate on a classical computer, we need update all  $N = 2^n$  coefficients in the wave function expansion (4.89). Plots of  $|\psi(x, t)|^2$  are shown in Fig. 4.7 for a few examples of single-particle quantum mechanical problems. In contrast to classical simulations, in a quantum computation we cannot access the wave function  $\psi(x, t)$  after a single run up to time  $t$ . Each run is followed by a standard projective measurement on the computational basis, giving the outcome  $x_i$  with probability  $|\langle i | \psi(t) \rangle|^2$  (more precisely,  $|\psi(t)\rangle$  is approximated by  $|\tilde{\psi}(t)\rangle$  as explained above). Therefore, the probability distribution  $|\psi(x, t)|^2$  may be reconstructed, up to statistical errors, only if the quantum simulation is repeated many times. If outcome  $x_i$  is obtained  $M_i$  times in  $M$  runs, we can estimate  $|\psi(x_i, t)|^2$  as  $\mathcal{N}^2 \frac{M_i}{M}$ , with the  $\mathcal{N}$  normalization factor defined in Eq. (4.91).

Figure 4.7 exhibits several interesting features of quantum mechanics. In Fig. 4.7(a) we can see the spreading (quadratic in time) of a Gaussian wave packet for an accelerated particle. In Fig. 4.7(b) interference fringes appear when a Gaussian packet impinges on a square barrier. Figure 4.7(c) shows the width oscillations for a squeezed state, moving back and forth between a minimum and a maximum value. Figure 4.7(d) illustrates the motion of a wave packet in a harmonic potential for  $x > 0$  and anharmonic potential for  $x < 0$ . We can see the deformation and spreading of the wave packet when it moves in the anharmonic part of the potential.

In the following subsections, we shall discuss two interesting dynamical systems, the baker's map and the sawtooth map, that can be simulated on a quantum computer without ancillary qubits. In these models interesting physical phenomena can already be studied with 3–10 qubits and a few tens–hundreds of quantum gates.

#### 4.7.2 \* The quantum baker's map

In this subsection, we show that the quantum baker's map can be simulated on a quantum computer in a particularly simple and efficient way. The quantum algorithm for the quantum baker's map computes the dynamical evolution of this system exponentially faster than any known classical computation.

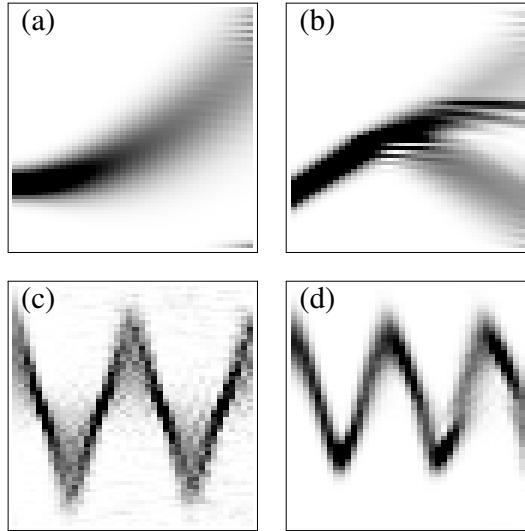


Fig. 4.7 Plots of  $|\psi(x, t)|^2$  with  $n = 6$  qubits. The time  $t$  is plotted on the horizontal axis, which is divided into 40 time steps; The vertical  $x$  axis is discretized by a grid of  $2^n = 64$  points, with  $-5 \leq x \leq 5$ . (a) Accelerated particle, (b) transmission and reflection through a square barrier, (c) harmonic potential, with a squeezed state of initial width twice that of a coherent state, and (d) anharmonic potential. The initial state  $\psi(x, 0)$  in all cases is a Gaussian wave packet,  $\psi(x, t) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \exp\left\{-\frac{(x-x_0)^2}{2\sigma^2} + \frac{i}{\hbar}[p_0(x-x_0)]\right\}$ . Parameter values: (a)  $x_0 = -2.5$ ,  $p_0 = 0$ ,  $\sigma = 0.5$ ,  $V(x) = -Fx$  ( $F = 4.8$ ), time step  $\epsilon = \pi/100$ , (b)  $x_0 = -3$ ,  $p_0 = 10$ ,  $\sigma = 0.5$ ,  $V(x) = V_0$  for  $-2.66 < x < -1.875$  ( $V_0 = 43.9$ ),  $V(x) = 0$  otherwise,  $\epsilon = \pi/200$ , (c)  $x_0 = 2.5$ ,  $p_0 = 0$ ,  $\sigma = 2$ ,  $V(x) = \frac{1}{2}\omega_0^2x^2$  ( $\omega_0 = 1$ ),  $\epsilon = \pi/10$ , (d)  $x_0 = 2.5$ ,  $p_0 = 0$ ,  $\sigma = 1$ ,  $V(x)$  as in (c) for  $x < 0$ ,  $V(x) = \alpha x^3$  ( $\alpha = -0.4$ ) for  $x > 0$ ,  $\epsilon = \pi/10$ . We set  $\hbar = m = 1$ . The figure is reproduced from Benenti and Strini (2008), with the permission of the American Association of Physics Teachers.

The classical baker's transformation maps the unit square  $0 \leq q, p < 1$  onto itself according to

$$(q, p) \rightarrow (\bar{q}, \bar{p}) = \begin{cases} (2q, \frac{1}{2}p), & \text{if } 0 \leq q \leq \frac{1}{2}, \\ (2q - 1, \frac{1}{2}p + \frac{1}{2}), & \text{if } \frac{1}{2} < q < 1. \end{cases} \quad (4.101)$$

This corresponds to compressing the unit square in the  $p$  direction and stretching it in the  $q$  direction, then cutting it along the  $p$  direction and finally stacking one piece on top of the other (similarly to the way a baker kneads dough). The baker's map is a paradigmatic model of classical chaos. Indeed, it exhibits sensitive dependence on initial conditions, which is the distinctive feature of classical chaos: any small error in determining the initial conditions is amplified exponentially in time. The dynamics is uniformly unstable, inducing contraction in the  $p$  direction and stretching in the  $q$  direction.

The baker's map can be quantized following Balazs and Voros (1989) and Saraceno (1990). We introduce the position ( $q$ ) and momentum ( $p$ ) operators, and denote the eigenstates of these operators by  $|q_j\rangle$  and  $|p_k\rangle$ , respectively. The corresponding eigenvalues are given by  $q_j = j/N$  and  $p_k = k/N$ , with  $j, k = 0, \dots, N-1$ ,

$N$  being the dimension of the Hilbert space. We take  $N = 2^n$ , where  $n$  is the number of qubits used to simulate the quantum baker's map on a quantum computer. The transformation between the position basis  $\{|q_0\rangle, \dots, |q_{N-1}\rangle\}$  and the momentum basis  $\{|p_0\rangle, \dots, |p_{N-1}\rangle\}$  is performed by means of a discrete Fourier transform  $F_n$ , defined by the matrix elements

$$\langle q_k | F_n | q_j \rangle \equiv \frac{1}{\sqrt{2^n}} \exp\left(\frac{2\pi i k j}{2^n}\right). \quad (4.102)$$

It can be shown that the quantized baker's map may be defined by the transformation (see Balazs and Voros, 1989)

$$|\psi\rangle \rightarrow |\bar{\psi}\rangle = T|\psi\rangle, \quad (4.103)$$

where  $|\bar{\psi}\rangle$  denotes the wave vector of the system after application of one map step to the state  $|\psi\rangle$ . The unitary transformation  $T$  defining the quantum baker's map is given by

$$T = F_n^{-1} \begin{bmatrix} F_{n-1} & 0 \\ 0 & F_{n-1} \end{bmatrix}, \quad (4.104)$$

where the matrix elements are to be understood relative to the position basis  $\{|q_j\rangle\}$  and  $F_{n-1}$  is the discrete Fourier transform defined by Eq. (4.102).

The unitary transformation  $T$  can be implemented on a quantum computer with  $n$  qubits. We have

$$T = F_n^{-1} (I \otimes F_{n-1}), \quad (4.105)$$

where  $F_{n-1}$  acts on the  $n - 1$  least significant qubits and  $I$  is the identity operator acting on the most significant qubit. Since the Fourier transform  $F_n$  can be implemented efficiently on a quantum computer with  $O(n^2)$  elementary gates, the quantum baker's map can also be simulated efficiently on a quantum computer with  $O(n^2)$  elementary gates per map iteration. Note that the best known classical algorithm to simulate the Fourier transform, the fast Fourier transform, requires  $O(n2^n)$  elementary operations. Therefore, the dynamics of the baker's map can be simulated on a quantum computer with an exponential gain with respect to any known classical computation.

#### 4.7.3 \* The quantum sawtooth map

An example of an interesting dynamical model that can be simulated efficiently (and without ancillary qubits) on a quantum computer is the so-called quantum sawtooth map. This map represents the dynamics of a periodically driven system and is derived from the Hamiltonian

$$H(\theta, I; \tau) = \frac{I^2}{2} + V(\theta) \sum_{j=-\infty}^{+\infty} \delta(\tau - jT), \quad (4.106)$$

where  $(I, \theta)$  are conjugate action-angle variables ( $0 \leq \theta < 2\pi$ ), with the usual quantization rules,  $\theta \rightarrow \theta$  and  $I \rightarrow I = -i\partial/\partial\theta$  (set  $\hbar = 1$ ) and  $V(\theta) = -k(\theta - \pi)^2/2$ .

This Hamiltonian is the sum of two terms,  $H(\theta, I; \tau) = H_0(I) + U(\theta; \tau)$ , where  $H_0(I) = I^2/2$  is just the kinetic energy of a free rotator (a particle moving on a circle parametrized by the coordinate  $\theta$ ), while  $U(\theta; \tau) = V(\theta) \sum_j \delta(\tau - jT)$  represents a force acting on the particle that is switched on and off instantaneously at time intervals  $T$ . Therefore, it is said that the dynamics described by Hamiltonian (4.106) is kicked. The (quantum) evolution from time  $tT^-$  (prior to the  $t$ -th kick) to time  $(t+1)T^-$  (prior to the  $(t+1)$ -th kick) is described by a unitary operator  $U$  acting on the wave function  $\psi$ :

$$\begin{aligned}\psi_{t+1} &= U\psi_t = U_T U_k \psi_t; \\ U_T &= e^{-iT\hat{I}^2/2}, \quad U_k = e^{ik(\theta-\pi)^2/2}.\end{aligned}\quad (4.107)$$

This map is called the quantum sawtooth map, since the force  $F(\theta) = -dV(\theta)/d\theta = k(\theta - \pi)$  has a sawtooth shape, with a discontinuity at  $\theta = 0$ .

In the following, we describe an exponentially efficient quantum algorithm for simulation of the map (4.107). It is based on the forward/backward quantum Fourier transform between action (momentum) and angle bases. Such an approach is convenient since the operator  $U$ , introduced in Eq. (4.107), is the product of two operators,  $U_k$  and  $U_T$ , diagonal in the  $\theta$  and  $I$  representations, respectively. This quantum algorithm requires the following steps for one map iteration:

- (1) We apply  $U_k$  to the wave function  $\psi(\theta)$ . In order to decompose the operator  $U_k$  into one- and two-qubit gates, we first of all write  $\theta$  in binary notation:

$$\theta = 2\pi \sum_{j=1}^n \alpha_j 2^{-j}, \quad (4.108)$$

with  $\alpha_i \in \{0, 1\}$ . Here  $n$  is the number of qubits, so that the total number of levels in the quantum sawtooth map is  $N = 2^n$ . From this expansion, we obtain

$$(\theta - \pi)^2 = 4\pi^2 \sum_{j_1, j_2=1}^n \left( \frac{\alpha_{j_1}}{2^{j_1}} - \frac{1}{2n} \right) \left( \frac{\alpha_{j_2}}{2^{j_2}} - \frac{1}{2n} \right). \quad (4.109)$$

This term can be put into the unitary operator  $U_k$ , giving the decomposition

$$e^{ik(\theta-\pi)^2/2} = \prod_{j_1, j_2=1}^n \exp \left[ i2\pi^2 k \left( \frac{\alpha_{j_1}}{2^{j_1}} - \frac{1}{2n} \right) \left( \frac{\alpha_{j_2}}{2^{j_2}} - \frac{1}{2n} \right) \right], \quad (4.110)$$

which is the product of  $n^2$  two-qubit gates (controlled phase-shift gates), each acting non-trivially only on the qubits  $j_1$  and  $j_2$ . In the computational basis  $\{|\alpha_{j_1} \alpha_{j_2}\rangle = |00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  each two-qubit gate can be written as  $\exp(i2\pi^2 k D_{j_1, j_2})$ , where  $D_{j_1, j_2}$  is a diagonal matrix:

$$D_{j_1, j_2} = \begin{bmatrix} \frac{1}{4n^2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2n} \left( \frac{1}{2^{j_2}} - \frac{1}{2n} \right) & 0 & 0 \\ 0 & 0 & -\frac{1}{2n} \left( \frac{1}{2^{j_1}} - \frac{1}{2n} \right) & 0 \\ 0 & 0 & 0 & \left( \frac{1}{2^{j_1}} - \frac{1}{2n} \right) \left( \frac{1}{2^{j_2}} - \frac{1}{2n} \right) \end{bmatrix}. \quad (4.111)$$

- (2) The change from the  $\theta$  to the  $I$  representation is obtained by means of the quantum Fourier transform, which, as we know, requires  $n$  Hadamard gates and  $\frac{1}{2}n(n - 1)$  controlled phase-shift gates.
- (3) In the  $I$  representation, the operator  $U_T$  has essentially the same form as the operator  $U_k$  in the  $\theta$  representation and can therefore be decomposed into  $n^2$  controlled phase-shift gates, similarly to Eq. (4.110).
- (4) We return to the initial  $\theta$  representation by application of the inverse quantum Fourier transform.

Thus, overall, this quantum algorithm requires  $3n^2 + n$  gates per map iteration ( $3n^2 - n$  controlled phase-shifts and  $2n$  Hadamard gates). This number is to be compared with the  $O(n2^n)$  operations required by a classical computer to simulate one map iteration by means of a fast Fourier transform. Thus, the quantum simulation of the quantum sawtooth map dynamics is exponentially faster than any known classical algorithm. Note that the resources required to the quantum computer to simulate the sawtooth map are only logarithmic in the system size  $N$ .

As an example of the efficiency of this quantum algorithm, Fig. 4.8 shows the Husimi function,<sup>5</sup> taken after 1000 map iterations. It is noted that  $n = 9$  qubits are sufficient to observe the appearance of integrable islands, while at  $n = 16$  these islands exhibit a complex hierarchical structure in the phase space. Of course, there remains the problem of extracting useful information from the quantum-computer wave function. This will be discussed in the next subsection.

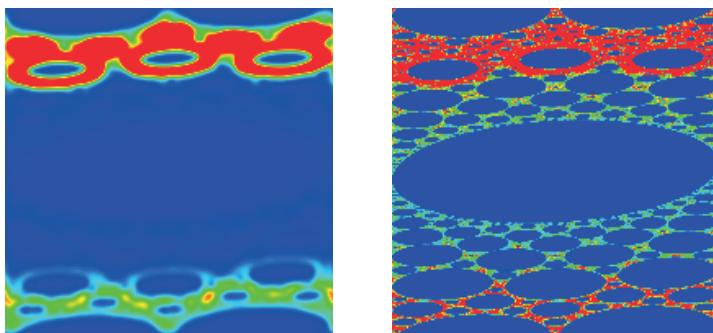


Fig. 4.8 Quantum phase space probability distribution (Husimi function) for the sawtooth map for  $n = 9$  (left) and  $n = 16$  (right) qubits, in action angle variables  $(I, \theta)$ , with  $-N/2 \leq I < N/2$  (vertical axis,  $N = 2^n$ ) and  $0 \leq \theta < 2\pi$  (horizontal axis), averaged in the interval  $950 \leq t \leq 1000$ , for  $T = 2\pi/N$  and  $kT = -0.1$  (here  $\hbar = 1$ ). A momentum eigenstate,  $|\psi(0)\rangle = |m_0\rangle$ , with  $m_0 = [0.38N]$  is considered as initial state at time  $t = 0$ . The color is proportional to the density: blue for zero and red for maximal density. The figure is reprinted with permission from Benenti et al. (2001). ©(2001) by the American Physical Society.

<sup>5</sup>The Husimi function is a “quantum phase-space probability distribution”, obtained at any given point  $(\theta, I)$  by the projection of the quantum state on the coherent state centred at that point. This corresponds to the smoothing of the Wigner function on the scale of the Planck constant (see Chang and Shi, 1986).

#### 4.7.4 Information extraction for dynamical quantum systems

Any quantum algorithm has to address the problem of efficiently extracting useful information from the quantum computer wave function. Indeed, the result of the simulation of a quantum system is the wave function of this system, encoded in the  $n$  qubits of the quantum computer. The problem is that, in order to measure all  $N = 2^n$  wave function coefficients by means of standard polarization measurements of the  $n$  qubits, one has to repeat the quantum simulation a number of times exponential in the number of qubits. This procedure would spoil any quantum algorithm, even in the case, like the present one, in which such algorithm could compute the wave function with an exponential gain with respect to any classical computation. Nevertheless, there are some important physical questions that can be answered in an efficient way.

The quantum computation can provide an exponential gain (with respect to any known classical computation) in problems that require the simulation of dynamics up to a time  $t$  which is independent of the number of qubits. In this case, provided that one can extract the relevant information in a number of measurements polynomial in the number of qubits, one should compare, in the example of the quantum sawtooth map,  $O(t(\log N)^2)$  elementary gates (quantum computation) with  $O(tN \log N)$  elementary gates (classical computation). This is, for instance, the case of the fidelity of quantum motion, which is a quantity of central interest in the study of the stability of quantum motion under perturbations. The fidelity  $f(t)$  (also called the Loschmidt echo), measures the accuracy with which a quantum state can be recovered by inverting, at time  $t$ , the dynamics with a perturbed Hamiltonian. It is defined as

$$f(t) = |\langle \psi_0 | U_\epsilon^\dagger(t) U(t) | \psi_0 \rangle|^2, \quad (4.112)$$

where  $U(t)$  is the time-evolution operator up to time  $t$ . Here the initial wave vector  $|\psi_0\rangle$  evolves forward in time with the Hamiltonian  $H$  up to time  $t$  and then evolves backward in time with a perturbed Hamiltonian  $H_\epsilon$  ( $U_\epsilon$  is the corresponding time-evolution operator). If the evolution operators  $U$  and  $U_\epsilon$  can be simulated efficiently on a quantum computer, as is the case in most physically interesting situations, then the fidelity of quantum motion can be evaluated with exponential speed up with respect to known classical computations.

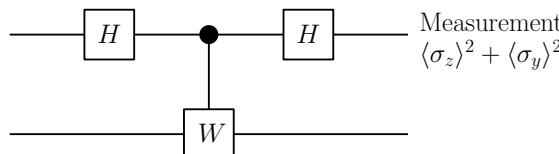


Fig. 4.9 Schematic drawing of a quantum circuit implementing a Ramsey-type quantum interferometer. The top line denotes a single ancillary qubit, the bottom line a set of  $n$  qubits,  $H$  the Hadamard gate and  $W$  a unitary transformation.

As shown in Fig. 4.9, it is possible to measure the fidelity by means of a Ramsey-type quantum interferometer method. A single ancillary qubit is needed, initially prepared in the state  $|0\rangle$ , while the input state for the other  $n$  qubits is a given initial state  $|\psi_0\rangle$  for a quantum dynamical system. Two Hadamard gates are applied to the ancillary qubit, and in between these operations a controlled- $W$  operation is applied ( $W$  is a unitary operator), namely  $W$  is applied to the other  $n$  qubits only if the ancillary qubit is in its  $|1\rangle$  state. As a result, one obtains the following final overall state for the  $n + 1$  qubits:

$$\frac{1}{2} [(|0\rangle + |1\rangle)|\psi_0\rangle + (|0\rangle - |1\rangle)W|\psi_0\rangle]. \quad (4.113)$$

If  $W = U_\epsilon^\dagger(t)U(t)$ , then one can derive the fidelity from polarization measurements of the ancillary qubit. Specifically, one obtains

$$\langle\sigma_z\rangle = \text{Re}[\langle\psi_0|W|\psi_0\rangle], \quad \langle\sigma_y\rangle = -\text{Im}[\langle\psi_0|W|\psi_0\rangle], \quad (4.114)$$

where  $\langle\sigma_z\rangle$  and  $\langle\sigma_y\rangle$  are the expectation values of the Pauli operators  $\sigma_z$  and  $\sigma_y$ . It is thus immediate to see that

$$f(t) = \langle\sigma_z\rangle^2 + \langle\sigma_y\rangle^2, \quad (4.115)$$

provided that the quantum algorithms implementing  $U$  and  $U_\epsilon$  are efficient, the fidelity can then be computed efficiently. Measuring the ancilla once provides an estimate of the fidelity to one bit of accuracy. Repeating the computation will build up a more accurate estimate, by combining all the outcomes.

Proceeding in a similar fashion, it is also possible to efficiently compute the dynamical correlation functions of the form

$$C(t) \equiv \langle\psi_0|A^\dagger(t)B(0)|\psi_0\rangle = \langle\psi_0|U^\dagger(t)A^\dagger(0)U(t)B(0)|\psi_0\rangle, \quad (4.116)$$

with the same computational efforts needed for the fidelity. The circuit implementing this measure is provided in Fig. 4.10. A single-qubit ancilla is needed to control the conditional application of  $B = B(0)$  and  $A^\dagger = A^\dagger(0)$ , between which the time evolution  $U(t)$  is performed. The operators  $A$  and  $B$  have to be expressible as a sum of unitary operators. After the application of the various gates, the required correlation function can be obtained from a simple polarization measurement of the ancilla, indeed it is immediate to see that

$$\langle\sigma_x\rangle = \text{Re}[\langle U^\dagger(t)AU(t)B\rangle], \quad \langle\sigma_y\rangle = \text{Im}[\langle U^\dagger(t)AU(t)B\rangle], \quad (4.117)$$

so that

$$\langle U^\dagger(t)AU(t)B\rangle = \langle\sigma_x\rangle + i\langle\sigma_y\rangle = 2\langle\sigma_+\rangle. \quad (4.118)$$

Note that, by replacing  $U(t)$  with the space translation operator, spatial correlations instead of time correlations can be obtained.

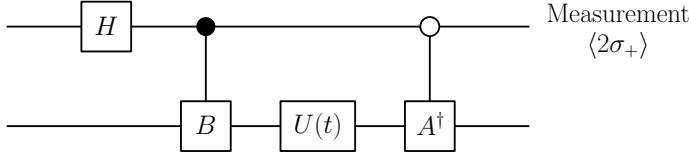


Fig. 4.10 Schematic drawing of a quantum circuit for measuring the correlation function in Eq. (4.116). The top line denotes a single ancillary qubit, the bottom line a set of  $n$  qubits, and  $H$  the Hadamard gate. After the ancilla has been measured in the  $\sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y)$  basis, the outcome  $\langle 2\sigma_+ \rangle$  provides an estimate of the correlation function.

## 4.8 Universal quantum simulation

There is an alternative strategy to efficiently simulate the time evolution of a large variety of discrete quantum systems, by means of a quantum computer. Such systems can be composed of several interacting particles or objects, a scenario which is very difficult to describe by means of analytical calculations, and is usually addressed by inefficient classical simulations. The essential requirement is the locality of the interactions, as we shall quantify below. As already noted before, this kind of simulation is an exponentially difficult problem, since just to record the generic state of  $n$  qubits in a classical computer memory would require  $2^n$  complex numbers. Moreover, calculating its time evolution under a given unitary dynamics requires the exponentiation of a  $2^n \times 2^n = 2^{2n}$  matrix.

The idea to bypass this exponentially growing amount of computational resources was originally put forward by Feynman (1982), who conjectured that it is possible to use a quantum system to directly simulate another one in an efficient way. The principle underlying the concept of *quantum simulator* lies in the fact that, according to this proposal, the states of the simulator would naturally obey the same equations of motion of the simulated system. This statement was formally proved by Lloyd (1996), thus validating the conjecture of an exponential speedup of the simulation of quantum systems by means of a quantum simulator, with respect to a classical computer. Below we detail the basic mechanism of universal quantum simulation.

We start by defining the basic processing job as the (unitary) time evolution of the initial wave function  $|\psi(0)\rangle$  of a quantum system, under a given Hamiltonian  $H$ :

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle, \quad (4.119)$$

where we supposed for simplicity that  $H$  is time-independent (the argument can be easily extended to explicitly time-dependent Hamiltonians). While obviously a quantum computer can store an arbitrary quantum state  $|\psi\rangle$  efficiently as compared with a classical computer, it is also crucial that the computation on this superposition state can be performed efficiently. By simply turning on and off the correct sequence of Hamiltonians, a system can be made to evolve according to any unitary operator. However, an arbitrary unitary operator requires exponentially many

parameters to specify it, so in general the algorithm is not efficient. A fundamental requirement, in order to guarantee the efficiency of the computation, is that the system under study has local interactions. Specifically, all the Hamiltonian evolutions satisfying this constraint can be written according to

$$H = \sum_{j=1}^{\ell} H_j, \quad (4.120)$$

where each of the  $\ell$  local Hamiltonians  $H_j$  acts on a limited portion of the system's degrees of freedom, containing at most  $k$  of the total of  $n$  variables (or objects). The  $k$  objects need to not be spatially localised: the requirement of locality here means only that  $k$  does not have to scale with  $n$ . The maximum number  $\ell_{\max}$  of possible distinct terms  $H_j$  entering the expansion (4.120) is thus given by a binomial coefficient and is polynomial in  $n$ , since

$$\ell_{\max} = \binom{n}{k} < \frac{n^k}{k!}. \quad (4.121)$$

However this is a generous upper bound in most practical cases: for short-range Hamiltonians in which each system interacts with at most  $k$  nearest neighbours, one has that  $\ell_{\max} \sim n$ .

In the same way as any classical simulation of the time evolution of dynamical systems is performed, the quantum simulation proceeds by dividing the total simulation time  $t$  into  $n_\tau$  small discrete slices. Each time step has a temporal duration  $\delta t = t/n_\tau$ , and can be approximated using a Trotter decomposition:

$$e^{-\frac{i}{\hbar} H t} = \left( e^{-\frac{i}{\hbar} H_1 \delta t} e^{-\frac{i}{\hbar} H_2 \delta t} \dots e^{-\frac{i}{\hbar} H_\ell \delta t} \right)^{n_\tau} + \frac{t^2}{2n_\tau} \sum_{i>j} [H_i, H_j] + \sum_{p=3}^{+\infty} \varepsilon(p). \quad (4.122)$$

It can be shown that the higher order error terms  $\varepsilon(p)$  are bounded by

$$\|\varepsilon(p)\|_{\sup} \leq \frac{n_\tau}{p!} \|H \delta t\|_{\sup}^p, \quad (4.123)$$

where  $\|A\|_{\sup}$  defines the supremum, that is, the maximum expectation value of the operator  $A$  over all the normalized quantum states. The total error in approximating  $e^{-\frac{i}{\hbar} H t}$  with the first term in the expansion (4.122) is bounded by  $\|n_\tau(e^{-\frac{i}{\hbar} H \delta t} - 1 - \frac{i}{\hbar} H \delta t)\|_{\sup}$ . As a consequence, by taking  $n_\tau$  sufficiently large, such error can be made as small as required. For a given threshold  $\epsilon$ , the second term in Eq. (4.122) implies that  $\epsilon \propto t^2/n_\tau$ . A reliable first order Trotter simulation thus requires  $n_\tau \propto t^2/\epsilon$ .

It remains to check that the simulation scales efficiently in the number of operations required. The size of the generic Hamiltonian  $H_j$  between  $k$  variables depends on the dimensions of the individual variables, but will be bounded by a maximum size  $g$ . Simulating  $e^{-\frac{i}{\hbar} H_j \delta t}$  requires  $g_j^2$  operations, where  $g_j \leq g$  is the dimension of the variables involved in  $H_j$ . In Eq. (4.122) each local operator  $H_j$  is simulated  $n_\tau$

times. Therefore, the total number of operations required for simulating  $e^{-\frac{i}{\hbar} H t}$  is bounded by  $n_\tau \ell g^2$ . Since  $n_\tau \propto t^2/\epsilon$ , we find that

$$\#\text{Op} = O(t^2 \ell g^2 / \epsilon). \quad (4.124)$$

The only dependence on the system size  $n$  is through  $\ell$ , and we have already observed that  $\ell$  is polynomial in  $n$ . The number of operations is thus efficient by the criterion of polynomial scaling in the problem size. We recall that the simulation method provided here is straightforward but very general. In particular it establishes certain conditions of local Hamiltonians under which it is possible in theory to carry out efficient quantum simulation.

## 4.9 A guide to the bibliography

Deutsch's algorithm was invented by Deutsch (1985) and extended to the  $n$ -qubit case by Deutsch and Jozsa (1992), see also Cleve *et al.* (1998). The extension discussed in Sec. 4.1.2 is due to Grassi and Strini (1999).

The quantum search algorithm was introduced by Grover (1996), see also Grover (1997). The optimality of the quadratic speedup is discussed in Boyer *et al.* (1998) and Zalka (1999). Further developments can be found in Brassard *et al.* (2002).

Useful discussions on the quantum Fourier transform can be found in Coppersmith (1994) and Ekert and Jozsa (1996). The generalization of the quantum Fourier transform over a generic finite Abelian group was found by Kitaev (1995). Other useful references are Jozsa (1998), Ekert and Jozsa (1998) and Bowden *et al.* (2002). The quantum wavelet transform is discussed in Fijany and Williams (1998). A unified approach to fast unitary transforms is described in Agaian and Klappenecker (2002). Signal-processing methods in quantum computing are discussed in Klappenecker and Rötteler (2001).

The phase estimation algorithm was introduced by Kitaev (1995) and a good description of this algorithm can be found in Cleve *et al.* (1998).

The quantum algorithm described in Sec. 4.5, which computes eigenvalues and eigenvectors of a given unitary operator, was introduced by Abrams and Lloyd (1999). Issues concerning the application of this algorithm for molecular Hamiltonians are discussed in Aspuru-Guzik *et al.* (2005).

Shor's algorithm for integer factoring was proposed in Shor (1994) and a detailed discussion of this algorithm can be found in Shor (1997). Other useful references are Kitaev (1995), Ekert and Jozsa (1996), Lomonaco (2000) and Beauregard (2003). A readable introduction to Shor's algorithm is Lavor *et al.* (2003).

The idea that a quantum computer might outperform a classical computer in simulating quantum mechanical systems was proposed by Feynman (1982) and further developed by Lloyd (1996), as discussed in Sec. 4.8. The method to simulate the Schrödinger equation in Sec. 4.7 is due to Zalka (1998) and Wiesner (1996). Benenti and Strini (2008) have shown that six to ten qubits would be sufficient for the simulation of several single-particle one-dimensional problems. The study

of quantum algorithms for the simulation of quantum chaos was started by Schack (1998) and further developed by Georgeot and Shepelyansky (2001a). The quantum algorithms for the simulation of the quantum baker and the quantum sawtooth map are due to Schack (1998) and Benenti *et al.* (2001), respectively. Other quantum algorithms for computing interesting physical quantities in dynamical models can be found in Benenti *et al.* (2003); Emerson *et al.* (2002, 2004). A quantum circuit for quantum-state tomography is discussed in Miquel *et al.* (2002). The quantum simulation of classical chaotic systems is discussed by Georgeot and Shepelyansky (2001b). The simulation of many-body Fermi systems on a quantum computer is investigated by Abrams and Lloyd (1997) and Ortiz *et al.* (2001). The simulation of spin systems is discussed in Sørensen and Mølmer (1999). The problem of simulating the equilibration of quantum systems on a quantum computer is addressed in Terhal and DiVincenzo (2000). The Ramsey-type quantum interferometer method of Sec. 4.7.4 is described in Gardiner *et al.* (1997). A review on quantum simulation is Georgescu *et al.* (2014).

**This page intentionally left blank**

# Chapter 5

## Quantum communication

In this chapter, we show that the basic properties of quantum mechanics can be put to practical use in the transmission of information. The most spectacular application is in the field of cryptography, the art of secret communication. After a short overview of classical cryptography, we discuss the unique contribution of quantum mechanics to cryptography. Quantum cryptography enables two communicating parties, named Alice (the *sender*) and Bob (the *receiver*), to detect whether the transmitted message has been intercepted by Eve (an *eavesdropper*). This is a consequence of a basic property of quantum mechanics, the “no-cloning theorem”: an unknown quantum state cannot be cloned. Then we illustrate two remarkable applications of quantum mechanics: dense coding and quantum teleportation. Dense coding uses entanglement to enhance the communication of classical information. If Alice and Bob share an entangled EPR pair of qubits, Alice can operate on her member of the pair and then send it to Bob: this single qubit carries two bits of classical information. Quantum teleportation again exploits entanglement and allows Alice to transmit a quantum bit to Bob by sending him only classical bits. Finally, after providing the necessary background, we discuss the extension of quantum cryptography from discrete to continuous variables, and hence from finite to infinite dimensional Hilbert spaces.

### 5.1 Classical cryptography

The origins of cryptography go back to before Christ, since when the need for secret communication has become evermore important. A significant example is the *Cæsar cypher*, used more than 2000 years ago by Julius Cæsar during the Gallic war. Such a code uses an alphabet in which each letter is shifted by a fixed number of steps. If there are, for instance,  $k = 26$  letters in the alphabet, then we have  $k - 1 = 25$  possible codes. If we label the letters as  $1, 2, \dots, k$ , then the code number  $j$  is obtained with  $i \rightarrow i + j \pmod{k}$ , for  $i = 1, 2, \dots, k$ . We say that the sender Alice encrypts her *plain text* into a *cypher text*, using a *secret key*. In the case of the Cæsar cypher, the key is  $j$ . It is clear that there are  $k - 1$  different codes, singled out by the number  $j$  (with  $j = 1, \dots, k - 1$ ). The case  $j = k$  is not acceptable

since the original letters are unchanged. This code was difficult to break in Cæsar's time, but of course it is easily breakable today. Note that Alice must first of all communicate the key to Bob over a *secure channel*, assumed inaccessible to the eavesdropper Eve. After this Alice sends the cypher text to Bob over an *insecure channel*.

An interesting variant of the Cæsar cypher replaces the  $k$  letters of the alphabet with one of their  $k!$  possible permutations. However, even this code is easy to break nowadays, since the letters of an alphabet appear with different frequencies in a text. Thus, a simple statistical analysis of the cypher text is sufficient to break the code.

### 5.1.1 The Vernam cypher

The first unbreakable code, the *Vernam cypher*, was invented in 1917 by Gilbert Vernam, even though the mathematical proof of its unbreakability was achieved only more than thirty years later by Shannon. Vernam's protocol is the following.

- (i) The plain text is written as a binary sequence of 0's and 1's.
- (ii) The secret key is a completely random binary sequence of the same length as the plain text.
- (iii) The cypher text is obtained by adding the secret key bitwise modulo 2 to the plain text.

If  $\{p_1, p_2, \dots, p_N\}$  denotes the plain text (with  $p_1, p_2, \dots, p_N$  binary digits) and  $\{k_1, k_2, \dots, k_N\}$  the secret key, then the cypher text  $\{c_1, c_2, \dots, c_N\}$  is obtained as follows:

$$c_i = p_i \oplus k_i \quad (i = 1, 2, \dots, N). \quad (5.1)$$

Let us consider a simple example:

001010011	plain text,
100111010	secret key,
101101001	cypher text.

(5.2)

The code is unbreakable, provided that the key is completely random, since in this case the cypher text is completely random, too. It gives no information whatsoever about the plain text. Since the secret key is shared by Alice and Bob, the latter can easily reconstruct the plain text. He simply adds the secret key bitwise modulo 2 to the cypher text:

$$p_i = c_i \oplus k_i \quad (i = 1, 2, \dots, N). \quad (5.3)$$

We stress that the secret key must be used only once (the Vernam cypher is also known as the *one time pad*). If it is reused and the eavesdropper is able to observe two cypher texts, then their bitwise addition modulo two is equal to the bitwise addition modulo two of the corresponding plain texts. Since in the plain texts there are always redundancies (they are not random binary sequences), the code becomes

breakable. Therefore, the main problem of cryptography is not the transmission of the cypher text but the distribution of the secret key. This distribution requires some kind of “trusted courier”; that is, the problem of the secrecy of communication is merely transferred to the problem of the secrecy of the key. The problem is that Eve could, at least in principle, find a way to read the key without leaving any trace of her action. Therefore, Alice and Bob can never be absolutely sure of the secrecy of the key. We shall see in Sec. 5.3 that quantum mechanics solves this problem, offering a unique way for secure *key distribution* and *key storage*.

The Vernam cypher requires the generation of a long random binary string (at least as long as the message to be transmitted), a non-trivial task in itself. Weaker cyphers using shorter keys are in principle breakable, but possibly hard to break. It is worth noting that the task of breaking sophisticated cyphers was between the motivations that stimulated the construction of electronic computers.

### 5.1.2 The public-key cryptosystem

Owing to the difficulty of supplying new random keys for every message, the Vernam cypher is nowadays used mainly for important diplomatic communications. For less delicate business, it is substituted by *public-key* cryptographic systems, whose principles were discovered in the middle of the 1970’s by Diffie and Hellman.

The fundamental difference between the traditional secret-key cryptosystem and the recent public-key cryptosystem is the following.

- (i) In the secret-key cryptosystem, Alice encrypts her message by means of a secret key. She sends the encrypted message to Bob, who owns the same secret key and can therefore decrypt the message. The security of the message resides in the secrecy of the key. Since the secret key must be at some time distributed between Alice and Bob, there is always a risk that the key is intercepted.
- (ii) In the public-key cryptosystem, Alice and Bob do not exchange any secret key. Bob makes public a key (the public key), used by Alice to encrypt the message. However, the message cannot be decrypted by this key, but only by another key (the private key), which is possessed by Bob alone. Therefore, the key-distribution problem is avoided. The public-key cryptosystem works as if Bob had constructed a safe. The safe has two keys, one public to lock it and another private to open it. Anyone may place a message in the safe, but only one person (Bob) can open the safe and take the message out.

The public-key cryptosystem requires the use of a *trap-door function*  $f$ , easy to compute but with inverse function  $f^{-1}$  hard to compute. Here the words easy and hard must be understood according to the theory of computational complexity (see Chap. 1):  $f$  can be computed with resources polynomial in the input size, while  $f^{-1}$  cannot be computed with polynomial resources (resources denote computational time, size of the hardware *etc.*). Any problem whose solution is hard to find but

easy to verify is in principle useful for cryptography. These problems lie in the computational class **NP**. Two keys are involved: a public key  $f$ , used by Alice to encrypt her text, and a secret key  $f^{-1}$ , possessed by Bob alone, who uses it to decrypt the message.

### 5.1.3 The RSA protocol

A famous example of public-key cryptosystem is the RSA cryptosystem, devised in 1977 by Rivest, Shamir and Adleman. The RSA protocol works as follows:

- (1) Bob chooses two “large enough” prime numbers  $p$  and  $q$ , and computes  $p q = N$ .
- (2) Bob chooses at random a number  $d$  that is co-prime with  $(p - 1)(q - 1)$ , that is, the greatest common divisor of  $d$  and  $(p - 1)(q - 1)$  is equal to 1.
- (3) Bob computes  $e$ , the inverse modulo  $(p - 1)(q - 1)$  of  $d$ :

$$e d|_{\text{mod}(p-1)(q-1)} = 1. \quad (5.4)$$

Note that from now on we can forget about  $p$  and  $q$ .

- (4) Bob publishes the pair  $(e, N)$ . This is the public key, which anybody can use to send messages to Bob.
- (5) The pair  $(d, N)$  is the private decryption key, possessed by Bob alone. Therefore, only Bob can decrypt the messages that were encrypted by means of the public key.
- (6) Alice divides her message into blocks and each block can be written as a number. For the  $i$ -th block, the number is  $m_i$ , with  $m_i < N$ . Alice encrypts each block as follows:

$$m_i \rightarrow m'_i = m_i^e|_{\text{mod } N}. \quad (5.5)$$

- (7) Bob decrypts the message by computing

$$m_i = m'^d_i|_{\text{mod } N}. \quad (5.6)$$

Indeed, elementary number theory tells us that  $m_i^{ed}|_{\text{mod } N} = m_i$  (see, e.g., Ekert *et al.*, 2001).

Note the advantages with respect to the Vernam cypher:

- (i) there is no need to distribute a secret key over a supposedly secure channel: the public key can be used by anybody who wishes to communicate with Bob while the secret key is possessed by Bob alone;
- (ii) the public key can be reused as many times as desired.

**Exercise 5.1** A crucial problem of cryptography is *authentication*: Bob needs to determine if the message received was really sent by Alice and not by someone else. Find how Alice can authenticate her message using a public-key cryptographic system.

The RSA code can be broken if one discovers the prime factors  $p$  and  $q$  of  $N$ . After this  $d$  can be easily computed since  $e$  is known. Therefore, the reliability of the method is based on the fact that there exist no known efficient (polynomial time) algorithms to find the factors of an integer  $N$ : The best classical algorithm known today for integer factorization, the number field sieve, requires  $\exp(O(n^{1/3}(\log n)^{2/3}))$  operations, where  $n = \log N$  is the input size. This means that the problem is in practice impossible to solve with current technology. To be more precise, we should add that it is not proved that the number field sieve algorithm is optimal for integer factorization. Furthermore, the possibility that a polynomial time algorithm can be discovered is not excluded, just as it cannot be excluded for any other **NP** problem. In any event, we should like to stress that, as we have seen in Sec. 4.6, such a polynomial time algorithm exists on a quantum computer. Therefore, a large-scale quantum computer, if constructed, would break the RSA encryption scheme. This is a clear demonstration that the security of public-key cryptosystems is not a sufficient guarantee for messages that must be kept secret for indefinitely long times.

## 5.2 The no-cloning theorem

There is one property of the classical bit that we take for granted: it can be copied. In contrast, we shall see in this section that the *generic* state of a qubit cannot be cloned. This is the content of the so-called no-cloning theorem of Dieks, Wootters and Zurek.

Let us first consider a concrete example, in which the qubit is the polarization state of a photon. We label the state of a photon as  $|0\rangle$  (or  $|\leftrightarrow\rangle$ ) when it is in a horizontally polarized state and as  $|1\rangle$  (or  $|\uparrow\rangle$ ) when it is in a vertically polarized state. A photon can also be polarized along a direction forming an angle  $\beta$  with respect to the horizontal (see Fig. 5.1, where  $x$  denotes the horizontal axis,  $y$  the vertical axis and  $z$  the direction of propagation of the photon). In this case, it is described by the wave function

$$|\psi\rangle = \cos\beta|\leftrightarrow\rangle + \sin\beta|\uparrow\rangle. \quad (5.7)$$

Now assume that such a photon is sent to a polarization analyzer (a birefringent crystal such as calcite). The photon emerges from the analyzer horizontally polarized or vertically polarized (see Fig. 5.1, where horizontally polarized photons pass straight through the crystal, while vertically polarized photons are deflected). These two mutually exclusive outcomes, which we call 0 and 1, take place with probabilities  $p_0 = |\langle\leftrightarrow|\psi\rangle|^2 = \cos^2\beta$  and  $p_1 = |\langle\uparrow|\psi\rangle|^2 = \sin^2\beta$ , respectively. Therefore, a measurement of the polarization state of a single photon gives a single bit of information, corresponding to the polarization state of the detected photon. We note that this is in perfect agreement with the measurement postulate introduced in Sec. 2.3 (Postulate II of quantum mechanics).

Assume now that a cloning machine exists. Then one could make an arbitrarily large number of copies of the state  $|\psi\rangle$  given by Eq. (5.7). Thus, it would be possible

to measure all these clones and obtain the angle  $\beta$  to any desired accuracy. Since the cloning machine can be thought of as part of the measurement apparatus, this would contradict the measurement postulate. This postulate implies that from the measurement of the polarization state of a photon we can obtain only a single bit of information. We obtain 0 with probability  $p_0 = \cos^2 \beta$  and 1 with probability  $p_1 = \sin^2 \beta$ . In contrast, if a cloning machine existed, from this measurement we could determine, to any desired accuracy, the parameter  $\beta$ . Therefore, from the simple measurement of the polarization state of a single photon one could extract an arbitrarily large amount of information (the bits necessary to represent  $\beta$  to the desired accuracy). We can conclude, on the basis of the measurement postulate of quantum mechanics, that a quantum cloning machine cannot exist.

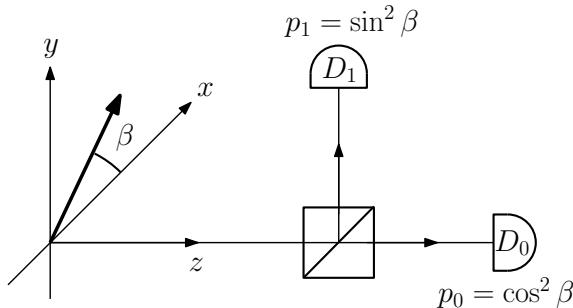


Fig. 5.1 Measurement of the polarization of a single photon. The two photodetectors are denoted by  $D_0$  and  $D_1$ .

We now give a more formal proof of the no-cloning theorem. This proof has to be considered weaker than the previous, since it takes advantage of the linearity of quantum mechanics.

**Theorem 5.1** *It is impossible to build a machine that operates unitary transformations and is able to clone the generic state of a qubit.*

**Proof.** Let us consider a system composed of the qubit to be cloned, a second qubit and the cloning machine. The first qubit is prepared in a generic state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (5.8)$$

with  $\alpha$  and  $\beta$  complex numbers, constrained by the normalization condition  $|\alpha|^2 + |\beta|^2 = 1$ . Initially, the second qubit and the cloning machine are prepared in some reference states, for instance  $|\phi\rangle$  and  $|A_i\rangle$ , respectively. The cloning machine should be able to perform a unitary transformation  $U$  such that

$$U(|\psi\rangle|\phi\rangle|A_i\rangle) = |\psi\rangle|\psi\rangle|A_{f\psi}\rangle = (\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle)|A_{f\psi}\rangle, \quad (5.9)$$

where the final state of the machine will in general depend on the state  $|\psi\rangle$  to be cloned. We show now that such a unitary transformation cannot exist. If the first qubit is in the state  $|0\rangle$ , the action of the cloning machine must be

$$U(|0\rangle|\phi\rangle|A_i\rangle) = |0\rangle|0\rangle|A_{f0}\rangle. \quad (5.10a)$$

Analogously, if the first qubit is prepared in the state  $|1\rangle$ ,

$$U(|1\rangle|\phi\rangle|A_i\rangle) = |1\rangle|1\rangle|A_{f1}\rangle. \quad (5.10b)$$

Therefore, the action of the cloning machine on a generic state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is given by

$$U((\alpha|0\rangle + \beta|1\rangle)|\phi\rangle|A_i\rangle) = \alpha U(|0\rangle|\phi\rangle|A_i\rangle) + \beta U(|1\rangle|\phi\rangle|A_i\rangle), \quad (5.11)$$

where we have invoked the *linearity* of quantum mechanics. We now insert Eqs. (5.10a) and (5.10b) into Eq. (5.11), obtaining the state

$$\alpha|0\rangle|0\rangle|A_{f0}\rangle + \beta|1\rangle|1\rangle|A_{f1}\rangle, \quad (5.12)$$

which is clearly different from the desired cloned state of Eq. (5.9).  $\square$

We stress that it is essential to consider a *generic* state. Indeed, if we know from the beginning that the quantum state of the qubit is prepared in one out of two orthogonal states, for instance  $|0\rangle$  or  $|1\rangle$ , then we can measure with certainty the state of the qubit and prepare as many copies of it as desired. In this case the qubit acts as a classical bit and we know that there exist classical cloning machines (photocopiers *etc.*).

**Exercise 5.2** A single qubit is in an unknown state  $|\psi_1\rangle$ . We guess at random that its state is  $|\psi_2\rangle$ . What is the average fidelity  $f$  of our guess, defined by  $f \equiv |\langle\psi_1|\psi_2\rangle|^2$ ?

### 5.2.1 Faster-than-light transmission of information?

It is interesting to show that the existence of a cloner would violate a basic principle of the theory of relativity, that is, information could be transmitted faster than light. Assume that a source produces EPR pairs in the Bell state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle). \quad (5.13a)$$

This state can also be written as

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_x|0\rangle_x + |1\rangle_x|1\rangle_x), \quad (5.13b)$$

where  $|0\rangle_x$  and  $|1\rangle_x$  are the eigenstates of the Pauli matrix  $\sigma_x$  with eigenvalues  $+1$  and  $-1$ , respectively. It is easy to check the equality of expressions (5.13a) and (5.13b), taking into account that

$$\begin{cases} |0\rangle_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ |1\rangle_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{cases} \quad (5.14)$$

A member of each EPR pair is sent to Alice, the other to Bob. Note that, in principle, Alice and Bob can be located arbitrarily far apart. Alice codes a message that she wants to send to Bob in a binary string. She then performs a measurement

on her member of each EPR pair and she chooses to measure  $\sigma_x$  or  $\sigma_z$  depending on the fact that the corresponding digit is 0 or 1 (we assume that Alice and Bob share at least as many EPR pairs as digits in their message). After this, the state of Bob's member of the EPR pair collapses onto an eigenstate of  $\sigma_x$  or  $\sigma_z$ . However, these states are not orthogonal and Bob cannot obtain any information about Alice's message from his measurements. In contrast, if a cloning machine existed, Bob could make an arbitrarily large number of copies of his EPR qubits and distinguish between eigenstates of  $\sigma_x$  and  $\sigma_z$  with any desired accuracy. Indeed, given the generic state of a qubit,  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , and a large number of copies of this qubit, Bob could estimate  $\alpha$  and  $\beta$  from measuring these qubits. Hence, in particular, he could determine whether this state were an eigenstate of  $\sigma_x$  or of  $\sigma_z$ . Therefore, superluminal transmission of information would be possible, in contradiction with a basic principle of the theory of relativity.

**Exercise 5.3** Show that, in the above example, independently of the decision of Alice to measure  $\sigma_x$  or  $\sigma_z$ , Bob obtains an up or down polarization state with equal probability, whatever direction he chooses for the measurement of the spin polarization.

It is interesting to note that within the framework of special relativity a faster than light transmission of information would violate the *causality principle*, stating that every “effect” follows a “cause” and that this principle must be valid in any admissible reference frame. In Newtonian mechanics, as time is absolute, if the causality principle is valid in a reference frame, then it is valid in every reference frame. In special relativity it is easy to show that if in a given inertial reference frame  $\Sigma$  there is an action propagating at speed  $V$  greater than the speed  $c$  of light in free space, then it exists another reference frame  $\Sigma'$ , moving with respect to  $\Sigma$  at speed  $\omega < c$ , in which the action propagates in the past.

The postulates of special relativity read as follows: (i) the geometry is Euclidean, (ii) the inertia principle is valid, (iii) the laws of physics are the same in all inertial frames of reference, (iv) the speed of light in free space is invariant for all inertial frames of reference. From these postulates it is possible to deduce the Lorentz transformations for two inertial reference frames  $\Sigma$  (Cartesian coordinates  $(x, y, z)$  and time  $t$ ) and  $\Sigma'$  (Cartesian coordinates  $(x', y', z')$  and time  $t'$ ). Let us consider, for instance, an observer in  $\Sigma$  measuring  $\Sigma'$  to move at velocity  $\omega$  along the coincident axes  $x$  and  $x'$  (with  $y$  parallel to  $y'$ ,  $z$  to  $z'$  and  $\Sigma \equiv \Sigma'$  at  $t = t' = 0$ ). We have the Lorentz transformation

$$\begin{cases} x' = \gamma(x - \omega t), \\ y' = y, \\ z' = z, \\ t' = \gamma(t - \omega x/c^2), \end{cases} \quad (5.15)$$

where  $\gamma = 1/\sqrt{1 - \beta^2}$ , with  $\beta = \omega/c$ .

The first consequence of Lorentz transformations is the *relativity of simultaneity*: it is not absolute whether two spatially separated events occur at the same time, but it depends on the reference frame. Two simultaneous events in  $\Sigma$ , of space-time coordinates  $A = (x = 0, t = 0)$  and  $B = (X, 0)$  (coordinates  $y$  and  $z$  are not relevant here) have in the reference frame  $\Sigma'$  coordinates  $A' = (0, 0)$  and  $B' = (\gamma X, -\gamma \omega X/c^2)$ . Note that, while  $\Delta t = t_B - t_A = 0$ ,  $\Delta t' = t'_{B'} - t'_{A'} = -\gamma \omega X/c^2$  is different from zero, provided  $X \neq 0$  (spatially separated events) and  $\omega \neq 0$  (relative motion of the two inertial frames).

Let us consider two events with coordinates  $A = (0, 0)$  and  $B = (X, T)$ , where  $T > 0$ , so that in  $\Sigma$  the event  $B$  occurs after  $A$ . The speed of a signal connecting the two events is (in reference frame  $\Sigma$ )  $V = X/T$ . The coordinates of the two events in  $\Sigma'$  are  $A' = (0, 0)$  and  $B' = (\gamma(X - \omega T), \gamma(T - \omega X/c^2))$ . In  $\Sigma'$  the temporal distance between the two events is  $t'_{B'} - t'_{A'} = \gamma(T - \omega X/c^2) = \gamma T(1 - \omega V/c^2)$ . Therefore in  $\Sigma'$  the event in  $B'$  is prior to the event in  $A'$  provided that  $1 - \omega V/c^2 < 0$ , namely  $V > c^2/\omega$ . Hence, if  $V > c$  it is possible to find a frame of reference with  $\omega < c$  in which the signal is propagating in the past, in contradiction with the causality principle.

### 5.2.2 \* The no-signalling condition

The no-signalling condition, namely the fact that one cannot exploit quantum entanglement and instantaneous wave-packet reduction for faster than light communication, is a general property of quantum mechanics.

Let  $\rho$  denote the density matrix describing the state of a composite system and  $\rho_1 = \text{Tr}_2 \rho$  the reduced density matrix describing subsystem 1, obtained after partial tracing of  $\rho$  over subsystem 2 (see Sec. 2.6.1). All probabilistic predictions concerning the possible outcomes of measurements performed on subsystem 1 can be described in terms of the reduced density matrix  $\rho_1$  alone. In what follows, we will show that  $\rho_1$  is unchanged by local operations acting on subsystem 2, and therefore no measurable information can be instantaneously transferred from subsystem 2 to subsystem 1 as a consequence of such operations. We can distinguish three instances:

- (i) *Local unitary transformation  $U_2$ .* We have

$$\text{Tr}_2 [(I_1 \otimes U_2) \rho (I_1 \otimes U_2)^\dagger] = \text{Tr}_2 [(I_1 \otimes U_2)^\dagger (I_1 \otimes U_2) \rho] = \rho_1, \quad (5.16)$$

where  $I_1$  denotes the identity operator for subsystem 1 (evolving trivially) and we used the cyclic property of the trace (see Sec. A.1).

- (ii) *Projective measurements,* described by the projectors  $P_{2,n}$  acting on the Hilbert space for the second subsystem. Such operators fulfill the completeness relation  $\sum_n P_{2,n} = I_2$ , with  $I_2$  the identity operator for subsystem 2. In this case

$$\text{Tr}_2 [\sum_n (I_1 \otimes P_{2,n}) \rho (I_1 \otimes P_{2,n})^\dagger] = \text{Tr}_2 [(I_1 \otimes \sum_n P_{2,n}^2) \rho] = \rho_1, \quad (5.17)$$

where we used for projectors the properties  $P_{2,n}^\dagger = P_{2,n}$  and  $P_{2,n}^2 = P_{2,n}$  (see Sec. A.1).

- (iii) *Quantum dynamical maps* representing more general evolutions for subsystem 2. Such maps can be described in terms of Kraus operators (see Sec. 7.1)  $E_{2,n}$ , with  $\sum_n E_{2,n}^\dagger E_{2,n} = I_2$ . We obtain

$$\text{Tr}_2 [\sum_n (I_1 \otimes E_{2,n}) \rho (I_1 \otimes E_{2,n})^\dagger] = \text{Tr}_2 [(I_1 \otimes \sum_n E_{2,n}^\dagger E_{2,n}) \rho] = \rho_1. \quad (5.18)$$

It might be interesting to note that the no-signalling condition has deep connections with the two-atom Fermi problem (Fermi, 1932). Suppose that there are two atoms, A and B, separated by a distance  $R$  and interacting with the electromagnetic field. Atom A

is in an excited state, while atom A is in its ground state. Assume that at time  $t = 0$  atom A emits a photon. Can such emission affect the excitation probability of atom B at times  $t < R/c$ , where  $c$  is the speed of light? Fermi's answer was negative, but his solution was based on approximations. The problem is still debated, see for instance Hegerfeldt (1994), Sabín *et al.* (2011), Zohar and Reznik (2011), Borrelli *et al.* (2012), and references therein.

### 5.2.3 \* Universal quantum cloning

The no-cloning theorem does not forbid the existence of a quantum cloning machine that *approximately* copies quantum mechanical states. Many schemes have been proposed that optimize some measure of the fidelity between the *imperfect* copies and the original state. Bužek and Hillery devised a machine that produces two identical copies of the original qubit, the quality of the copies being independent of the input state. It can be shown that the quantum cloning machine of Bužek and Hillery is optimal, in the sense that it maximizes the average fidelity between the input and output states (see Gisin *et al.*, 2002 and Bruß *et al.*, 1998). The fidelity is a measure of the quality of the copy and is defined by

$$f = \langle \psi | \rho | \psi \rangle, \quad (5.19)$$

where  $|\psi\rangle$  is the state to be copied and  $\rho$  is the density matrix describing the copy. We have  $0 \leq f \leq 1$  and the maximum value  $f = 1$  is taken when  $\rho = |\psi\rangle\langle\psi|$ . We note that Eq. (5.19) generalizes the definition of fidelity of two pure states  $|\psi\rangle, |\phi\rangle$ , given by  $f = |\langle\psi|\phi\rangle|^2$  (see exercise 3.4). If  $\rho = \sum_k p_k |\phi_k\rangle\langle\phi_k|$ , then Eq. (5.19) gives  $f = \sum_k p_k F_k$ , where  $F_k = |\langle\psi|\phi_k\rangle|^2$ . Therefore,  $f$  is the weighted sum of the pure-state fidelities  $F_k$ .

Let us describe the working of Bužek–Hillery cloning machine. Given a qubit in a generic unknown pure state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (5.20)$$

we consider the copying network shown in Fig. 5.2. This circuit can be decomposed into two parts: (i) the preparation of a specific state of the quantum copier and (ii) the copying process. It can be seen from Fig. 5.2 that only part (ii) depends on the state  $|\psi\rangle$  to be copied. Let us first look at the preparation stage. The gates labelled by  $\theta_i$  denote the application of the rotation matrices

$$R_y(-2\theta_i) = \begin{bmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{bmatrix}, \quad (5.21)$$

where, as discussed in Sec. 3.4.1,  $R_y(-2\theta_i)$  corresponds to a counterclockwise rotation through an angle  $-2\theta_i$  about the  $y$ -axis of the Bloch sphere and the angles  $\theta_i$  are chosen as

$$\cos 2\theta_1 = \frac{1}{\sqrt{5}}, \quad \cos 2\theta_2 = \frac{\sqrt{5}}{3}, \quad \cos 2\theta_3 = \frac{2}{\sqrt{5}}. \quad (5.22)$$

At the end of the preparation stage, the two qubits initially in the state  $|00\rangle$  are transformed into the state vector

$$|\Phi\rangle = \frac{1}{\sqrt{6}} (2|00\rangle + |01\rangle + |11\rangle). \quad (5.23)$$

It can be checked that the copying part of the circuit in Fig. 5.2 transforms the state  $|\psi\rangle|\Phi\rangle$  into

$$|A_0\rangle|0\rangle + |A_1\rangle|1\rangle, \quad (5.24)$$

with

$$\begin{aligned} |A_0\rangle &= \alpha \sqrt{\frac{2}{3}} |00\rangle + \beta \sqrt{\frac{1}{6}} (|10\rangle + |01\rangle), \\ |A_1\rangle &= \beta \sqrt{\frac{2}{3}} |11\rangle + \alpha \sqrt{\frac{1}{6}} (|10\rangle + |01\rangle). \end{aligned} \quad (5.25)$$

Since the three qubits are now entangled, we must trace over two of them to obtain the (mixed) state describing the third. Let us first trace over the bottom qubit of Fig. 5.2, obtaining the density matrix

$$|A_0\rangle\langle A_0| + |A_1\rangle\langle A_1|. \quad (5.26)$$

Then, by further tracing over one of the first two qubits, it is possible to check that each of the two qubits at the output of the quantum copier (the two top qubits in Fig. 5.2) is described by the same reduced density operator

$$\rho = \frac{2}{3} |\psi\rangle\langle\psi| + \frac{1}{6} I. \quad (5.27)$$

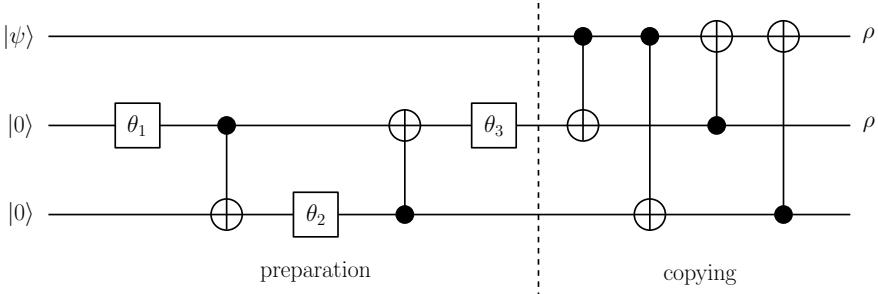


Fig. 5.2 A quantum copying network: two imperfect copies of the state  $|\psi\rangle$ , described by the density matrix  $\rho$ , are recovered at the output. The  $\theta_i$ -symbols stand for the rotation matrices  $R_y(-2\theta_i)$  defined by Eq. (5.21). Note that, to simplify notation, on the left-hand side of the circuit we show the state vectors instead of the corresponding density matrices  $|\psi\rangle\langle\psi|$  and  $|0\rangle\langle 0|$ .

It is easy to check that, independently of the initial state  $|\psi\rangle$ , the (optimal) fidelity of the copy  $\rho$  is given by

$$f = \langle\psi|\rho|\psi\rangle = \langle\psi|(\frac{2}{3}|\psi\rangle\langle\psi| + \frac{1}{6}I)|\psi\rangle = \frac{5}{6}. \quad (5.28)$$

**Exercise 5.4** Check that the quantum circuit in Fig. 5.2 produces two copies described by the density matrix (5.27).

### 5.2.4 \* The universal-NOT gate

As for quantum cloning, the universal-NOT (U-NOT) gate represents a transformation that cannot be carried out exactly for an unknown quantum state. The classical NOT gate changes the value of a bit:  $0 \rightarrow 1$  and  $1 \rightarrow 0$ . The quantum NOT gate turns the states of the computational basis into states that are orthogonal to them:

$$\text{NOT } |0\rangle = e^{i\delta}|1\rangle, \quad \text{NOT } |1\rangle = |0\rangle. \quad (5.29)$$

Since global phases have no physical meaning,  $e^{i\delta}|1\rangle \equiv |1\rangle$ ; however, for a generic input state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , we have  $\text{NOT } |\psi\rangle = \alpha e^{i\delta}|1\rangle + \beta|0\rangle$  and therefore the relative phase  $\delta$  is observable. A universal-NOT gate is required to transform a generic input state  $|\psi\rangle$  into the state orthogonal to it:  $|\psi^\perp\rangle = \beta^*|0\rangle - \alpha^*|1\rangle$ . In the Bloch sphere representation,  $|\psi\rangle$  and  $|\psi^\perp\rangle$  are antipodes of each other. Hence the U-NOT gate corresponds to the inversion of the Bloch sphere. No unitary transformation can perform this operation. While the NOT gate turns the states of the computational basis into their orthogonal states, for a generic state we have  $\text{NOT } |\psi\rangle \neq |\psi^\perp\rangle$ . The problem resides in the complex conjugates which appear in the orthogonal state. If  $\alpha, \beta$  are real, then we have  $\text{NOT } |\psi\rangle = |\psi^\perp\rangle$ , provided  $\delta = \pi$ . In general, the transformation U-NOT :  $|\psi\rangle \rightarrow |\psi^\perp\rangle$  is *antiunitary*, namely it is an *antilinear* map (i.e, for all  $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{V}$  and all  $\alpha_1, \alpha_2 \in \mathbb{C}$ ,  $\text{U-NOT}(\alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle) = \alpha_1^*\text{U-NOT } |\psi_1\rangle + \alpha_2^*\text{U-NOT } |\psi_2\rangle$ ) and moreover  $\langle \text{U-NOT } \psi_1 | \text{U-NOT } \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle^*$ .

Since only unitary transformations are allowed in quantum mechanics, the U-NOT gate cannot be carried out perfectly. However, similarly to the case of cloning it is possible to approximate the forbidden U-NOT gate, generating an optimal output state  $\rho_{\text{out}}$ . That is, given a generic input state  $|\psi\rangle$  we maximize the fidelity  $f = \langle \psi^\perp | \rho_{\text{out}} | \psi^\perp \rangle$  between the obtainable state  $\rho_{\text{out}}$  and the desired state  $|\psi^\perp\rangle$ . It can be shown (see Bužek *et al.*, 1999) that the optimal fidelity is  $f = \frac{2}{3}$ , this value being independent of the input state  $|\psi\rangle$ .

We first consider a measurement-based scenario. The input state  $|\psi\rangle$  is measured along the direction of a randomly chosen vector

$$|\eta\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle. \quad (5.30)$$

The polarization measurement is positive with probability  $|\langle \eta | \psi \rangle|^2$ , and in this case the output is taken to be  $|\eta^\perp\rangle$ . In contrast, the polarization measurement is negative with probability  $1 - |\langle \eta | \psi \rangle|^2 = |\langle \eta^\perp | \psi \rangle|^2$  and in this case the output is  $|\eta\rangle$ . The output state (for a selected direction  $|\eta\rangle$ ) is therefore

$$\rho_{\text{out}}(\eta) = |\langle \eta | \psi \rangle|^2 |\eta^\perp\rangle \langle \eta^\perp| + |\langle \eta^\perp | \psi \rangle|^2 |\eta\rangle \langle \eta|. \quad (5.31)$$

After averaging over all possible directions we obtain the final output state

$$\rho_{\text{out}} = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \rho_{\text{out}}(\eta) = \frac{1}{3} |\psi^\perp\rangle \langle \psi^\perp| + \frac{1}{3} I. \quad (5.32)$$

It is easy to check that the fidelity  $f = \langle \psi^\perp | \rho_{\text{out}} | \psi^\perp \rangle = \frac{2}{3}$ , independently of the input state  $|\psi\rangle$ .

We now consider a unitary transformation on a larger system, including the qubit prepared in a generic input state  $|\psi\rangle$  and ancillary qubits. The quantum circuit is essentially the same as for the Bužek–Hillery cloning machine, see Fig. 5.2. While the two top qubits end up as clones, if we apply the NOT gate defined in Eq. (5.29) to the third qubit, with the phase  $\delta = \pi$ , then we produce for that qubit the output state  $\rho_{\text{out}} = \frac{1}{3} |\psi^\perp\rangle\langle\psi^\perp| + \frac{1}{3} I$ , whose fidelity with the state  $|\psi^\perp\rangle$  is  $f = \frac{2}{3}$ . Hence the quantum circuit of Fig. 5.2 acts as both an optimal cloner and an optimal universal-NOT gate.

### 5.3 Quantum cryptography

In classical physics it is impossible to know with certainty if the eavesdropper Eve is monitoring a message. The reason is that classical information can be copied without changing the original message. Indeed, information must be encoded in some physical system (a piece of paper, radio signals *etc.*), whose properties can, in principle, be measured *passively*. The modifications induced in the system can be made as small as permitted by the available technology. In contrast, in quantum mechanics the measurement process in general disturbs the system for fundamental reasons. This is a consequence of the Heisenberg uncertainty principle (see Sec. 2.3). Indeed, if one considers a pair of non-commuting observables, the measurement of one observable necessarily disturbs (randomizes) the other. In this section, we shall see that this inherently quantum property allows *intrusion detection*: Alice and Bob can discover if Eve is eavesdropping their communication. This possibility can be used to create a secret key between two parties. The resulting key allows Alice and Bob to communicate secretly by means of classical cryptosystems like the Vernam cypher. In the following we describe two protocols for quantum-key distribution, the BB84 protocol and the E91 protocol.

#### 5.3.1 The BB84 protocol

The protocol BB84, discovered by Bennett and Brassard in 1984, requires four states and two binary alphabets:  $|0\rangle$  and  $|1\rangle$  (the  $z$ -alphabet),  $|+\rangle \equiv |0\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ,  $|-\rangle \equiv |1\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  (the  $x$ -alphabet). The letters of the  $z$ - and  $x$ -alphabets are associated with the eigenstates of the Pauli matrices  $\sigma_z$  and  $\sigma_x$ , respectively. The description of the BB84 protocol follows (and a simple example is given in Table 5.1):

- (1) Alice generates a random sequence of 0's and 1's.
- (2) Alice encodes each data bit in a qubit,  $|0\rangle$  or  $|+\rangle = |0\rangle_x$  if the corresponding bit is 0,  $|1\rangle$  or  $|-\rangle = |1\rangle_x$  if the corresponding bit is 1. For each bit, Alice chooses randomly between the  $x$ - and the  $z$ -alphabet, by means of a fair coin

(e.g., if the coin lands heads Alice chooses the  $x$ -alphabet, while the  $z$ -alphabet is chosen when the coin lands tails).

- (3) The resulting string of qubits is sent by Alice and received by Bob.
- (4) For each qubit, Bob decides at random what axis (alphabet) to use for the measurement,  $x$  or  $z$ . In the first case, he measures the spin polarization along the  $x$  axis, in the latter along the  $z$  axis. Note that half of the time Bob chooses the same axis as Alice. In this case, assuming that there are no eavesdroppers or noise effects, Alice and Bob share the same bit (here we summarize under the word noise effects such as imperfect state preparation or detection, interactions of the transmitted qubit with the environment, etc.). In contrast, if Bob chooses an axis different from Alice, the bit resulting from his measurement agrees with the bit sent by Alice only half of the time. For instance, if Bob receives the qubit  $|-\rangle$  and measures  $\sigma_z$ , the outcomes 0 and 1 have equal probability.

From now on Alice and Bob exchange only classical information over a public channel.

- (5) Bob communicates to Alice over a classical public channel which alphabet he used for each qubit measurement. Of course, he does not communicate the results of these measurements.
- (6) Alice communicates to Bob over a classical public channel which alphabet she used for each transmitted qubit (again, not the results of these measurements).
- (7) Alice and Bob delete all bits corresponding to the cases in which they used different alphabets. After this they share the so-called *raw key*. This key is the same for Alice and Bob, insofar as Eve and noise were absent.

By means of the following steps, Alice and Bob distill the secret key starting from the raw key.

- (8) Over a public communication channel, Alice and Bob announce and compare a part of their raw key. From this comparison they can estimate the *error rate*  $R$  due to eavesdroppers or noise effects. If this rate is too high they restart the protocol from the beginning. If not, they perform *information reconciliation* and *privacy amplification* on the remaining bits of their raw key.
- (9) Information reconciliation is just classical error correction over a public transmission channel. We shall describe classical error-correcting codes in Chap. 9. Here, we limit ourselves to the illustration of a simple scheme for information reconciliation. Alice and Bob divide the remaining bits of their raw key into subsets of length  $l$ . This length is chosen in order that it is unlikely to have more than one error per subset ( $Rl \ll 1$ , with  $R$  the previously estimated error rate). For each subset, Alice and Bob make parity checks (the parity  $P$  of a binary string  $\{b_1, b_2, \dots, b_l\}$  is defined as  $P = b_1 \oplus b_2 \oplus \dots \oplus b_l$ ), discarding each time the last bit. If the parities of a given subset are different for Alice and Bob, then they locate and delete the erroneous bit by binary search in the following way. They bisect the subset and check the parities of the new blocks

( $P_1 = b_1 \oplus b_2 \oplus \dots \oplus b_{(l-1)/2}$  and  $P_2 = b_{(l-1)/2+1} \oplus b_{(l-1)/2+2} \oplus \dots \oplus b_{l-1}$ ). They repeat the bisection for the block in which Alice's and Bob's parities differ and so on. Note that each time Alice and Bob delete the last bit of the blocks whose parity is publicly announced. In this way, they avoid Eve obtaining any amount of information from their parity checks. At the end, with high probability, Alice and Bob share the same string of bits.

- (10) Privacy amplification reduces Eve's information about the final secret key to arbitrarily small values. Let us illustrate a simple privacy amplification protocol. Alice and Bob estimate from the error rate  $R$  obtained previously the maximum number of bits  $k$  known by Eve. Let  $s$  be a security parameter. Then Alice and Bob choose at random  $n - k - s$  subsets of their key, where  $n$  denotes the number of bits in the key. The parities of these subsets become the final secret key. This key is more secure than the previous one, since Eve must know something about each bit of a subset in order to obtain information about its parity. It can be shown that Eve's residual information is  $O(2^{-s})$ .

Note that Eve can choose different eavesdropping strategies:

- (i) *intercept and resend*: Eve intercepts and measures the qubits sent by Alice and resends them to Bob;
- (ii) *translucent attack*: Eve has probes (ancillary qubits) interacting with the qubits sent by Alice and she measures the state of these probes;
- (iii) *collective attack*: Eve manipulates not a single qubit at a time but a block of qubits.

Independently of the eavesdropping strategy, it is possible to show that quantum-key distribution is secure, in the sense that it is possible to guarantee that Eve's information about the final key is arbitrarily small (see Nielsen and Chuang, 2000).

Table 5.1 An example of the BB84 protocol.

Alice's data bits	1	0	0	0	1	1	0	1	0	1
Alice's alphabet	$x$	$z$	$x$	$z$	$x$	$x$	$x$	$z$	$z$	$x$
Transmitted qubits	$ 1\rangle_x$	$ 0\rangle$	$ 0\rangle_x$	$ 0\rangle$	$ 1\rangle_x$	$ 1\rangle_x$	$ 0\rangle_x$	$ 1\rangle$	$ 0\rangle$	$ 1\rangle_x$
Bob's alphabet	$x$	$z$	$x$	$x$	$z$	$x$	$z$	$x$	$z$	$z$
Measurement outcomes	1	0	0	0	0	1	0	0	0	1
Bob's data bits	1	0	0	0	0	1	0	0	0	1
Raw key	1	0	0			1			0	

We stress that the validity of the BB84 protocol is based on the Heisenberg principle. The two alphabets are associated with two non-commuting observables,  $\sigma_x$  and  $\sigma_z$ . Eve cannot measure both the polarization along  $x$  and along  $z$  for the same qubit. For instance, if she measures  $\sigma_z$  for the qubit  $|0\rangle_x$ , she obtains

the outcomes 0 or 1 with equal probability. Thus, she has irreversibly randomized the polarization originally sent by Alice. We also stress the importance of the no-cloning theorem: it guarantees that Eve cannot distinguish with certainty between non-orthogonal quantum states. If a quantum cloning machine existed, Eve could make a large number of copies of each qubit sent by Alice and distinguish with arbitrary accuracy between eigenstates of  $\sigma_x$  and  $\sigma_z$ . For instance, assume that Eve measures  $\sigma_z$  for the qubit and all its copies. If she received  $|1\rangle$ , she always obtains outcome 1. On the other, if she received  $|1\rangle_x$ , she obtains outcomes 0 and 1 with equal probabilities. Finally, Eve could resend a copy of the intercepted qubit to Bob. Therefore, if it were possible to violate the no-cloning theorem, Eve could intercept the qubits sent by Alice and resend them to Bob, leaving no trace of her intrusion.

**Exercise 5.5** Assume that Eve intercepts every qubit sent by Alice, measures its polarization along some axis and resends it to Bob. What is the error rate introduced in the raw key?

**Exercise 5.6** Show that it is not possible to gain information about which one of two non-orthogonal quantum states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  was sent by Alice without disturbing the state.

Finally, we note that one of the main drawbacks of quantum cryptography is that no mechanism is known for authentication. Thus, a classical secret key is required for this purpose. Indeed, in order to be sure that they are not communicating with someone else, Alice and Bob need to send an authentication key over a classical secure channel. After this they can implement a quantum protocol like BB84 and “expand” the existing authentication key.

### 5.3.2 The E91 protocol

Now we discuss the protocol E91 (Ekert, 1991), a quantum cryptosystem that uses entangled EPR pairs.

- (1) A source  $S$  emits a pair of qubits (spin  $\frac{1}{2}$  particles) in the EPR state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (5.33)$$

The qubits are sent in opposite directions; the first is received by Alice, the second by Bob (see Fig. 5.3). Note that a third party is not strictly necessary: Alice could produce EPR pairs and then send a member of each pair to Bob.

- (2) Alice and Bob can discover if Eve has intercepted the transmission of the EPR pairs by exploiting the quantum correlations of EPR pairs. They measure the spin polarization along one of the three directions  $\hat{a}_1, \hat{a}_2, \hat{a}_3$  (Alice) and  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  (Bob) (see Fig. 5.4). For each EPR pair, Alice and Bob choose at random their measurement axes between  $\hat{a}_1, \hat{a}_2, \hat{a}_3$  and  $\hat{b}_1, \hat{b}_2, \hat{b}_3$ , respectively. Let us denote  $p_{\pm\pm}(\hat{a}_i, \hat{b}_j)$  the probability that Alice’s polarization measurement along

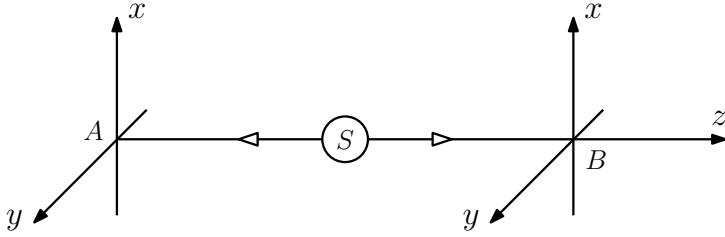


Fig. 5.3 A schematic picture of the E91 protocol. The EPR source, Alice, and Bob are denoted by  $S$ ,  $A$  and  $B$ .

the direction  $\hat{a}_i$  gives the result  $\pm 1$  and Bob's measurement along  $\hat{b}_i$  gives  $\pm 1$ . We define the correlation coefficients:

$$E(\hat{a}_i, \hat{b}_j) = p_{++}(\hat{a}_i, \hat{b}_j) + p_{--}(\hat{a}_i, \hat{b}_j) - p_{+-}(\hat{a}_i, \hat{b}_j) - p_{-+}(\hat{a}_i, \hat{b}_j). \quad (5.34)$$

From the discussion of Sec. 2.5 on Bell's inequalities, we know that

$$C \equiv E(\hat{a}_1, \hat{b}_1) - E(\hat{a}_1, \hat{b}_3) + E(\hat{a}_3, \hat{b}_1) + E(\hat{a}_3, \hat{b}_3) = -2\sqrt{2}, \quad (5.35)$$

that is, quantum mechanics violates the CHSH inequality, which reads  $|C| \leq 2$  (see Sec. 2.5).

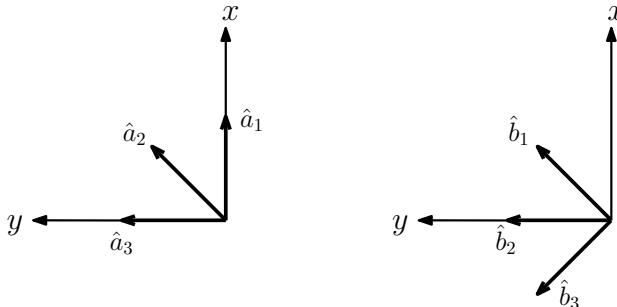


Fig. 5.4 Directions of the measurement axes for Alice (left) and Bob (right). The angles between these directions and the  $x$  axis are 0,  $\frac{\pi}{4}$ ,  $\frac{\pi}{2}$  for  $\hat{a}_1$ ,  $\hat{a}_2$ ,  $\hat{a}_3$  and  $\frac{\pi}{4}$ ,  $\frac{\pi}{2}$ ,  $\frac{3\pi}{4}$  for  $\hat{b}_1$ ,  $\hat{b}_2$ ,  $\hat{b}_3$ .

- (3) Alice and Bob announce over a public channel the axis chosen for each measurement. Then they make public the outcomes of the measurements in the cases in which their polarization axes did not coincide. This allows Alice and Bob to check the equality (5.35). If  $C > -2\sqrt{2}$ , then Eve attacked the EPR pairs or there were noise effects (note that it is possible to show that  $|C|$  cannot be larger than  $2\sqrt{2}$ ; for a proof of this result see, e.g., Preskill, 1998a). In the absence of such effects, that is,  $C = -2\sqrt{2}$ , Alice and Bob's measurements along the same axis are perfectly anticorrelated, namely,

$$E(\hat{a}_2, \hat{b}_1) = E(\hat{a}_3, \hat{b}_2) = -1. \quad (5.36)$$

The outcomes of these measurements are the raw key shared by Alice and Bob (the keys agree if Bob negates his outcomes,  $0 \rightarrow 1$  and  $1 \rightarrow 0$ ). After this

Alice and Bob can perform key reconciliation and privacy amplification as in the BB84 protocol.

We note that it is not necessary to test the relation (5.35). Alice and Bob could simply perform measurements along  $x$  or  $z$ . The decision is random, each choice occurring with probability  $\frac{1}{2}$ . After the measurement process, Alice and Bob announce over a public channel which observable they measured for each EPR pair. In the cases in which their measurement axes agree, the outcomes are perfectly anticorrelated. Alice and Bob discard the other bits, thus remaining with a shared raw key. After this, they proceed as in the BB84 protocol. It is interesting to note that the secret key is not generated by either Alice or Bob. The key is undetermined until Alice and Bob measure their halves of the shared EPR pairs. Then the secret key arises from a fundamentally random process, the quantum measurement. Finally, we stress that the E91 protocol is potentially interesting for *key storage*. The problem is the following: once the secret key has been established, Alice and Bob must store it in their safes, until they need it. However, the key is a string of *classical bits* and, in principle, can be copied. It may be very difficult to crack the safe, but always possible. No fundamental reasons exclude this possibility. However, if Alice and Bob were able to store EPR pairs, they could wait to establish the secret key until needed. Of course, the implementation of such key storage is hampered by the fact that one should be able to protect the EPR pairs from noise effects, due to interactions with their environment, for long times. This possibility is beyond the reach of present technology.

#### 5.4 Dense coding

Dense coding is the simplest example of the application of quantum entanglement to communication. It allows Alice to send two bits of classical information to Bob by sending him only a single qubit. The dense coding protocol works as follows (see the schematic picture in Fig. 5.5 and the quantum circuit implementing the protocol in Fig. 5.6):

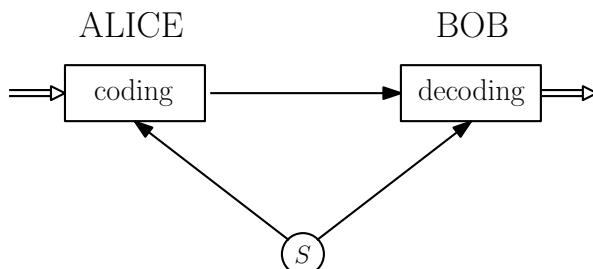


Fig. 5.5 A schematic picture of the dense coding protocol. The double lines denote two classical bits and the single lines a quantum bit.

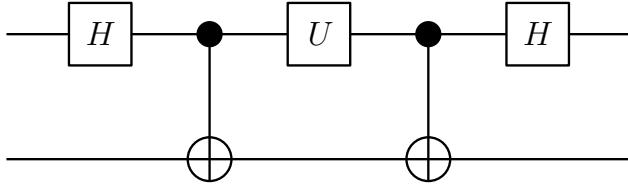


Fig. 5.6 A quantum circuit implementing the dense coding protocol.

- (1) A source  $S$  generates an EPR pair shared by Alice and Bob. The EPR pair is prepared, for instance, in the state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (5.37)$$

The Bell state  $|\phi^+\rangle$  is obtained from the state  $|00\rangle$  after application of a Hadamard gate and a CNOT gate:

$$\text{CNOT} (H \otimes I) |00\rangle = |\phi^+\rangle. \quad (5.38)$$

In the computational basis with basis vectors  $|00\rangle$ ,  $|01\rangle$ ,  $|10\rangle$  and  $|11\rangle$  (note that the first label refers to Alice's half of the EPR pair, the second to Bob's half), the transformation (5.38) has the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (5.39)$$

Note that, as shown in Fig. 5.5, a source  $S$  creates the EPR pair and then sends one member of the pair to Alice and the other to Bob. It is important to stress that Alice and Bob can be located arbitrarily far apart.

- (2) There are four possible values of the two classical bits that Alice wishes to send to Bob: 00, 01, 10 and 11. They determine the unitary operation  $U$  that Alice performs on her half of the EPR pair:  $U = I$ ,  $\sigma_x$ ,  $\sigma_z$  or  $i\sigma_y$  (we remind the reader that  $I$  denotes the identity and  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$  the Pauli matrices). The reason for choosing one of these four unitary transformations will become clear in the following.

As we have said, the operator  $U$  in the circuit of Fig. 5.6 is determined by the value of the two classical bits that Alice wishes to communicate to Bob. If she wishes to communicate 00, she operates the identity on her half of the EPR pair. This gives the trivial transformation

$$I \otimes I |\phi^+\rangle = |\phi^+\rangle. \quad (5.40)$$

If she wishes to communicate 01, she operates the Pauli matrix  $\sigma_x$  on her half of the EPR pair, obtaining

$$\sigma_x \otimes I |\phi^+\rangle = |\psi^+\rangle, \quad (5.41)$$

which corresponds to the following matrix representation in the computational basis:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}. \quad (5.42)$$

If she wishes to communicate 10, she operates  $\sigma_z$ , obtaining

$$\sigma_z \otimes I |\phi^+\rangle = |\phi^-\rangle, \quad (5.43)$$

with matrix representation

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}. \quad (5.44)$$

Finally, if she wishes to communicate 11, she operates  $i\sigma_y$ , obtaining

$$i\sigma_y \otimes I |\phi^+\rangle = |\psi^-\rangle, \quad (5.45)$$

with matrix representation

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}. \quad (5.46)$$

Up to this point, the circuit in Fig. 5.6 has constructed one of the four Bell states defined in Sec. 3.5 ( $|\phi^+\rangle$ ,  $|\psi^+\rangle$ ,  $|\phi^-\rangle$  and  $|\psi^-\rangle$ ).

- (3) Alice sends her half of the EPR pair to Bob.
- (4) Bob performs the appropriate unitary operations on the EPR pair, measures the two qubits and obtains the two classical bits. First of all, Bob transforms the Bell states into states of the computational basis. As was described in Sec. 3.5, the appropriate circuit for this operation is the inverse of the circuit in Fig. 3.8, which is also the first part of the dense-coding circuit represented in Fig. 5.6. Since Hadamard and CNOT gates are self-inverse, Bob operates

$$(CNOT(H \otimes I))^{-1} = (H \otimes I) CNOT, \quad (5.47)$$

which has matrix representation

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}. \quad (5.48)$$

It is easy to check that

$$\begin{aligned} B|\phi^+\rangle &= |00\rangle, & B|\psi^+\rangle &= |01\rangle, \\ B|\phi^-\rangle &= |10\rangle, & B|\psi^-\rangle &= |11\rangle. \end{aligned} \quad (5.49)$$

Eventually, Bob measures the two qubits in the computational basis, obtaining with unit probability the two desired classical bits.

We stress that dense coding is not possible in classical physics, since a classical bit also has a well-defined value prior to its measurement. In quantum mechanics, there is entanglement. When Alice operates on her half of the EPR pair, she acts not on an isolated qubit, but on an entangled two-qubit system.

## 5.5 Quantum teleportation

Quantum teleportation is one of the most amazing applications of quantum physics to the realm of information theory: it allows for the transmission of quantum information from Alice to Bob, even though Alice sends only classical information to Bob. This possibility could be of practical interest for quantum computation, for example in the transfer of quantum information between different units of a quantum computer. Let us consider the simplest example of teleportation: Alice owns a two level system in some unknown state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (5.50)$$

and she wishes to send this qubit to Bob using only a classical communication channel: she can send only classical bits, not quantum bits. At first sight, the task seems desperate since a measurement of the system would uncontrollably perturb its state and from this measurement Alice can obtain only a single bit of information. We note that describing  $|\psi\rangle$  requires an infinite amount of classical information, since this quantum state lives in a continuous space (it is parametrized by two complex parameters  $\alpha$  and  $\beta$ ). However, quantum teleportation solves the problem, provided that Alice and Bob share a maximally entangled pair of qubits. The protocol for quantum teleportation is outlined in the following (the quantum circuit implementing teleportation is shown in Fig. 5.7):

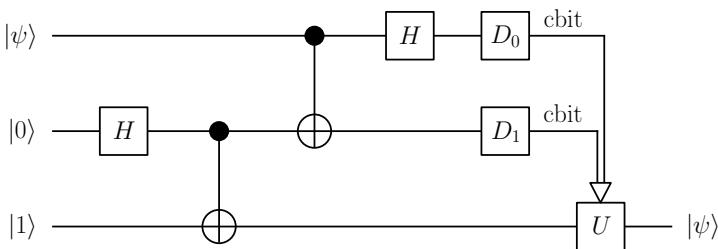


Fig. 5.7 A quantum circuit for teleportation. The first line represents the qubit to be teleported, the second line a qubit possessed by Alice, and the third a qubit possessed by Bob. The measurement performed by Alice (by means of the detectors  $D_0$  and  $D_1$ ) gives two *cbits* (classical bits) of information, which control the unitary transformation  $U$  performed by Bob.

- (1) The first two gates of the circuit in Fig. 5.7 create the Bell state

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \quad (5.51)$$

Indeed,

$$\text{CNOT} (H \otimes I) |01\rangle = |\psi^+\rangle. \quad (5.52)$$

They are operated by a source  $S$  that generates the EPR pair. Then the first half of the EPR pair is sent to Alice and the second half to Bob. Therefore, Alice owns two qubits (the state  $|\psi\rangle$  and half of the EPR pair) and Bob a single qubit (the second half of the EPR pair). Note that, as usual, Alice and Bob can be very far apart. The three-qubit state is given by the direct product

$$\begin{aligned} |\psi\rangle \otimes |\psi^+\rangle &= (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ &= \frac{\alpha}{\sqrt{2}}(|001\rangle + |010\rangle) + \frac{\beta}{\sqrt{2}}(|101\rangle + |110\rangle). \end{aligned} \quad (5.53)$$

- (2) Alice allows the qubit  $|\psi\rangle$  to interact with her half of the EPR pair. This step is necessary. Indeed, if Alice simply performed a measurement in the computational basis, the quantum state  $|\psi\rangle$  would collapse onto  $|0\rangle$  or  $|1\rangle$  and Alice would not obtain enough information to reconstruct it. The way out is a measurement in the Bell basis, given by the states  $|\phi^+\rangle$ ,  $|\phi^-\rangle$ ,  $|\psi^+\rangle$  and  $|\psi^-\rangle$  defined in Sec. 3.5.1. These states constitute a complete orthonormal set and therefore one can expand the states of the computational basis over this basis. This gives

$$\begin{cases} |00\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle + |\phi^-\rangle), \\ |11\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle - |\phi^-\rangle), \\ |01\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle + |\psi^-\rangle), \\ |10\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle - |\psi^-\rangle). \end{cases} \quad (5.54)$$

We insert these relations into Eq. (5.53), obtaining

$$\begin{aligned} |\psi\rangle \otimes |\psi^+\rangle &= \frac{\alpha}{2}(|\phi^+\rangle + |\phi^-\rangle)|1\rangle + \frac{\alpha}{2}(|\psi^+\rangle + |\psi^-\rangle)|0\rangle \\ &\quad + \frac{\beta}{2}(|\psi^+\rangle - |\psi^-\rangle)|1\rangle + \frac{\beta}{2}(|\phi^+\rangle - |\phi^-\rangle)|0\rangle \\ &= \frac{1}{2}|\psi^+\rangle(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2}|\psi^-\rangle(\alpha|0\rangle - \beta|1\rangle) \\ &\quad + \frac{1}{2}|\phi^+\rangle(\alpha|1\rangle + \beta|0\rangle) + \frac{1}{2}|\phi^-\rangle(\alpha|1\rangle - \beta|0\rangle). \end{aligned} \quad (5.55)$$

Therefore, Alice must perform a Bell measurement, obtaining one out of the four states  $|\psi^\pm\rangle$  or  $|\phi^\pm\rangle$ , with equal probability  $p = \frac{1}{4}$ . Note that, as we have seen in Sec. 3.5.1, the Bell measurement can be transformed into a standard measurement in the computational basis, provided that one applies the unitary transformation

$$(H \otimes I) \text{ CNOT} \quad (5.56)$$

before the measurement. This transforms  $|\phi^+\rangle$  to  $|00\rangle$ ,  $|\psi^+\rangle$  to  $|01\rangle$ ,  $|\phi^-\rangle$  to  $|10\rangle$  and  $|\psi^-\rangle$  to  $|11\rangle$  (see again Sec. 3.5.1). Thus, Alice applies the unitary transformation (5.56) to the two qubits that she possesses. This leads to the following global state for the three qubits:

$$\begin{aligned} &\frac{1}{2}|01\rangle(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2}|11\rangle(\alpha|0\rangle - \beta|1\rangle) \\ &+ \frac{1}{2}|00\rangle(\alpha|1\rangle + \beta|0\rangle) + \frac{1}{2}|10\rangle(\alpha|1\rangle - \beta|0\rangle). \end{aligned} \quad (5.57)$$

- (3) Alice measures the two qubits in her possession in the computational basis. The four possible outcomes (00, 01, 10 and 11) give two bits of classical information. As can be seen from Eq. (5.57), if the outcome of Alice's measurement is 00, then the state of Bob's particle collapses onto  $\alpha|1\rangle + \beta|0\rangle$ . Analogously, outcomes 01, 10 and 11 leave the post-measurement state of Bob's particle in  $\alpha|0\rangle + \beta|1\rangle$ ,  $\alpha|1\rangle - \beta|0\rangle$  and  $\alpha|0\rangle - \beta|1\rangle$ , respectively.
- (4) Alice sends these two classical bits to Bob.
- (5) Bob receives these two bits of *classical information*, telling him which of the four possible outcomes of her measurement Alice obtained. Depending on this *classical* message, Bob performs one out of four possible unitary operations  $U$  on his qubit to recover the state  $|\psi\rangle$ . If Alice obtained 00, Bob performs  $U = \sigma_x$ . Analogously, 01, 10 and 11 drive  $U = I$ ,  $U = i\sigma_y$  and  $U = \sigma_z$ , respectively.

We stress that teleportation does not allow one to communicate quantum information faster than light. Indeed, Alice must send two bits of classical information to allow Bob to reconstruct the state  $|\psi\rangle$ . This information is transmitted by classical means, at a speed not greater than that of light. Note also that it is the information about the quantum state, the qubit, that passes from Alice to Bob and not the physical system itself. The physical systems implementing the qubit can be very different in Alice's and Bob's laboratories.

We also emphasize that teleportation is fully consistent with the no-cloning theorem. The quantum state  $|\psi\rangle$  is in Bob's possession at the end of the teleportation process, but the original state is left in  $|0\rangle$  or  $|1\rangle$ , depending on the result of Alice's measurement. The unknown quantum state  $|\psi\rangle$  vanishes in one place and reappears in another.

It is also interesting to note that dense coding and quantum teleportation can be obtained by the same quantum circuit, “cut” in different positions (see Fig. 5.8).

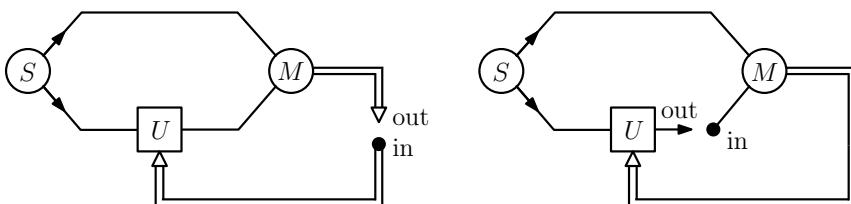


Fig. 5.8 Diagrams representing dense coding (left) and teleportation (right). Double lines represent two classical bits, single lines a quantum bit,  $S$  the EPR source,  $M$  the measurement process,  $U$  the unitary transformation driven by the two classical bits, “in” and “out” the input and output of the circuits.

Finally, we would like to emphasize that teleportation plays a very important role in a number of quantum computation protocols (Gottesman and Chuang, 1999; Knill *et al.*, 2001). It is a powerful tool for transferring quantum states from one

system to another, as would be required in a quantum computer made of several independent units. In particular, it has been proved that teleportation, together with single-qubit operations, is sufficient to achieve universal quantum computation (Gottesman and Chuang, 1999).

**Exercise 5.7** Study the quantum circuit of Fig. 5.9. It achieves teleportation

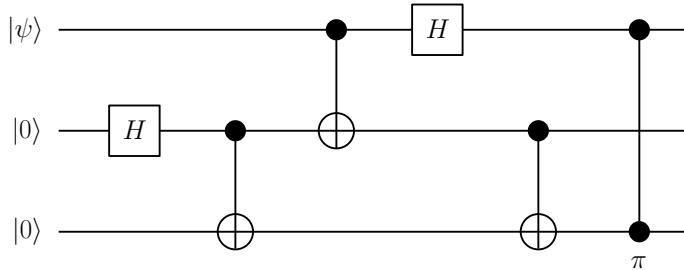


Fig. 5.9 A quantum circuit implementing intraportation. The last gate is a controlled phase shift through an angle  $\pi$ .

via quantum computation (see Brassard *et al.*, 1998). Show that in the output the state  $|\psi\rangle$  is recovered in the third line of the circuit. This circuit is sometimes called *intraportation*, since the CNOT gates are performed between the first-second and second-third qubits. Therefore, in order to implement these CNOT gates, the first two qubits cannot be arbitrarily far away from the third one.

**Exercise 5.8** Study the circuit of Fig. 5.10 for the teleportation of an EPR pair

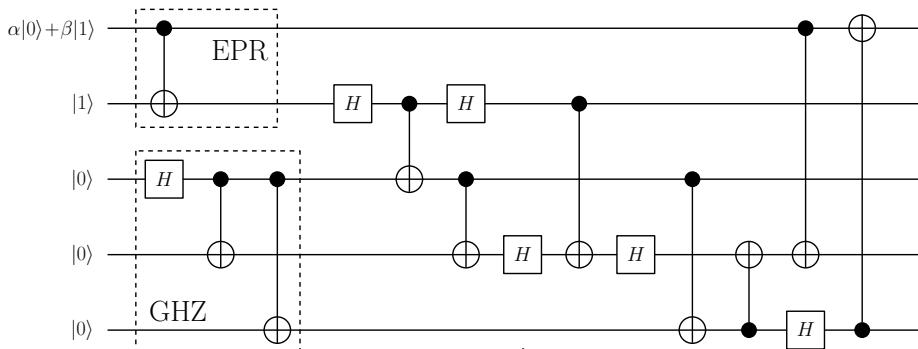


Fig. 5.10 A quantum circuit implementing the quantum teleportation of an entangled pair.

(see Gorbachev and Trubilko, 2000). The first quantum gates generate the entangled state

$$\alpha|01\rangle + \beta|10\rangle \quad (5.58)$$

and the GHZ (Greenberger, Horne and Zeilinger) state

$$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \quad (5.59)$$

Show that at the end the EPR state (5.58) is recovered in the last two lines of the circuit in Fig. 5.10.

### 5.5.1 \* Conclusive teleportation

The wording conclusive teleportation refers to the situation in which the two communicating parties share an entangled pure state which is not maximally entangled. In this case, the conclusive teleportation protocol described below allows them with some positive probability to teleport an unknown quantum state with unit fidelity. That is, either the quantum state is faithfully teleported or the protocol fails.

Alice wants to teleport to Bob the unknown state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ . Alice and Bob share a generic entangled two-qubit state, which can always be written by means of the Schmidt decomposition in the form

$$|\chi\rangle = p|00\rangle + q|11\rangle. \quad (5.60)$$

Without loss of generality, we can assume the coefficients  $p, q \in \mathbb{R}$  and  $0 < p < q < 1$ . It is easy to check that the three-qubit state may be rewritten as

$$\begin{aligned} |\psi\rangle \otimes |\chi\rangle = & \frac{1}{\sqrt{2}} |\phi^+\rangle (p\alpha|0\rangle + q\beta|1\rangle) + \frac{1}{\sqrt{2}} |\phi^-\rangle (p\alpha|0\rangle - q\beta|1\rangle) \\ & + \frac{1}{\sqrt{2}} |\psi^+\rangle (p\beta|0\rangle + q\alpha|1\rangle) + \frac{1}{\sqrt{2}} |\psi^-\rangle (-p\beta|0\rangle + q\alpha|1\rangle). \end{aligned} \quad (5.61)$$

As in the standard teleportation protocol, Alice owns the first two qubits, and Bob the third qubit. We assume that Alice and Bob know the coefficients  $p$  and  $q$ .

If the communicating parties used the standard teleportation protocol, then the final state owned by Bob would depend on  $p$  and  $q$ . For instance, let us assume Alice obtains the state  $|\phi^+\rangle$ , which happens with probability

$$\langle\psi| \otimes \langle\chi| (|\phi^+\rangle \langle\phi^+| \otimes I_B) |\psi\rangle \otimes |\chi\rangle = \frac{1}{2} (p^2|\alpha|^2 + q^2|\beta|^2), \quad (5.62)$$

where  $I_B$  denotes the identity for Bob's qubit. The state of Bob's particle is then projected onto

$$|\psi_{\text{out}}\rangle = \frac{p\alpha|0\rangle + q\beta|1\rangle}{\sqrt{p^2|\alpha|^2 + q^2|\beta|^2}}, \quad (5.63)$$

whose fidelity  $|\langle\psi|\psi_{\text{out}}\rangle|^2$  with the state to be teleported is in general smaller than one.

A different teleportation protocol allows a “conclusive” answer, namely either the teleportation succeeds with unit fidelity or it fails. To understand this protocol, it is convenient to write the initial state as follows:

$$\begin{aligned} |\psi\rangle \otimes |\chi\rangle = & \frac{1}{2} (p|00\rangle + q|11\rangle)(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2} (p|00\rangle - q|11\rangle)(\alpha|0\rangle - \beta|1\rangle) \\ & + \frac{1}{2} (q|01\rangle + p|10\rangle)(\beta|0\rangle + \alpha|1\rangle) + \frac{1}{2} (q|01\rangle - p|10\rangle)(-\beta|0\rangle + \alpha|1\rangle). \end{aligned} \quad (5.64)$$

Alice first measures whether the state of her two qubits is in the subspace spanned by  $|00\rangle$  and  $|11\rangle$ , or in the subspace spanned by  $|01\rangle$  and  $|10\rangle$  (the two possibilities

have equal probability). Let us assume that the measurement selects the  $\{|00\rangle, |11\rangle\}$  subspace (a similar analysis can be done if the subspace  $\{|01\rangle, |10\rangle\}$  is selected). Alice now performs a second, POVM measurement that *conclusively* distinguishes between the two non-orthogonal states  $p|00\rangle + q|11\rangle$  and  $p|00\rangle - q|11\rangle$ , or fails. For this purpose, it is sufficient to use the POVM elements  $F_0$ ,  $F_1$ , and  $F_2$  of Eq. (2.164), with  $r = p/q$ . As discussed in Sec. 2.9.1, such a POVM can never give a wrong identification of the two non-orthogonal states, but we cannot conclude anything when the outcome corresponds to the POVM element  $F_2$ . In that case the teleportation protocol fails, otherwise unit-fidelity teleportation is achieved. By computing the probabilities of the various outcomes as explained in Sec. 2.9.1, it can be checked that the failure probability is  $q^2 - p^2$ . Therefore, perfect teleportation is always obtained for  $p = q$ , when Alice and Bob share maximally entangled states, whereas the protocol always fails in the limit  $p \rightarrow 0$  ( $q \rightarrow 1$ ), when the state  $|\chi\rangle$  is separable.

## 5.6 Quantum mechanics with continuous variables

This section introduces basic aspects of quantum mechanics for continuous variable systems which will be useful later in the book (the necessary mathematical tools are provided in App. A.2). A continuous variable system has an infinite-dimensional Hilbert space and is characterized by observables which can have a continuous spectrum. The simplest class of such systems is provided by single particles living in a one-dimensional space (such as a free particle moving along a line). The corresponding wave function  $\psi(x)$  resides in the infinite-dimensional Hilbert space  $\mathcal{L}^2(\mathbb{R})$  and has unit norm:  $\|\psi\| = 1$ . The average value of any observable  $O$  is given by  $\langle O \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) O \psi(x)$ . In the following we shall focus on prototypical systems, such as the harmonic oscillator. Readers with a background in elementary quantum mechanics and quantum optics can safely skip this section.

### Position-momentum uncertainty relations

First of all, we recall the Heisenberg's uncertainty principle in the mathematically rigorous formulation of Robertson, Eq. (2.43), for the position-momentum operators  $X$  and  $P$ . As proved in App. A.2,  $X$  and  $P$  do not commute,  $[X, P] = i\hbar$  and therefore from Eq. (2.43) we obtain

$$\Delta X \Delta P \geq \frac{|\langle \psi | [X, P] | \psi \rangle|}{2} = \frac{\hbar}{2}, \quad (5.65)$$

for any state  $|\psi\rangle$  of the system; we remind the reader that, for any operator  $A$ , the variance  $\langle (\Delta A)^2 \rangle = \langle (A - \langle A \rangle)^2 \rangle$  and  $\Delta A = \sqrt{\langle (\Delta A)^2 \rangle}$ . In three dimensions, given the position operator  $\mathbf{R} = (R_1, R_2, R_3) \equiv (X, Y, Z)$  and the momentum operator  $\mathbf{P} = (P_1, P_2, P_3) \equiv (P_x, P_y, P_z)$ , we can easily prove that  $[R_i, P_j] = i\hbar \delta_{ij}$ . Therefore, from (2.43) we can obtain that

$$\Delta X \Delta P_x \geq \frac{\hbar}{2}, \quad \Delta Y \Delta P_y \geq \frac{\hbar}{2}, \quad \Delta Z \Delta P_z \geq \frac{\hbar}{2}. \quad (5.66)$$

### Free particle

Let us consider a free particle moving along a line. We assume that the particle is localized at  $x_0$ , with the wave function  $\psi(x) \in \mathcal{L}^2(\mathbb{R})$  given by a Gaussian wave packet:

$$\psi(x) = \langle x | \psi \rangle = \frac{1}{\sqrt{\sqrt{\pi} \delta}} \exp \left[ -\frac{(x - x_0)^2}{2\delta^2} \right]. \quad (5.67)$$

It is easy to check (see exercise 5.9) that the mean values of position and momentum are

$$\langle X \rangle = x_0, \quad \langle P \rangle = 0, \quad (5.68)$$

and the corresponding variances are

$$\langle (\Delta X)^2 \rangle = \langle (X - \langle X \rangle)^2 \rangle = \frac{\delta^2}{2}, \quad \langle (\Delta P)^2 \rangle = \langle (P - \langle P \rangle)^2 \rangle = \frac{\hbar^2}{2\delta^2}. \quad (5.69)$$

In the momentum-space representation the wave function  $\tilde{\psi}(p)$  is given by the Fourier transform of  $\psi(x)$  and again has a Gaussian shape:

$$\begin{aligned} \tilde{\psi}(p) &= \langle p | \psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx \exp \left( -\frac{ipx}{\hbar} \right) \psi(x) \\ &= \sqrt{\frac{\delta}{\sqrt{\pi}\hbar}} \exp \left( -\frac{\delta^2 p^2}{2\hbar^2} \right) \exp \left( -\frac{ipx_0}{\hbar} \right). \end{aligned} \quad (5.70)$$

The density matrix corresponding to the Gaussian wave packet (5.67) reads

$$\langle x | \rho | x' \rangle = \langle x | \psi \rangle \langle \psi | x' \rangle = \frac{1}{\sqrt{\pi}\delta} \exp \left( -\frac{(x - x_0)^2 + (x' - x_0)^2}{2\delta^2} \right) \quad (5.71)$$

and is drawn in Fig. 5.11.

**Exercise 5.9** Check that the Gaussian wave packet

$$\psi(x, t) = \frac{1}{\sqrt{\sqrt{\pi} (1 + i \frac{\hbar t}{m\delta^2}) \delta}} \exp \left\{ -\frac{(x - x_0 - \frac{p_0 t}{m})^2}{2\delta^2 (1 + i \frac{\hbar t}{m\delta^2})} + \frac{i}{\hbar} \left[ p_0(x - x_0) - \frac{p_0^2}{2m} t \right] \right\} \quad (5.72)$$

is solution of the Schrödinger equation for a free particle moving in one dimension:

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = H\psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t). \quad (5.73)$$

Show that

$$\langle X \rangle = x_0 + \frac{p_0}{m} t, \quad \langle P \rangle = p_0; \quad (5.74)$$

namely, that the wave packet moves with constant velocity  $p_0/m$ . Finally, check that

$$\langle (\Delta X)^2 \rangle = \frac{\delta^2}{2} \left[ 1 + \frac{\hbar^2 t^2}{m^2 \delta^4} \right], \quad \langle (\Delta P)^2 \rangle = \frac{\hbar^2}{2\delta^2}. \quad (5.75)$$

This implies that

$$\Delta X \Delta P = \frac{\hbar}{2} \sqrt{1 + \frac{\hbar^2 t^2}{m^2 \delta^4}}, \quad (5.76)$$

where we have defined  $\Delta X = \sqrt{\langle (\Delta X)^2 \rangle}$  and  $\Delta P = \sqrt{\langle (\Delta P)^2 \rangle}$ . Therefore, (5.72) is a minimum uncertainty wave packet ( $\Delta X \Delta P = \hbar/2$ ) only at time  $t = 0$ .

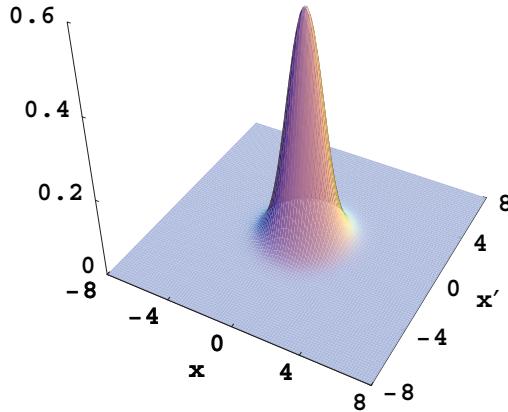


Fig. 5.11 The density matrix corresponding to a Gaussian wave packet centred at  $x_0 = 0$ . In this figure we set  $\delta = 1$ .

### Harmonic oscillator

Let us consider the harmonic oscillator, whose Hamiltonian reads

$$H = \frac{P^2}{2m} + \frac{m\omega^2 X^2}{2}, \quad (5.77)$$

where the operator  $P$  reads, in the position representation,  $P = -i\hbar(d/dx)$ . The stationary states of the harmonic oscillator are the eigenfunctions  $|n\rangle$  of the Hamiltonian operator (5.77); that is,

$$H|n\rangle = E_n|n\rangle. \quad (5.78)$$

As shown in quantum mechanics textbooks, the eigenvalues read

$$E_n = \hbar\omega \left(n + \frac{1}{2}\right) \quad (5.79)$$

and the corresponding stationary states are given by

$$\phi_n(x) \equiv \langle x|n\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{2^{n/2}\sqrt{n!}} H_n \left(\sqrt{\frac{m\omega}{\hbar}}x\right) \exp\left(-\frac{1}{2}\frac{m\omega}{\hbar}x^2\right), \quad (5.80)$$

where  $H_n$  denotes the  $n$ -th Hermite polynomial.<sup>1</sup> Note that the energy of the ground state,  $E_0 = \frac{1}{2}\hbar\omega$ , is known as the *zero-point energy* and can be seen as a consequence of the Heisenberg principle. Indeed, it can be seen analogously to exercise 5.9 that the product of the uncertainties  $\Delta X$  and  $\Delta P$  is equal to  $\frac{\hbar}{2}$ , namely the minimum uncertainty permitted by the Heisenberg principle.

---

<sup>1</sup>The Hermite polynomials satisfy the recurrence relation

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - 2nH_{n-1}(\xi)$$

and the first few Hermite polynomials are

$$H_0(\xi) = 1, \quad H_1(\xi) = 2\xi, \quad H_2(\xi) = 4\xi^2 - 2, \quad H_3(\xi) = 8\xi^3 - 12\xi, \quad \dots$$

Hamiltonian (5.77) can also be written as

$$H = \hbar\omega (a^\dagger a + \frac{1}{2}), \quad (5.81)$$

where

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega X + iP), \quad a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega X - iP). \quad (5.82)$$

Note that  $[a, a^\dagger] = 1$ . It can be shown (see exercise 5.10) that the action of  $a$  and  $a^\dagger$  on the stationary state  $|n\rangle$  is as follows:

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \quad (5.83)$$

Due to their action on the *Fock states*  $|n\rangle$ ,  $a^\dagger$  and  $a$  are known as the creation and annihilation operators, respectively. It follows from (5.83) that the *number operator*  $N = a^\dagger a$  has the property  $a^\dagger a|n\rangle = n|n\rangle$ ; that is,  $N$  has the same eigenstates as  $H$  (the Fock states are therefore the eigenstates of the number operator).

**Exercise 5.10** Prove Eq. (5.83).

#### Coherent states

To be specific, we discuss the *coherent states* of the harmonic oscillator for a quantized single mode of the electromagnetic field, described by the Hamiltonian  $H = \hbar\omega (a^\dagger a + \frac{1}{2})$ . In this case  $N = a^\dagger a$  is the photon number operator and  $|n\rangle$  ( $N|n\rangle = n|n\rangle$ ) is a  $n$ -photon state. The coherent states are defined as the eigenstates  $|\alpha\rangle$  of the (non-Hermitian) annihilation operator  $a$ ; that is,

$$a|\alpha\rangle = \alpha|\alpha\rangle, \quad \alpha \in \mathbb{C}. \quad (5.84)$$

It can be shown (see exercise 5.11) that:

- (i) The representation of a coherent state in the Fock basis (that is, in the basis of the Fock states) is given by

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (5.85)$$

- (ii) The mean number of photons in the coherent state  $|\alpha\rangle$  is given by  $\bar{n} = \langle\alpha|N|\alpha\rangle = \langle\alpha|a^\dagger a|\alpha\rangle = |\alpha|^2$  while the root mean square deviation in the photon number  $\Delta n = \sqrt{\bar{n}}$ . The photon-number distribution of light in a coherent state is Poissonian. Indeed, from the Fock representation (5.85) of a coherent state we obtain

$$p(n) = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2} = \frac{\bar{n}^n}{n!} e^{-\bar{n}}. \quad (5.86)$$

- (iii) During its time evolution, a coherent state remains a minimum uncertainty wave packet and the temporal evolution of the mean values of position and momentum are the same as for a classical particle. That is, for an initial state  $|\psi(0)\rangle = |\alpha\rangle$ , up to a global phase factor,  $|\psi(t)\rangle = |\alpha(t)\rangle$ , where  $\alpha(t) = e^{-i\omega t}\alpha$ .

(iv) Coherent states are not orthogonal,

$$\langle \alpha | \alpha' \rangle = e^{-|\alpha|^2/2} e^{-|\alpha'|^2/2} e^{\alpha^* \alpha'} \neq \delta(\alpha - \alpha'), \quad (5.87)$$

and satisfy the closure relation

$$\frac{1}{\pi} \int \int d\text{Re}(\alpha) d\text{Im}(\alpha) |\alpha\rangle \langle \alpha| = I. \quad (5.88)$$

Therefore the coherent states are an *overcomplete basis*, since any coherent state can be expanded in terms of the other states:

$$|\alpha\rangle = \frac{1}{\pi} \int \int d\text{Re}(\alpha') d\text{Im}(\alpha') |\alpha'\rangle \langle \alpha'| \alpha \rangle \quad (5.89)$$

$$= \frac{1}{\pi} \int \int d\text{Re}(\alpha') d\text{Im}(\alpha') e^{-|\alpha|^2/2} e^{-|\alpha'|^2/2} e^{\alpha'^* \alpha} |\alpha'\rangle. \quad (5.90)$$

**Exercise 5.11** Prove properties (i)–(iv) stated above for a coherent state.

The ground state  $|0\rangle$  is the so-called *vacuum state*, in which the photon number is equal to zero. It fulfills the conditions  $N|0\rangle = 0$  and  $a|0\rangle = 0$ . The second condition tells us that the vacuum state is also a coherent state  $|\alpha\rangle$  with  $\alpha = 0$ .

To investigate the properties of coherent states, it is useful to introduce the *displacement operator*

$$D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a). \quad (5.91)$$

Since  $i(\alpha a^\dagger - \alpha^* a)$  is Hermitian,  $D(\alpha)$  is unitary. It can be shown (see exercise 5.12) that coherent states are displacements of the vacuum state:

$$|\alpha\rangle = D(\alpha)|0\rangle. \quad (5.92)$$

**Exercise 5.12** Show that for infinitesimal  $\delta\alpha$

$$D^\dagger(\delta\alpha) a D(\delta\alpha) = a + \delta\alpha I \quad (5.93)$$

and conclude that for any complex number  $\alpha$

$$D^\dagger(\alpha) a D(\alpha) = a + \alpha I. \quad (5.94)$$

Then prove that

$$a D(-\alpha)|\alpha\rangle = 0, \quad (5.95)$$

so that  $D(-\alpha)|\alpha\rangle$  is the vacuum state, implying Eq. (5.92).

Let us now define the *quadratures*

$$X_1 = \frac{1}{2}(a + a^\dagger), \quad X_2 = \frac{1}{2i}(a - a^\dagger). \quad (5.96)$$

These dimensionless operators are Hermitian and obey the commutation relation

$$[X_1, X_2] = \frac{i}{2}. \quad (5.97)$$

Since for a harmonic oscillator we derive from Eq. (5.82)

$$X = \sqrt{\frac{2\hbar}{m\omega}} X_1, \quad P = \sqrt{2\hbar m\omega} X_2, \quad (5.98)$$

it is clear that  $X_1$  and  $X_2$  can be regarded as dimensionless position and momentum operators of the electromagnetic oscillator. For a coherent state  $|\alpha\rangle$ ,  $\langle\alpha|X_1|\alpha\rangle = \text{Re}(\alpha)$  and  $\langle\alpha|X_2|\alpha\rangle = \text{Im}(\alpha)$ . From the commutator (5.97) we obtain the uncertainty relation

$$\Delta X_1 \Delta X_2 \geq \frac{1}{4}. \quad (5.99)$$

For a coherent state  $\Delta X_1 \Delta X_2 = \frac{1}{4}$  (see exercise 5.13), that is, coherent states are minimum uncertainty states.

**Exercise 5.13** Compute  $\Delta X_1$  and  $\Delta X_2$  for coherent and for Fock states.

\* A note on the physical meaning of coherent states

The introduction of coherent states as eigenstates of the annihilation operator does not immediately highlight their physical meaning. In what follows, inspired by Goldman *et al.* (1960), we show that coherent states can be obtained from the solution of the dynamics of a quantum harmonic oscillator driven by a classical force. More precisely: (i) the oscillator is initially in its ground state (which is a coherent state); (ii) we apply an arbitrary time-dependent force; (iii) the dynamical evolution of the quantum harmonic oscillator corresponds to the evolution of the classical harmonic oscillator, with additionally the same quantum fluctuations due to the Heisenberg uncertainty principle as for the ground state ( $\Delta X_1 = \Delta X_2 = \frac{1}{2}$ ). That is, the quantum solution is a Gaussian identical to the vacuum coherent state but centered on the classical solution. We conclude from this result that we can obtain a generic coherent state by means of a classical force  $f(t)$ , acting on a quantized single mode of the electromagnetic field initially in the vacuum state. The proof of this result follows.

The Hamiltonian of a driven quantum harmonic oscillator reads

$$\begin{aligned} H &= \frac{P^2}{2m} + \frac{m\omega^2 X^2}{2} - f(t)X \\ &= \hbar\omega \left( a^\dagger a + \frac{1}{2} \right) - f(t)\sqrt{\frac{\hbar}{2m\omega}} (a + a^\dagger). \end{aligned} \quad (5.100)$$

We look for solutions to the Schrödinger equation  $i\hbar|\dot{\psi}\rangle = H|\psi\rangle$  of the form

$$|\psi(t)\rangle = c(t)e^{\alpha(t)a^\dagger} e^{\beta(t)a} e^{\gamma(t)a^\dagger a} |\psi(0)\rangle. \quad (5.101)$$

We obtain

$$\begin{aligned} |\dot{\psi}(t)\rangle &= \dot{c}(t)e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle + c(t)\dot{\alpha}a^\dagger e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle \\ &\quad + c(t)\dot{\beta}e^{\alpha a^\dagger} ae^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle + c(t)\dot{\gamma}e^{\alpha a^\dagger} e^{\beta a^\dagger} ae^{\gamma a^\dagger a} |\psi(0)\rangle, \end{aligned} \quad (5.102)$$

which can be conveniently written as

$$|\dot{\psi}(t)\rangle = G(t)e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle, \quad (5.103)$$

where  $G(t)$  is an operator to be determined. For this purpose, we must reorder some operators in Eq. (5.102), in such a way that all the four terms in the right-hand side of (5.102) end with  $e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle$ . We have

$$\begin{aligned} e^{\alpha a^\dagger} a &= \left( I + \alpha a^\dagger + \frac{\alpha^2}{2!} a^{\dagger 2} + \dots \right) a \\ &= a \left( I + \alpha a^\dagger + \frac{\alpha^2}{2!} a^{\dagger 2} + \dots \right) - \alpha \left( I + \alpha a^\dagger + \dots \right) = (a - \alpha I) e^{\alpha a^\dagger}, \end{aligned} \quad (5.104)$$

where we have used the relation  $[a^{\dagger n}, a] = -n (a^\dagger)^{n-1}$ . The third term in the right-hand side of (5.102) can then be written as

$$c\dot{\beta}(a - \alpha I) e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a} |\psi(0)\rangle. \quad (5.105)$$

The last term in (5.102) is manipulated in a similar way as follows:

$$\begin{aligned} c\dot{\gamma}e^{\alpha a^\dagger} e^{\beta a} a^\dagger a e^{\gamma a^\dagger a} &= c\dot{\gamma}e^{\alpha a^\dagger} (a^\dagger a + \beta a) e^{\beta a} e^{\gamma a^\dagger a} \\ &= c\dot{\gamma} \left( a^\dagger a - \alpha a^\dagger + \beta a - \alpha\beta I \right) e^{\alpha a^\dagger} e^{\beta a} e^{\gamma a^\dagger a}. \end{aligned} \quad (5.106)$$

By inserting (5.105) and (5.106) into (5.102) we obtain

$$G(t) = c\dot{\gamma}a^\dagger a + c(\dot{\alpha} - \alpha\dot{\gamma})a^\dagger + c(\dot{\beta} + \beta\dot{\gamma})a + (\dot{c} - c\dot{\gamma}\alpha\beta)I. \quad (5.107)$$

After substitution of (5.107) into (5.103), we can derive from the Schrödinger equation the relations

$$\begin{aligned} \dot{\gamma} &= -i\omega, \quad \dot{\alpha} + i\omega\alpha = \frac{i}{\sqrt{2\hbar m\omega}} f(t), \\ \dot{\beta} - i\omega\beta &= \frac{i}{\sqrt{2\hbar m\omega}} f(t), \quad \frac{\dot{c}}{c} = -i\frac{\omega}{2} + \alpha(\dot{\beta} - i\omega\beta). \end{aligned} \quad (5.108)$$

From (5.101) we obtain the boundary conditions  $\alpha(0) = \beta(0) = 0$  and  $|c(0)| = 1$ . We then obtain from integration of (5.108)

$$\begin{aligned} \gamma &= -i\omega t, \quad \beta = -\alpha^*, \\ \alpha(t) &= \frac{i}{\sqrt{2\hbar m\omega}} e^{-i\omega t} \int_{-\infty}^t dt' f(t') e^{i\omega t'}, \\ c(t) &= e^{-i\frac{\omega}{2}t} \exp \left( -\frac{1}{2\hbar m\omega} \int_{-\infty}^t dt' f(t') e^{-i\omega t'} \int_{-\infty}^{t'} dt'' f(t'') e^{-i\omega t''} \right). \end{aligned} \quad (5.109)$$

We can then write the solution to the Schrödinger equation as follows:

$$|\psi(t)\rangle = c(t) e^{\alpha a^\dagger} e^{-\alpha^* a} e^{-i\omega a^\dagger a t} |\psi(0)\rangle. \quad (5.110)$$

If we assume that at  $t = 0$  the oscillator is in the ground state, then

$$e^{-\alpha^* a} e^{-i\omega a^\dagger a t} |\psi(0)\rangle = |\psi(0)\rangle. \quad (5.111)$$

Then Eq. (5.110) becomes

$$|\psi(t)\rangle = c(t)e^{\alpha(t)a^\dagger} |0\rangle = c(t) \sum_{n=0}^{\infty} \frac{[\alpha(t)]^n}{\sqrt{n!}} \frac{a^{\dagger n}}{\sqrt{n!}} |0\rangle = |\alpha(t)\rangle. \quad (5.112)$$

The Hamilton equations of motion for the driven classical harmonic oscillator are  $\dot{x} = \frac{p}{m}$  and  $\dot{p} = -m\omega^2x + f(t)$ . We define the classical variable

$$\alpha_c = \sqrt{\frac{m\omega}{2\hbar}} \left( x + i \frac{1}{m\omega} p \right). \quad (5.113)$$

It is easy to see that  $\alpha_c$  obeys the same equation we have written for  $\alpha$  in (5.108) and therefore we can identify  $\alpha(t)$  and  $\alpha_c(t)$ , provided we have as initial conditions the ground state  $[\alpha(0) = 0]$  in the quantum case and  $\alpha_c(0) = 0$  in the classical case  $[x(0) = p(0) = 0]$ .

### Squeezed states

Given two observables  $A$  and  $B$  which satisfy the commutation relation  $[A, B] = iC$ , we know from the Heisenberg's uncertainty principle (2.43) that  $\Delta A \Delta B \geq \frac{1}{2} |\langle \psi | C | \psi \rangle|$ . We say that a state  $|\psi\rangle$  is an ideal *squeezed state* if for one of the observables (for instance  $A$ )

$$(\Delta A)^2 < \frac{1}{2} |\langle \psi | C | \psi \rangle| \quad (5.114)$$

and moreover the minimum uncertainty relation

$$\Delta A \Delta B = \frac{1}{2} |\langle \psi | C | \psi \rangle| \quad (5.115)$$

is fulfilled. For the quadratures  $X_1$  and  $X_2$  of a single mode of the electromagnetic field these conditions read

$$\Delta X_1 < \frac{1}{2}, \quad \Delta X_1 \Delta X_2 = \frac{1}{4} \quad (5.116)$$

(of course, also  $\Delta X_2 < \frac{1}{2}$  could give a squeezed state when  $\Delta X_1 > \frac{1}{2}$  and  $\Delta X_1 \Delta X_2 = \frac{1}{4}$ ). Reducing the uncertainty in one of the quadratures can be used to enhance precision measurements and finds applications in a variety of fields, for example in the detection of gravitational waves. Note that neither a Fock state nor a coherent state are squeezed states. A Fock state does not satisfy both conditions in Eq. (5.116), while a coherent state is a minimum uncertainty state, but  $\Delta X_1 = \Delta X_2 = \frac{1}{4}$  (see exercise 5.13).

To generate a squeezed state, we can consider the Hamiltonian

$$H = i\hbar (ga^{\dagger 2} - g^*a^2), \quad (5.117)$$

where  $g$  is a constant.<sup>2</sup> Starting from the vacuum state, the following field state is generated:

$$|\psi(t)\rangle = \exp [(ga^{\dagger 2} - g^*a^2)t] |0\rangle, \quad (5.118)$$

<sup>2</sup>This Hamiltonian can be obtained in the framework of *nonlinear optics*, when a non-linear crystal is pumped by a laser with frequency twice the frequency of the mode of interest, namely incoming pump photons of frequency  $2\nu$  are converted into two photons of lower frequency  $\nu$ , for details see for instance Scully and Zubairy (1997).

and this leads us to define the unitary *squeezing operator*

$$S(z) = \exp\left(\frac{1}{2}z^*a^2 - \frac{1}{2}za^{\dagger 2}\right), \quad (5.119)$$

where  $z \equiv 2gt$  is a complex number,  $z = re^{i\phi}$ . It is easy to check that

$$S^\dagger(z) = S^{-1}(z) = S(-z). \quad (5.120)$$

It is useful to define the squeezed creation and annihilation operators  $\tilde{a} = S(z)aS^\dagger(z)$  and  $\tilde{a}^\dagger = S(z)a^\dagger S^\dagger(z)$ . Using the Baker-Campbell-Hausdorff formula (A.101), we find (see exercise 5.14)

$$\begin{aligned} \tilde{a} &= a \cosh r + a^\dagger e^{i\phi} \sinh r, \\ \tilde{a}^\dagger &= ae^{-i\phi} \sinh r + a^\dagger \cosh r. \end{aligned} \quad (5.121)$$

This transformation mapping the operators  $a$  and  $a^\dagger$  into  $\tilde{a}$  and  $\tilde{a}^\dagger$  is canonical since, as it can be readily checked,  $[\tilde{a}, \tilde{a}^\dagger] = 1$ . It is known as *Bogoliubov transformation* and it describes the evolution of Gaussian states when the Hamiltonian is at most bilinear in the original creation and annihilation operators, as is the case in Eq. (5.117).

**Exercise 5.14** By using the Baker-Campbell-Hausdorff formula (A.101), with  $A = \frac{1}{2}z^*a^2 - \frac{1}{2}za^{\dagger 2}$  and  $B = a$ , check Eq. (5.121) up to second order in  $r$ .

By acting with the squeezing operator on the vacuum, we obtain the *squeezed vacuum*

$$|z\rangle \equiv S(z)|0\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} (-e^{i\phi} \tanh r)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle, \quad (5.122)$$

which reduces to the standard vacuum for  $r = 0$ . A remarkable property of the squeezed vacuum (5.122) is that it contains only states with an even number of photons. For this reason the squeezed states are also known as two-photon states.<sup>3</sup>

Note that all minimum-uncertainty states are *displaced squeezed vacuums* (for a proof of this statement, see for instance Leonhardt (1997)):

$$|\alpha, z\rangle = D(\alpha)S(z)|0\rangle \quad (5.123)$$

(note that  $D(\alpha)S(z)|0\rangle \neq S(z)D(\alpha)|0\rangle$ ).

**Exercise 5.15** Using the relation (see for instance Barnett and Radmore (1997))

$$S(z) = \exp\left(-\frac{\nu}{2\mu}a^{\dagger 2}\right) \mu^{-(a^\dagger a + \frac{1}{2})} \exp\left(\frac{\nu^*}{2\mu}a^2\right), \quad (5.124)$$

with  $\mu = \cosh r$  and  $\nu = e^{i\phi} \sinh r$ , prove Eq. (5.122).

---

<sup>3</sup>Hamiltonian (5.117) describes the creation of photons in pairs, since each pump photon is converted into two photons of half the pump frequency.

To prove that a squeezed state is a minimum uncertainty state, we first generalize the quadratures (5.96) by defining

$$X_\theta = \frac{1}{2} (a e^{-i\theta} + a^\dagger e^{i\theta}) . \quad (5.125)$$

We have

$$\langle z|X_\theta|z\rangle = \frac{1}{2} e^{-i\theta} \langle 0|S(z)^\dagger a S(z)|0\rangle + \frac{1}{2} e^{i\theta} \langle 0|S^\dagger(z) a^\dagger S(z)|0\rangle = 0 , \quad (5.126)$$

where the last equality follows from Eq. (5.121). Analogously, we obtain

$$\Delta X_\theta^2 = \langle z|X_\theta^2|z\rangle = \frac{1}{4} \left[ \sin^2 \left( \theta - \frac{\phi}{2} \right) e^{2r} + \cos^2 \left( \theta - \frac{\phi}{2} \right) e^{-2r} \right] . \quad (5.127)$$

If we choose  $\theta_1 = \phi/2$  and  $\theta_2 = (\phi + \pi)/2$ , we have

$$\Delta X_{\theta_1} = \frac{1}{2} e^{-r}, \quad \Delta X_{\theta_2} = \frac{1}{2} e^r, \quad \Delta X_{\theta_1} \Delta X_{\theta_2} = \frac{1}{4} , \quad (5.128)$$

so that we always have a minimum uncertainty state, squeezed provided  $r \neq 0$  (for  $r = 0$  the state is coherent).

### Position representation for squeezed states

For a better intuitive understanding of coherent and squeezed states, it is useful to have other representations besides the above discussed Fock-state representation. The quadrature representation of a squeezed state  $|z\rangle$  is obtained by projecting the state on the ket  $|x_\theta\rangle$  ( $X_\theta|x_\theta\rangle = x_\theta|x_\theta\rangle$ ), and in this way one obtains Gaussian states, see for instance Barnett and Radmore (1997).

Here we follow a simpler approach and directly look for Gaussian solutions of the Schrödinger equation for the harmonic oscillator in the position representation:

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + \frac{m\omega^2 x^2}{2} \psi(x, t) . \quad (5.129)$$

A Gaussian solution to the Schrödinger equation is given by

$$\psi(x, t) = N(t) \exp \left\{ -\frac{1}{2} A(t)[x - \bar{x}(t)]^2 + i\chi(x, t) \right\} , \quad (5.130)$$

where

$$N^2(t) = \frac{\alpha}{\sqrt{\pi}} \frac{1}{C(t) + i\sqrt{BS(t)}} , \quad (5.131)$$

$$C(t) = \cos(\omega t + \phi), \quad S(t) = \sin(\omega t + \phi), \quad B = \frac{\hbar^2 \alpha^4}{m^2 \omega^2} , \quad (5.132)$$

$$A(t) = A_r(t) + iA_i(t) , \quad (5.133)$$

$$A_r(t) = \frac{m\omega\sqrt{B}}{\hbar[C^2(t) + BS^2(t)]}, \quad A_i(t) = \frac{m\omega(1-B)C(t)S(t)}{\hbar[C^2(t) + BS^2(t)]} , \quad (5.134)$$

$$\bar{x}(t) = \bar{x}_0 C(t), \quad \chi(x, t) = -\frac{\omega m}{\hbar} \bar{x}_0 x S(t) + \frac{m\omega}{2\hbar} \bar{x}_0^2 C(t) S(t) . \quad (5.135)$$

Different values of the parameters  $\alpha$ ,  $\bar{x}_0$  and  $\phi$  correspond to different solutions to the Schrödinger equation.<sup>4</sup> The wave function in the momentum representation is given by the Fourier transform of the wave function in the position representation. Since the Fourier transform of a Gaussian is a Gaussian, also the wave function in the momentum representation is a Gaussian wave packet. Due to the time dependence of  $A(t)$ , it is not obvious how the Gaussian packet (5.130) evolves. A clear illustration is provided by the phase space representations of quantum mechanics.

<sup>4</sup>Note that, for the sake of simplicity, we have not considered in (5.130) the most general Gaussian wave packet for the harmonic oscillator, see Dodonov *et al.* (1988).

### Phase-space representations

The cornerstone of quantum mechanics in phase space is the *Wigner function*. The Wigner function provides a pictorial phase-space representation of the abstract notion of a quantum state and allows us to compute the quantum mechanical expectation values of observables in terms of phase space-averages. The Wigner phase space distribution function of a quantum state described by a density operator  $\rho$  reads

$$W(x, p) \equiv \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dy e^{-\frac{i}{\hbar} py} \left\langle x + \frac{y}{2} \middle| \rho \middle| x - \frac{y}{2} \right\rangle, \quad (5.136)$$

where  $x$  and  $p$  are position and momentum but could be any pair of conjugate variables. Note that for, the sake of simplicity, we consider here the case of a single particle moving along a straight line. For a pure state, described by a state vector  $|\psi\rangle$  ( $\rho = |\psi\rangle\langle\psi|$ ), expression (5.136) reduces to

$$W(x, p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dy e^{-\frac{i}{\hbar} py} \psi\left(x + \frac{y}{2}\right) \psi^*\left(x - \frac{y}{2}\right). \quad (5.137)$$

The factor  $\frac{1}{2\pi\hbar}$  ensures the normalization:

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp W(x, p) = 1. \quad (5.138)$$

Note, however, that the Wigner function cannot be interpreted as a probability distribution in the phase space, since  $W$  is real but not in general positive definite.<sup>5</sup> The fact that one cannot define a true phase space representation in quantum mechanics is rooted in the Heisenberg uncertainty principle. Since a particle cannot simultaneously have well defined values for position and momentum, one cannot define a probability that a particle has position  $x$  and momentum  $p$ .<sup>6</sup> The  $x$  and  $p$  probability distributions are given by the marginals of the Wigner function:

$$\int_{-\infty}^{\infty} dp W(x, p) = \langle x | \rho | x \rangle, \quad \int_{-\infty}^{\infty} dx W(x, p) = \langle p | \rho | p \rangle. \quad (5.139)$$

For a pure state  $|\psi\rangle$ ,  $\langle x | \rho | x \rangle = |\psi(x)|^2$  and  $\langle p | \rho | p \rangle = |\tilde{\psi}(p)|^2$ .

**Exercise 5.16** Prove Eq. (5.139).

For a Gaussian wave packet,

$$\psi(x) = \frac{1}{\sqrt{\sqrt{\pi}\delta}} \exp\left(-\frac{(x-x_0)^2}{2\delta^2}\right) \exp\left(\frac{i}{\hbar} p_0 x\right), \quad (5.140)$$

---

<sup>5</sup>Actually, in the case of pure states only for Gaussian wave packets the Wigner function is positive everywhere, see Hudson (1974).

<sup>6</sup>The Wigner function is referred to as a “quasi-probability distribution”, since it bears some resemblance to a classical phase-space distribution functions. It provides useful insights in particular into the connection between classical and quantum mechanics, see for instance Schleich (2001); Zachos *et al.* (2005).

it can be readily checked from Eq. (5.137) that the Wigner function is Gaussian both as a function of  $x$  and  $p$ :

$$W(x, p) = \frac{1}{\pi\hbar} \exp \left( -\frac{(x - x_0)^2}{\delta^2} - \frac{\delta^2(p - p_0)^2}{\hbar^2} \right). \quad (5.141)$$

For the squeezed wave packet (5.130), a contour level  $W_{\max}/e$  of the Wigner function is shown in Fig 5.12 ( $W_{\max}$  is the maximum value of  $W$ ). This curve is a (parametrically dependent on  $t$ ) ellipse in phase space, whose equation can be written in the form

$$\begin{aligned} a[x - \bar{x}(t)]^2 + b[p - m\dot{x}(t)]^2 + 2c[x - \bar{x}(t)][p - m\dot{x}(t)] &= 1, \\ a = \frac{A_r^2 + A_i^2}{A_r}, \quad b = \frac{1}{\hbar^2 A_r}, \quad c = \frac{A_i}{\hbar A_r}. \end{aligned} \quad (5.142)$$

Note that the area of the ellipse remains constant:  $A = \pi\hbar$ . Owing to the term proportional to  $(x - \bar{x})(p - m\dot{x})$ , the ellipse in general does not have its axes parallel to the coordinate axes  $x, p$ . This term disappears when  $A_i = 0$ , that is, for  $B = 1$  ( $\alpha^2 = \frac{m\omega}{\hbar}$ ). This special case corresponds to coherent states, for which a contour plot of the Wigner function becomes, in the coordinate plane  $(x, \frac{p}{m\omega})$ , a rigidly moving circle:

$$[x - \bar{x}(t)]^2 + \frac{[p - m\dot{x}(t)]^2}{m^2\omega^2} = \frac{\hbar}{m\omega}. \quad (5.143)$$

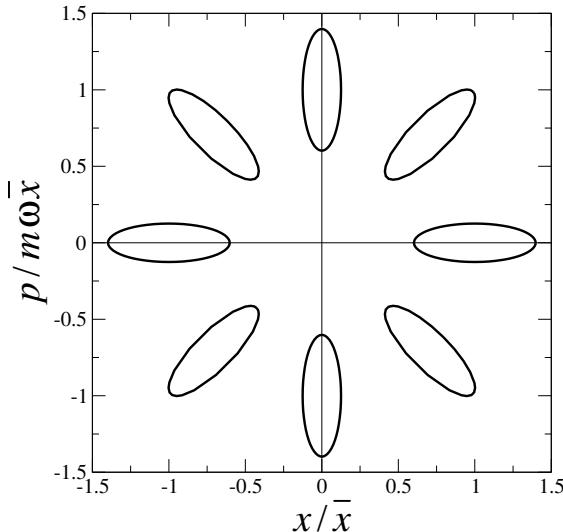


Fig. 5.12 Contour level  $W = W_{\max}/e$  of the Wigner function (corresponding to the value  $g_F = 0$  of the Fermi function described below), for the squeezed state (5.130) of a harmonic oscillator, with  $\frac{m\omega\bar{x}^2}{\hbar} = 20$ ,  $B = 0.1$  and  $\omega t + \phi$  ranging from 0 to  $\frac{7\pi}{4}$ , in steps of  $\frac{\pi}{4}$ .

A different phase space approach is based on the *Fermi g<sub>F</sub> function*. As pointed out by Fermi (1930), the state of a quantum system may be defined in two completely

equivalent ways: by its wave function  $\psi(x) = \langle x|\psi\rangle$  or by measuring a physical quantity  $g_F(x, p)$ . Given the measurement outcome  $g_F(x, p) = \bar{g}$ ,  $\psi(x)$  is obtained as solution of the eigenvalue equation  $g_F(x, p)\psi(x) = \bar{g}\psi(x)$ , where  $p = -i\hbar\partial_x$ . On the other hand, given the wave function  $\psi(x)$  it is always possible to find an operator  $g_F(x, p)$  such that

$$g_F(x, p)\psi(x) = 0. \quad (5.144)$$

Using the polar decomposition

$$\psi(x) = R(x)e^{\frac{i}{\hbar}S(x)}, \quad (5.145)$$

where  $R(x)$  and  $S(x)$  are real functions [ $R(x) \geq 0$  for any  $x$ ], it is easy to check that identity (5.144) is fulfilled by taking

$$g_F(x, -i\hbar\partial_x) = [-i\hbar\partial_x - S'(x)]^2 + \hbar^2 \frac{R''(x)}{R(x)}. \quad (5.146)$$

Equation (5.144) implies that the corresponding physical quantity  $g_F(x, p)$  takes the value  $\bar{g} = 0$ . The equation

$$g_F(x, p) = [p - S'(x)]^2 + \hbar^2 \frac{R''(x)}{R(x)} = 0 \quad (5.147)$$

defines a curve in the two-dimensional phase space. In other words, as expected from Heisenberg uncertainty principle, we cannot identify a quantum particle by means of a phase-space point  $(x, p)$  but we need a curve,  $g_F(x, p) = 0$ .

The phase space Fermi function  $g_F(x, p)$  and the Wigner function  $W(x, p)$  are at first sight unrelated. On the other hand, for a Gaussian wave packet the  $g_F(x, p) = 0$  curve is an ellipse of area  $\pi\hbar$ . Indeed, given the wave packet (5.140), we have

$$R(x) = \frac{1}{\sqrt{\sqrt{\pi}\delta}} \exp\left(-\frac{(x - x_0)^2}{2\delta^2}\right), \quad S(x) = p_0x, \quad (5.148)$$

and therefore from Eq. (5.147) we obtain

$$g_F(x, p) = \frac{\hbar^2(x - x_0)^2}{\delta^4} + (p - p_0)^2 - \frac{\hbar^2}{\delta^2}. \quad (5.149)$$

It is clear from Eqs. (5.141) and (5.149) that for Gaussian packet we have

$$W(x, p) = \frac{1}{\pi e \hbar} \exp\left(-\frac{\delta^2}{\hbar^2} g_F(x, p)\right). \quad (5.150)$$

Therefore, there is a one to one correspondence between the “equipotential curves”  $g_F(x, p) = K$  and  $W(x, p) = C$ , with  $C = \frac{1}{\pi e \hbar} \exp\left(-\frac{\delta^2}{\hbar^2} K\right)$ . In particular, the  $g_F = 0$  curve coincides with the curve  $W = \frac{1}{\pi e \hbar} = \frac{W_{\max}}{e}$ , with  $W_{\max} = W(x_0, p_0)$  maximum value of  $W$ .

**Exercise 5.17** Compute the Fermi  $g_F$  function for the Gaussian wave packet (5.72) representing the evolution of a free particle.

### 5.6.1 \* General framework for Gaussian states

We are now in the position to describe the basic features of a generic continuous-variable system, made of  $N$  bosonic modes that are associated with the corresponding creation and annihilation operators  $\{a_j, a_j^\dagger\}_{j=1}^N$ . The global Hilbert space  $\mathcal{H} = \otimes_{j=1}^N \mathcal{H}_j$  can be decomposed as a tensor product of single-mode Hilbert spaces  $\mathcal{H}_j$ . Each of them is spanned by the Fock basis  $\{|n_j\rangle\}_{n_j=0}^\infty$  characterizing the eigenstates of the number operator  $n_j = a_j^\dagger a_j$ . Over these states the action of the bosonic operators is well defined, and follows the usual rules of Eq. (5.83) for each mode. Arranging the creation and annihilation bosonic operators in a vectorial form:  $\vec{\mathcal{A}} = [a_1, a_1^\dagger, a_2, a_2^\dagger, \dots, a_N, a_N^\dagger]^T$ , we can express the bosonic commutation relations in a compact form as

$$[\mathcal{A}_j, \mathcal{A}_l] = \Omega_{jl}, \quad (5.151)$$

where we introduced the  $2N \times 2N$  matrix

$$\Omega = \bigoplus_{j=1}^N \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (5.152)$$

satisfying

$$\Omega = -\Omega^T = -\Omega^{-1}. \quad (5.153)$$

The bosonic system can be alternatively described by means of the quadratures  $\{X_j, P_j\}_{j=1}^N$ , which are defined according to Eqs. (5.96) for each mode:

$$X_j = \frac{1}{2}(a_j + a_j^\dagger), \quad P_j = \frac{1}{2i}(a_j - a_j^\dagger). \quad (5.154)$$

Similarly to the creation and annihilation operators, we can arrange the quadratures in a vectorial form:  $\vec{\mathcal{Q}} = [X_1, P_1, X_2, P_2, \dots, X_N, P_N]^T$ , so that the canonical commutation relations  $[X_j, P_l] = \frac{i}{2}\delta_{jl}$  can be expressed according to

$$[\mathcal{Q}_j, \mathcal{Q}_l] = \frac{i}{2}\Omega_{jl}, \quad (5.155)$$

$\Omega_{jl}$  being the matrix elements of  $\Omega$  as defined in (5.152). The quadrature operators can be seen as observables with continuous eigenspectra, whose eigenvalues are used as continuous variables to describe the bosonic system.

A generic quantum state of the continuous-variable system is represented by a density operator  $\rho \in \mathcal{H}$  which can be characterized by its statistical moments with respect to the quadrature variables. Let us define its first moment  $\vec{d}$ , that is, the mean value whose components are given by:

$$d_j \equiv \langle \mathcal{Q}_j \rangle = \text{Tr}(\mathcal{Q}_j \rho), \quad (5.156)$$

and its second moment  $\Sigma$ :

$$\Sigma_{jl} \equiv 2\langle \{\Delta \mathcal{Q}_j, \Delta \mathcal{Q}_l\} \rangle = 2\text{Tr}([\mathcal{Q}_j \mathcal{Q}_l + \mathcal{Q}_l \mathcal{Q}_j] \rho) - 4\text{Tr}(\mathcal{Q}_j \rho) \text{Tr}(\mathcal{Q}_l \rho), \quad (5.157)$$

where  $\Delta \mathcal{Q}_j = \mathcal{Q}_j - d_j$  and  $\{\cdot, \cdot\}$  denotes the anticommutator. The vector  $\vec{d}$  of all the first moments is called the *displacement vector*, while the matrix  $\Sigma$  of all the

second moments is called the *covariance matrix*. Remarkably, Gaussian states (i.e., bosonic states having a Gaussian Wigner representation) can be fully characterized by the first two moments. A pure state is Gaussian if and only if its Wigner function is non-negative.

It is always possible to arbitrarily adjust first moments by means of local unitary operations (displacements in phase space), namely applications of the single-mode operators of Eq. (5.91) in such a way to locally re-center the Wigner function corresponding to each single mode. Since such operations preserve all the informationally relevant properties, such as measures of correlations, unless specified, hereafter we will assume that the states under consideration are characterized by  $\vec{d} = 0$ .

Concerning the covariance matrix  $\Sigma$ , from its definition it is clear that such  $2N \times 2N$  matrix is real and symmetric. Moreover the commutation relations (5.155), together with the requirement that the associated density matrix is positive semidefinite, impose the following constraint:

$$\Sigma + \frac{i}{4}\Omega \geq 0, \quad (5.158)$$

thus implying its positiveness ( $\Sigma > 0$ ) (see Simon *et al.*, 1994). More in general, it can be shown that inequality (5.158) is a necessary and sufficient condition that  $\Sigma$  has to fulfill in order to characterize a physical density matrix of any continuous-variable system. Inequality (5.158) provides the generalization of the uncertainty principle on the canonical operators in the so-called Robertson-Schrödinger form (see Sec. 2.3.4). Indeed, from its block diagonal elements, one can easily derive the usual Heisenberg relation for position and momentum:  $\Delta X_j \Delta P_j \geq \frac{1}{4}$ .

### Gaussian operations

Let us now introduce the most general unitary transformation that describes the evolution of Gaussian states into the same class of Gaussian states. It is easy to demonstrate that such evolution  $U \equiv \exp(-iH)$  is generated via a Hamiltonian  $H$  which is at most quadratic in the bosonic creation and annihilation field operators  $\mathcal{B}_j$ . This means that

$$H = i \left[ \sum_j \alpha_j a_j^\dagger + \sum_{j,l} A_{jl} a_j^\dagger a_l + \sum_{j,l} B_{jl} a_j^\dagger a_l^\dagger - \text{H.c.} \right], \quad (5.159)$$

where  $\vec{\alpha} \in \mathbb{C}^n$ , while  $A$  and  $B$  are  $N \times N$  complex matrices. The Hamiltonian  $H$  can be divided into three blocks,  $H_1$ ,  $H_2$  and  $H_3$ , each of them describing a specific operation whose meaning and action can be described as follows:

- (i) the block

$$H_1 = i \left[ \sum_j (\alpha_j a_j^\dagger - \alpha_j^* a_j) \right] \quad (5.160)$$

generates the so-called *displacement operators*, which have been introduced in Eq. (5.91);

(ii) the block

$$H_2 = i \left[ \sum_{j,l} (A_{jl} a_j^\dagger a_l - A_{j,l}^* a_l^\dagger a_j) \right] \quad (5.161)$$

generates phase rotations (for  $j = l$ ) (phase shifter transformation) and mixing of the various modes (for  $j \neq l$ ). In the case of two bosonic modes, say  $(a, a^\dagger)$  and  $(b, b^\dagger)$ , the latter describes a so-called beam splitter transformation, which is the simplest example of an interferometer, defined by  $B(\theta) = \exp[\theta(a^\dagger b - ab^\dagger)]$ ;

(iii) the remaining block

$$H_3 = i \left[ \sum_{j,l} (B_{jl} a_j^\dagger a_l^\dagger - B_{j,l}^* a_l a_j) \right] \quad (5.162)$$

is responsible for the squeezing of one mode (for  $j = l$ ), as in Eq. (5.119), or of many modes (for  $j \neq l$ ).

Using the Baker-Campbell-Hausdorff expansion of Eq (A.101) and using the canonical commutation relations for the bosonic modes, one can show, that in the Heisenberg picture, the unitary operation generated by Hamiltonian (5.159) corresponds to a linear transformation which maps the bosonic operators  $\{a_j, a_j^\dagger\}_{j=1,\dots,N}$  into the new operators  $\{b_j, b_j^\dagger\}_{j=1,\dots,N}$ . Namely:

$$\vec{b} \equiv U^\dagger \vec{a} U = F \vec{a} + G \vec{a}^\dagger + \vec{\alpha}, \quad (5.163)$$

where we defined the vectors with all the bosonic annihilation operators  $\vec{a} = [a_1, a_2, \dots, a_N]^T$  and  $\vec{b} = [b_1, b_2, \dots, b_N]^T$ .

**Exercise 5.18** By imposing the bosonic commutation relations to the transformed operators  $\{b_j, b_j^\dagger\}_{j=1,\dots,N}$  show that matrices  $F$  and  $G$  of Eq. (5.163) satisfy the constraints:

$$F G^T = G F^T, \quad F F^\dagger - G G^\dagger = I_N. \quad (5.164)$$

The system of equations in (5.163), with  $F$  and  $G$  satisfying (5.164), are known as the Bogoliubov equations for a  $N$ -mode Bosonic system, and generalize Eq. (5.121) to  $N$  modes. Together with the adjoint relations, these can be written in a compact matrix form as:

$$\begin{bmatrix} \vec{b} \\ \vec{b}^\dagger \end{bmatrix} = \begin{bmatrix} F & G \\ G^* & F^* \end{bmatrix} \begin{bmatrix} \vec{a} \\ \vec{a}^\dagger \end{bmatrix} + \begin{bmatrix} \vec{\alpha} \\ \vec{\alpha}^\dagger \end{bmatrix} \quad (5.165)$$

Using Eqs. (5.164), it is easy to invert the homogeneous term of the Bogoliubov equations, by noting that

$$\begin{bmatrix} F & G \\ G^* & F^* \end{bmatrix}^{-1} = \begin{bmatrix} F^\dagger & -G^T \\ -G^\dagger & F^T \end{bmatrix}. \quad (5.166)$$

Looking at the action of the unitary operation  $U$  on the quadrature operators, this can be seen in a more elegant way as an affine map:

$$\vec{\mathcal{Q}}' \equiv U^\dagger \vec{\mathcal{Q}} U = S \vec{\mathcal{Q}} + \vec{\ell}, \quad (5.167)$$

where  $\vec{\ell} \in \mathbb{R}^{2N}$  and  $S$  is a  $2N \times 2N$  real matrix. The conditions (5.164) are equivalent to require that matrix  $S$  is symplectic, that is a matrix which satisfies the equality

$$S \Omega S^T = \Omega, \quad (5.168)$$

with  $\Omega$  being the skew-symmetric matrix (5.152). This is consistent with the fact that the bosonic commutation relations are preserved by the transformation. Finally, it is of fundamental importance to remind that the mapping of the statistical moments  $\vec{d}$  and  $\Sigma$  induced by  $U$ :

$$\vec{d}' \rightarrow \vec{d}' = S \vec{d} + \vec{\ell}, \quad (5.169)$$

$$\Sigma \rightarrow \Sigma' = S \Sigma S^T, \quad (5.170)$$

can completely characterize its action over a generic Gaussian state.

The decomposition of Gaussian transformations over elementary symplectic transformations is dictated by the fact that mathematically they form a group. As we have briefly discussed before, such transformations are physically implementable by means of phase shifters, beam splitters and squeezers. Let us now focus on few paradigmatic examples of non-trivial multi-mode Gaussian states and the corresponding operations on them.

### Thermal decomposition

Thanks to Williamson's theorem (see Sec. A.1.7.2), a generic  $N$ -mode covariance matrix  $\Sigma$  can be transformed into a diagonal form by a symplectic transformation. Namely, there exists a symplectic matrix  $S$  such that

$$\Sigma = S \bigoplus_{j=1}^N \begin{bmatrix} \nu_j & 0 \\ 0 & \nu_j \end{bmatrix} S^T. \quad (5.171)$$

The symplectic spectrum  $\{\nu_j\}_{j=1}^N$  equals the modulus of the  $2N$  real eigenvalues of the matrix  $i\Omega\Sigma$ . We note that the constraint (5.158) implies that  $\nu_j \geq 1$ .

In the space of density operators, the symplectic transformation (5.171) corresponds to a decomposition of an arbitrary Gaussian state into a direct sum of single-mode states with a covariance matrix  $\bar{\Sigma}_j = (2\bar{n}_j + 1)I_2$ , where  $\bar{n}_j = \frac{1}{2}(\nu_j - 1) \geq 0$ . One can show that the number-state representation of each of such states is

$$\rho_{\text{th}}(\bar{n}_j) \equiv \frac{e^{-\beta_j a^\dagger a}}{\text{Tr}[e^{-\beta_j a^\dagger a}]} = \frac{1}{1 + \bar{n}_j} \sum_{m=0}^{+\infty} \left( \frac{\bar{n}_j}{1 + \bar{n}_j} \right)^m |m\rangle\langle m|. \quad (5.172)$$

To obtain a physical interpretation of this result, note that  $\rho_{\text{th}}(\bar{n}_j)$  can be seen as a single bosonic mode at thermal equilibrium, at temperature  $T = (k_B\beta)^{-1}$  and angular frequency  $\omega$  (where  $k_B$  denotes the Boltzmann constant), such that  $\bar{n} = (e^{\beta\hbar\omega} - 1)^{-1}$  is the average number of particles according to the Bose-Einstein

statistics. In fact, as we shall see later in Sec. 6.10.1, the significance of  $\rho_{\text{th}}(\bar{n})$  resides in the observation that it maximizes the so-called von Neumann entropy functional  $S(\rho) = -\text{Tr}(\rho \log \rho)$  at a fixed number of bosons  $\bar{n} = \text{Tr}(a^\dagger a \rho)$ .

In passing we also note that the state  $\rho_{\text{th}}(0)$  denotes the vacuum state and has a covariance matrix which coincides with the identity:  $\Sigma = I_2$ . Therefore, using Eq. (5.157) for its diagonal elements  $\Sigma_{11} = \Sigma_{22} = 1$ , we find  $\Delta X^2 = \Delta P^2 = \frac{1}{4}$ . This condition (also called the vacuum noise) provides the minimum uncertainty point, according to the Heisenberg relation  $\Delta X \Delta P \geq \frac{1}{4}$ , which is reached symmetrically by the two quadratures.

### Two-mode Gaussian states

Gaussian states of two bosonic modes ( $N = 2$ ) represent the simplest states where properties such as classical and quantum correlations can be addressed, as we will see later in Sec. 6.10. The generic covariance matrix of such states can be written as

$$\Sigma = \begin{bmatrix} \Sigma_A & \Upsilon_{AB} \\ \Upsilon_{AB}^T & \Sigma_B \end{bmatrix}, \quad (5.173)$$

where  $\Sigma_A = \Sigma_A^T$  and  $\Sigma_B = \Sigma_B^T$  respectively are the covariance matrices of the two partial systems corresponding to the first mode ( $A$ ) and the second mode ( $B$ ), while  $\Upsilon_{AB}$  describes the correlations between the two modes. All of them are  $2 \times 2$  real matrices. Let us now define the following invariants under symplectic transformations:  $A = \det(\Sigma_A)$ ,  $B = \det(\Sigma_B)$ ,  $C = \det(\Upsilon_{AB})$ , and  $D = \det(\Sigma)$ . In order to have a physical state, one must require  $A, B \geq 1$ . According to Williamson's decomposition, the symplectic eigenvalues  $\nu_\pm$  of  $\Sigma$  are (see Serafini *et al.*, 2004):

$$\nu_\pm^2 = \frac{1}{2} \left( \Delta \pm \sqrt{\Delta^2 - 4D} \right), \quad \text{with } \Delta = A + B + 2C. \quad (5.174)$$

Moreover, it can be shown that the constraint (5.158) here translates into the conditions:  $\det(\Sigma) \geq 1$  and  $\Delta \leq 1 + \det(\Sigma)$ .

## 5.7 Quantum cryptography with continuous variables

Although quantum cryptography was developed for qubits, the use of continuous variables presents important advantages with respect to qubit-based protocols. These advantages are mainly of practical nature, since quantum cryptography with continuous variables requires only standard quantum communication technology. In what follows, we discuss the basic principles of continuous-variable quantum key distribution (CVQKD), focusing in particular on the main differences with discrete-variable quantum key distribution (DVQKD).

DVQKD is based on single-photon Fock states, emitted on demand. Unfortunately, these states are difficult to realize experimentally. However, single-photon Fock states can be approximated by means of *faint laser pulses*: a laser outputs a coherent state, given by Eq. (5.85), and this state is attenuated to a very low mean

photon number  $\bar{n} \ll 1$ . For instance, if  $\bar{n} = 0.1$ , the coherent state (5.85) reads  $|\alpha\rangle \approx 0.95|0\rangle + 0.30|1\rangle + 0.07|2\rangle + \dots$ , where  $\alpha = \sqrt{\bar{n}} = \sqrt{0.1}$ . This means that most pulses are empty: the probability that the attenuated coherent state contains no photons is  $p_0 \approx (0.95)^2 \approx 0.90 \approx 1 - \bar{n}$ . A single photon is found with probability  $p_1 \approx (0.30)^2 = 0.09$ . Therefore, the probability of having a nonempty pulse is  $p(n > 0) = 1 - p_0$  and the probability that a nonempty pulse contains more than one photon is  $p(n > 1 | n > 0) = p(n > 1)/p(n > 0) = (1 - p_0 - p_1)/(1 - p_0) \approx \bar{n}/2 = 0.05$  (this means that 5% of the nonempty pulses contain more than one photon). This strategy of course reduces the transmission rate. Moreover the photodetectors must be active for all pulses, including the empty ones. Therefore, the problem of dark counts (that is, there is a click in the detector without an arriving photon) becomes more important when  $\bar{n}$  is small.

CVQKD uses coherent or squeezed states. Coherent states are obtained by means of standard laser pulses, while squeezed states are generated when a nonlinear crystal is pumped by a laser field.

In DVQKD information is encoded on properties of single photons. For instance, a natural way to code the four states of the BB84 protocol is to use photons polarized at  $-45^\circ$ ,  $0^\circ$ ,  $+45^\circ$  and  $90^\circ$ . If, for instance, a photon polarized at  $+45^\circ$  is sent (by Alice) and the measurement takes place in the diagonal basis, then the outcome is deterministic. On the other hand, if the receiver (Bob) chooses the horizontal-vertical basis, he obtains randomly one of the two possible outcomes. In CVQKD, information is carried by properties of light that are continuous, such as the values of the quadratures  $X_1$  and  $X_2$  of a coherent state. It is the fact that  $X_1$  and  $X_2$  are a pair of non-commuting observables that allows intrusion detection, similarly to DVQKD. Many protocols for continuous variables have been proposed in the literature, which differ in the choice of the states that are prepared: single-mode coherent or squeezed states, two-mode squeezed states. For instance, in some protocols Alice prepares and sends to Bob a large number of coherent states  $|\alpha_1\rangle, \dots, |\alpha_N\rangle$ , where  $\alpha_i$  are independent and identically distributed complex Gaussian variables with zero mean and a given variance. Other protocols are based on a discrete encoding of Gaussian states. For instance, an alphabet of  $n$  coherent states  $|\alpha_k\rangle = |ae^{i2\pi k/n}\rangle$  can be used, where  $k$  encodes the secret key (such states have fixed amplitude  $a$  and relative phases  $2\pi k/n$ ).

State measurement (by the receiver Bob) in CVQKD is completely different from what we have discussed for DVQKD. In DVQKD single photons are detected (polarization measurements), while in CVQKD information is encoded both in the amplitude and in the phase of the transmitted (coherent or squeezed) states. Since only relative phases have physical meaning, Alice and Bob must exchange synchronization signals to set a common reference frame for phases. Moreover, direct photon count experiments by means of photodetectors only provide information on the

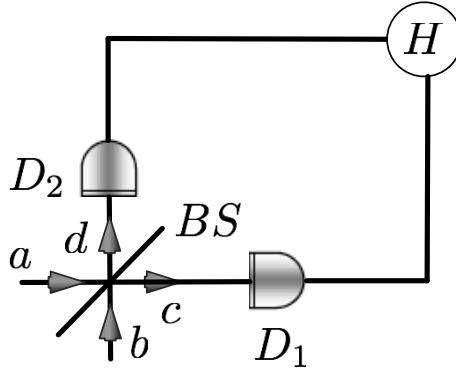


Fig. 5.13 Schematic drawing of balanced homodyne detection:  $a$  is the input mode,  $b$  the local oscillator mode,  $c$  and  $d$  the output modes,  $BS$  is a 50:50 beam splitter,  $D_1$  and  $D_2$  are two photodetectors, and  $H = c^\dagger c - d^\dagger d$  the operator which determines the output homodyne signal.

amplitude of the field. Method such as *homodyne detection* and *double homodyne detection* are instead used.<sup>7</sup>

In homodyne detection the signal mode  $a$  interferes with a second mode  $b$  (the local oscillator) at the same frequency, excited in a coherent state  $|\beta_l\rangle$ , with amplitude  $|\beta_l|$  and phase  $\phi_l$ . The two modes interfere at a beam splitter of transmittivity  $T$  and reflectivity  $R$  ( $T + R = 1$ ), see Fig. 5.13. Denoting  $c$  and  $d$  the output modes reaching photodetectors  $D_1$  and  $D_2$ , we have

$$\begin{aligned} c &= \sqrt{T}a + i\sqrt{1-T}b, \\ d &= i\sqrt{1-T}a + \sqrt{T}b, \end{aligned} \quad (5.175)$$

where the factor  $i = e^{i\frac{\pi}{2}}$  stands for a  $\frac{\pi}{2}$  phase shift between the reflected and the transmitted waves. The signal measured by the two photodetectors are determined by the operators  $c^\dagger c$  and  $d^\dagger d$ : the mean number of photons measured at modes  $c$  and  $d$  are  $\langle c^\dagger c \rangle$  and  $\langle d^\dagger d \rangle$ , respectively. We consider *balanced* homodyne detection. In this scheme, a 50:50 beam splitter ( $T = \frac{1}{2}$ ) is used and the difference of the two photodetector measurements is obtained. The measured homodyne output signal is then obtained using Eq. (5.175):

$$\langle H \rangle = \langle c^\dagger c - d^\dagger d \rangle = i\langle a^\dagger b - b^\dagger a \rangle = 2|\beta_l|\langle X_{\phi_l+\frac{\pi}{2}} \rangle, \quad (5.176)$$

where to derive the last equality we have used the definition of generalized quadratures, Eq. (5.125). To measure simultaneously both quadratures, the input signal must be splitted and sent to two homodyne detectors (double homodyne). In CVQKD, depending on the protocol (homodyne or double homodyne), Bob for each input signal measures either a randomly chosen quadrature ( $X_1$  or  $X_2$ ) or both quadratures. In the first case he informs Alice of his choice and stores a single

<sup>7</sup>Note that double homodyne detection is often referred to as *heterodyne detection*. However, we prefer to avoid this terminology since in radio technology heterodyne detection refers to the case where the frequency of the local oscillator is different from the frequency of the signal.

measurement outcome. In the latter case, for each received signal Bob stores two measurement outcomes (both quadratures). In both cases, Alice and Bob compare a part of their quadrature values and then, similarly to DVQKD, they perform information reconciliation and privacy amplification to extract the final secret key.

## 5.8 A guide to the bibliography

The no-cloning theorem is due to Dieks (1982) and to Wootters and Zurek (1982). The no-signalling condition is discussed in Ghirardi (2013). The no-cloning theorem does not forbid the existence of imperfect cloning machines, discussed, for instance, in Bužek and Hillery (1996), Gisin and Massar (1997) and Bruß *et al.* (1998). The (imperfect) universal-NOT gate is discussed in Bužek *et al.* (1999), for its experimental realization see De Martini *et al.* (2002).

A book on classical cryptography is Welsh (1997). The quantum cryptographic protocols discussed in Sec. 5.3 are due to Bennett and Brassard (1984) and Ekert (1991), see also Bennett *et al.* (1992). Another interesting protocol using non-orthogonal quantum states was introduced by Bennett (1992). A review of quantum cryptography is Gisin *et al.* (2002). Very readable introductions are Bennett *et al.* (1991), Bruß and Lütkenhaus (2000) and Lomonaco (2001).

Quantum teleportation was discovered by Bennett *et al.* (1993), for conclusive teleportation see Brassard *et al.* (2004). The quantum dense coding protocol is due to Bennett and Wiesner (1992).

Quantum mechanics with continuous variable systems is discussed in standard textbooks such as Cohen-Tannoudji *et al.* (1977). For coherent and squeezed states see quantum optics textbooks, such as Leonhardt (1997) and Scully and Zubairy (1997). For the Wigner and other phase-space representations of quantum mechanics see Schleich (2001) and Zachos *et al.* (2005). The Fermi  $g_F$  function was introduced by Fermi (1930); in the discussion of the  $g_F$  for Gaussian wave packets we followed Benenti and Strini (2009a).

Reviews on quantum information with continuous variables are Braunstein and van Loock (2005), Weedbrook *et al.* (2012) and Holevo (2011).

## Chapter 6

# Entanglement and non-classical correlations

The presence of correlations shared among the subparts of a given physical system, and which do not have a classical counterpart, is a clear signature of non-classicality for a quantum state. As we saw in Chap. 2, such correlations can be captured by entanglement, one of the most intriguing features predicted by quantum theory, which points directly to a violation of the Bell's inequalities. In this chapter we provide a rigorous definition of entanglement, and review its basic properties together with the general framework that is needed to quantify it for bipartite states. We present the common criteria that are used to detect whether a state is entangled or not, and discuss the most common entanglement quantifiers in the bipartite context. The latter follow an introduction of the cornerstone concepts of classical Shannon entropy and quantum von Neumann entropy. Subsequently, the multipartite scenario is briefly addressed. Although entanglement is the most remarkable among non-classical correlations, even some separable (not entangled) mixed states can exhibit a non-classical behaviour. In this wider context we discuss the quantum discord, which is expressed as the difference between two entropic quantities and is extractable by performing a local measure on one of the subsystems. This, and other related quantifiers, enlighten the impossibility of recovering the information contained in a composite quantum system after the measurement process. A special-topic section on the different definitions of entropy used in physics closes the chapter.

### 6.1 Definition of entanglement

The concept of entanglement, first introduced in Chap. 2, arises from the necessity to distinguish the quantum correlations that can occur in many-party quantum states, from the classical ones. While such distinction has been the source of an intense debate for many years, only recently the modern theory of quantum information has developed a rigorous way to define the amount of classical correlations. The key to this formalism relies on the concept of the so-called LOCC (operations); that is, *local operations* possibly supplemented by *classical communication*. Local operations are unitary transformations or (generalized) measurements performed by Alice or Bob

on their members of the shared quantum state. Classical communication enables Alice and Bob to share the results of the local quantum operations, in order to select the successful, maximally entangled cases. We define classical correlations as those that can be generated by LOCC. Any other correlations that cannot be simulated classically are associated to quantum effects, and labelled as quantum correlations.

Specifically, suppose to be able to process a given quantum state by LOCC. Then, if the processed state can be used for some task that cannot be simulated by classical correlations, such as violating a Bell's inequality, we must not attribute these effects to the LOCC processing that we have performed, but to quantum correlations that were already present in the initial state. This poses the basis for an operative definition of entanglement, which can be seen as “the correlations that may not be created by LOCC alone”.

### 6.1.1 Basic properties

Given the above operational definition of entanglement, it is possible to assess some general statements which should always hold, irrespective of the specific invoked measure.

- (i) *A separable state contains no entanglement:* We already defined the separable pure states of a generic bipartite system as those which are written in the form of Eq. (2.59). Specifically, a pure state  $|\psi\rangle_{AB}$  of a bipartite system is separable if and only if it can be written as

$$|\psi\rangle_{AB} = |\alpha\rangle_A \otimes |\beta\rangle_B, \quad (6.1)$$

with states  $|\alpha\rangle_A$  and  $|\beta\rangle_B$  describing the components of the two subsystems, say Alice's ( $A$ ) and Bob's ( $B$ ) one. It is easy to generalize such notion to mixed states. Suppose indeed that Alice and Bob agree over the phone on the local preparation of their two subsystems. Therefore, a bipartite mixed state is separable if and only if it can be written as

$$\rho_{AB} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}, \quad \text{with } p_k \geq 0 \text{ and } \sum_k p_k = 1, \quad (6.2)$$

$\rho_A^{(k)}$  and  $\rho_B^{(k)}$  being density matrices for the corresponding two subsystems.

The concept of separability can be stated in the general framework of a  $n$ -partite system, made of parties  $A, B, C, \dots$ , according to the following:

$$\rho_{ABC\dots}^{\text{sep}} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)} \otimes \rho_C^{(k)} \otimes \dots \quad (6.3)$$

These states can be generated by means of LOCC only: the party  $A$  samples from the probability distribution  $\{p_1, p_2, \dots\}$ , informs all the other parties of the  $k$ -th outcome, and then each party  $X$  locally creates  $\rho_X^{(k)}$ . This implies that all the correlations contained in Eq. (6.3) can be described classically, therefore the state is not entangled. A separable system of the form (6.3) always satisfies Bell's inequalities.

- (ii) *All non separable states are entangled:* Given that separable states are those and only those that may be created using LOCC alone, it is clear that a quantum state cannot be generated by LOCC alone if and only if it is entangled. This amounts to say that terms non-separable and entangled are synonymous.
- (iii) *Entanglement cannot increase under LOCC:* If a quantum state  $\rho$  can be transformed into another quantum state  $\sigma$  with LOCC, then anything doable with  $\sigma$  and LOCC can also be achieved with  $\rho$  and LOCC (that is, the amount of violation of the Bell's inequalities for  $\sigma$  cannot be greater than that for  $\rho$ ). This also implies that any two states that are related by local unitaries must have an equal amount of entanglement.
- (iv) *There are bipartite maximally entangled states:* Since a given state can be more entangled than another, it is reasonable to ask whether there is a state that is more entangled than all the others. While for multipartite states the issue is still debated (see Sec. 6.8), for bipartite states the answer is positive: all maximally entangled states are unitarily equivalent to the generalized EPR pair

$$|\phi_d^+\rangle_{AB} = \frac{1}{\sqrt{d}} (|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B + \dots + |d-1\rangle_A \otimes |d-1\rangle_B), \quad (6.4)$$

where  $d$  is the dimension of each of the two subsystems. Indeed, as will be detailed in Sec. 6.5, any state of a bipartite system can be synthesized from such EPR states using only LOCC.

Having stated the basic requirements for the entanglement, one would be tempted to build up a more formal derivation of what is a measure of entanglement and how it could be quantified. However before focusing on this issue, we will address a much simpler question: Given a generic quantum state, pure or mixed, is it separable or entangled? We will frame all our discussion in the much simpler, yet very instructive, case of bipartite systems, and often refer to the two parties with the usual names of Alice (A) and Bob (B). The multipartite situation is much more involved and requires a separate treatment; we will postpone it to Sec. 6.8.

## 6.2 Bipartite separability criteria

The formal answer to the previous question is actually straightforward: a generic bipartite state is said to be separable if and only if can be written in the form of Eq. (6.2). This means that Alice and Bob can prepare it in a “classical” manner; that is, by means of LOCC.

However, given a density matrix  $\rho_{AB}$ , it is in general a non-trivial task to operatively prove whether a decomposition of the form (6.2) exists or not, unless it reduces to a global pure state  $|\psi\rangle_{AB}$ . Indeed in the latter case it is sufficient to perform the Schmidt decomposition of the state, and to extract its Schmidt number. Separability is a necessary and sufficient condition for the state  $|\psi\rangle_{AB}$  to have a

Schmidt number equal to one, as already outlined in Sec. 2.7. We thus need separability criteria for general mixed states  $\rho_{AB}$  that are easy to test. Several criteria have been proposed; let us introduce the most commonly employed ones.

### 6.2.1 The Peres separability criterion

The Peres criterion provides a necessary condition for the existence of decomposition (6.2), in other words, a violation of this criterion is a sufficient condition for entanglement. This criterion is based on the *partial transpose* operation. Introducing an orthonormal basis  $\{|i\rangle_A|\alpha\rangle_B\}$  in the Hilbert space  $\mathcal{H}_{AB}$  associated with the bipartite system  $A + B$ , the density matrix  $\rho_{AB}$  has matrix elements  $(\rho_{AB})_{i\alpha;j\beta} = {}_A\langle i| \rho_{AB} |j\rangle_A |\beta\rangle_B$ . The partial transpose density matrix is constructed by only taking the transpose in either the Latin or Greek indices (Latin indices refer to Alice's subsystem and Greek indices to Bob's). For instance, the partial transpose with respect to Alice is given by

$$(\rho_{AB}^{T_A})_{i\alpha;j\beta} = (\rho_{AB})_{j\alpha;i\beta}. \quad (6.5)$$

Since a separable state  $\rho_{AB}$  can always be written in the form (6.2) and the density matrices  $\rho_A^{(k)}$  and  $\rho_B^{(k)}$  have non-negative eigenvalues, then the overall density matrix  $\rho_{AB}$  also has non-negative eigenvalues. The partial transpose of a separable state reads

$$\rho_{AB}^{T_A} = \sum_k p_k [\rho_A^{(k)}]^T \otimes \rho_B^{(k)}. \quad (6.6)$$

Since the transpose matrices  $[\rho_A^{(k)}]^T = [\rho_A^{(k)}]^\star$  are Hermitian non-negative matrices with unit trace, they are also legitimate density matrices for Alice. It follows that none of the eigenvalues of  $\rho_{AB}^{T_A}$  is non-negative. This is a necessary condition for decomposition (6.2) to hold. It is then sufficient to have at least one negative eigenvalue of  $\rho_{AB}^{T_A}$  to conclude that the state  $\rho_{AB}$  is entangled.

As an example, we consider the so-called Werner state

$$(\rho_W)_{AB} = \frac{1}{4}(1-q)I + q|\psi^-\rangle\langle\psi^-|, \quad (6.7)$$

where  $0 \leq q \leq 1$ ,  $I$  is the identity in the Hilbert space  $\mathcal{H}_{AB}$  and  $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$  is a state of the Bell basis (see Sec. 3.5.1). In the basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  the density matrix  $(\rho_W)_{AB}$  reads

$$(\rho_W)_{AB} = \begin{bmatrix} \frac{1-q}{4} & 0 & 0 & 0 \\ 0 & \frac{1+q}{4} & -\frac{q}{2} & 0 \\ 0 & -\frac{q}{2} & \frac{1+q}{4} & 0 \\ 0 & 0 & 0 & \frac{1-q}{4} \end{bmatrix}. \quad (6.8)$$

Taking the partial transpose yields

$$(\rho_W)^{T_A}_{AB} = \begin{bmatrix} \frac{1-q}{4} & 0 & 0 & -\frac{q}{2} \\ 0 & \frac{1+q}{4} & 0 & 0 \\ 0 & 0 & \frac{1+q}{4} & 0 \\ -\frac{q}{2} & 0 & 0 & \frac{1-q}{4} \end{bmatrix}. \quad (6.9)$$

This latter matrix has eigenvalues  $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1+q}{4}$  and  $\lambda_4 = \frac{1-3q}{4}$ . As  $\lambda_4 < 0$  for  $\frac{1}{3} < q \leq 1$ , we may conclude that the Werner state is entangled for these values of the parameter  $q$ .

It can be shown (M. Horodecki *et al.*, 1996) that for composite states of dimension  $2 \times 2$  and  $2 \times 3$ , the Peres criterion provides a necessary and sufficient condition for separability; that is, the state  $\rho_{AB}$  is separable if and only if  $\rho_{AB}^{T_A}$  is non-negative. This result teaches us, for instance, that the Werner state is separable for  $0 \leq q \leq \frac{1}{3}$ . However, for higher dimensional systems, states exist for which all eigenvalues of the partial transpose are non-negative, but that are non-separable (P. Horodecki, 1997). These states are known as *bound entangled states* since they cannot be distilled by means of LOCC to form a maximally entangled state (M. Horodecki *et al.*, 1998).

We stress that the Peres criterion is more sensitive than Bell's inequality for detecting quantum entanglement; that is, there are states detected as entangled by the Peres criterion that do not violate Bell's inequalities (Peres, 1996).

**Exercise 6.1** Show that for a separable two-qubit state  $\rho_{AB}$  the following inequality is satisfied:

$$\langle(\Delta\Sigma_x)^2\rangle + \langle(\Delta\Sigma_y)^2\rangle + \langle(\Delta\Sigma_z)^2\rangle \geq 4, \quad (6.10)$$

where  $\Sigma_i = \sigma_A^i \otimes I_B + I_A \otimes \sigma_B^i$ ,  $\langle(\Delta\Sigma_i)^2\rangle = \langle(\Sigma_i)^2\rangle - \langle\Sigma_i\rangle^2$  ( $i = x, y, z$ ) and the angle brackets denote the expectation value over  $\rho_{AB}$ . Show that this criterion allows us to conclude that the Werner state (6.7) is entangled when  $\frac{1}{3} < q \leq 1$ .

### 6.2.2 Positive maps

The Peres criterion is equivalent to demand the positivity of the operator  $T_A \otimes I_B$  applied on the matrix  $\rho_{AB}$ , where  $T_A$  denotes the transposition operation relative to Alice's subsystem and  $I_B$  the identity relative to Bob's subsystem [the matrix elements of the transformed state  $\rho'_{AB} = (T_A \otimes I_B)[\rho_{AB}]$  are defined in Eq. (6.5)]. More formally, a generic map  $\Theta_A$  acting on the Hilbert space of Alice is *positive* if and only if it maps any non-negative operator  $\rho_A$  into a non-negative one:

$$\rho'_A = \Theta_A[\rho_A] \geq 0 \quad (\text{positivity}), \quad (6.11)$$

meaning that the eigenvalues of the matrix  $\rho'_A$  are non negative. Moreover the map  $\Theta_A$  is said to be *completely positive* if and only if the following stronger condition is satisfied, for the identity map  $I_B$  acting on any finite-dimensional system  $B$ :

$$\rho'_{AB} = (\Theta_A \otimes I_B)[\rho_{AB}] \geq 0 \quad (\text{complete positivity}), \quad (6.12)$$

meaning that the eigenvalues of the matrix  $\rho'_{AB}$  are non negative. The example above on the Werner state (6.7) shows that the partial transposition  $T_A$  is positive, but not completely positive.

Having all these notions at hand, it is possible to define a more powerful necessary and sufficient condition for separability. The state  $\rho_{AB}$  is separable if and only if, for any positive but not completely positive map  $\Theta_A$ , we have:

$$(\Theta_A \otimes I_B)[\rho_{AB}] \geq 0, \quad (6.13)$$

meaning that the eigenvalues of the matrix  $[\Theta_A \otimes I_B]\rho_{AB}$  are non negative (M. Horodecki *et al.*, 1996). Note that unfortunately the stronger criterion of positive maps is non-operational, in the sense that we do not have a complete characterization of the set of all positive (and not completely positive) maps  $\Theta_A$ .

### 6.2.3 Entanglement witnesses

Entanglement witnesses are tools that can be used to try to determine whether a state is separable or not. Specifically, a state  $\rho_{AB}$  is entangled if and only if there exists an Hermitian operator  $W$  satisfying the following requirements:

$$(i) \quad \text{Tr}(W\rho_{AB}) < 0, \quad (6.14)$$

$$(ii) \quad \text{Tr}(W\rho_{AB}^{\text{sep}}) \geq 0 \quad \forall \rho_{AB}^{\text{sep}} \text{ separable state}. \quad (6.15)$$

The observables  $W$  satisfying the conditions in Eqs. (6.14)–(6.15) for a given state  $\rho_{AB}$  are called *entanglement witness*, because they are able to “detect” the entanglement contained in  $\rho_{AB}$  (Terhal, 2000). The central interest of entanglement witnesses is that, for any entangled state, there exists a witness that detects it.

A simple example of entanglement witness is the Hermitian swap operator

$$V = \sum_{i,j=0}^{d-1} |i\rangle_{AA}\langle j| \otimes |i\rangle_{BB}\langle j|, \quad (6.16)$$

where  $d$  is the dimension of the local Hilbert space of Alice and of Bob. Indeed it is easily verified that, for a pure product state, it is positive:

$$({}_A\langle\psi| \otimes {}_B\langle\phi|)V(|\psi\rangle_A \otimes |\phi\rangle_B) = {}_A\langle\psi|\phi\rangle_A{}_B\langle\phi|\psi\rangle_B = |\langle\psi|\phi\rangle|^2 \geq 0. \quad (6.17)$$

This fact ensures that, for a generic mixed separable state  $\rho_{AB}^{\text{sep}}$  of the form (6.2), the swap must be non negative and the condition (ii) in Eq. (6.15) is satisfied. Moreover we have that  $V = P^{(+)} - P^{(-)}$ , where  $P^{(\pm)} = \frac{1}{2}(I \pm V)$  respectively correspond to the projectors onto the symmetric and the antisymmetric subspaces of the global Hilbert space. It is thus easy to find states that correspond to negative eigenvalues of  $V$ . For example, take the antisymmetric Bell state  $|\psi^-\rangle$ :

$$\text{Tr}(V|\psi^-\rangle\langle\psi^-|) = \frac{1}{2}(|01\rangle - |10\rangle)V(|01\rangle - |10\rangle) = -1, \quad (6.18)$$

therefore  $V$  also satisfies the requirement (i) in Eq. (6.14).

#### Geometric representation

The usefulness of entanglement witnesses as a criterion to detect the entanglement can be appreciated by adopting a geometric approach. Here we will not provide a formal argument, but rather give an intuitive explanation that allows to understand why entanglement witnesses work.

A generic density matrix  $\rho$  can be seen as a vector  $|\rho\rangle$  in a vector space that is referred to as the Hilbert-Schmidt space. In order to have a convenient geometric interpretation of this vector space, we define the scalar product on this space:

$$\langle\rho|\sigma\rangle = \text{Tr}(\rho^\dagger\sigma). \quad (6.19)$$

Now we observe that, given two arbitrary separable states  $\rho_1^{\text{sep}}$  and  $\rho_2^{\text{sep}}$ , any convex combination of such states is again separable; that is,

$$\rho_3^{\text{sep}} = \lambda \rho_1^{\text{sep}} + (1 - \lambda) \rho_2^{\text{sep}} \quad \text{with } 0 \leq \lambda \leq 1. \quad (6.20)$$

From a geometric point of view, this means that any two points belonging to the (convex) set  $\mathcal{S}$  of separable states are connected by a straight line which is entirely contained into  $\mathcal{S}$ . This situation can be pictorially shown as in Fig. 6.1, where the shaded shape  $\mathcal{A}_1$  represents a convex set such as  $\mathcal{S}$  (namely it has no trough), while the shaded shape  $\mathcal{A}_2$  is not convex (it has a trough on its upper right side). Convexity implies that it is always possible to find lines separating the shaded area from their surrounding. Consequently, for any point  $P_1$  lying outside the convex set  $\mathcal{S}$ , one can find a straight line  $W$  that separates it from the shaded area. While, for the shape on the left, this is the case, for the shape on the right this is not always possible (see, e.g., the point  $P_2$  inside the trough). Although the set of separable states has a higher dimension and is more complicated than the shapes depicted in Fig. 6.1, the two-dimensional sketch is sufficient to grasp the basic mechanism of entanglement witnesses.

A separable state  $\rho^{\text{sep}}$  is characterized by the requirement  $\text{Tr}(W\rho^{\text{sep}}) \geq 0$ . The condition that  $\text{Tr}(W\rho)$  vanishes requires  $\rho$  to be a linear combination of operators  $\mathcal{O}_i$  that are orthogonal to  $W$ :

$$\rho = \sum_i c_i \mathcal{O}_i, \quad \text{with } \text{Tr}(W\mathcal{O}_i) = 0. \quad (6.21)$$

This amounts to say that the equation  $\text{Tr}(W\rho) = 0$  defines an hyperplane in the space of the operators analogous to the straight line drawn in Fig. 6.1 (left). The sign of  $\text{Tr}(W\rho)$  indicates on which side of the hyperplane  $\rho$  is situated, therefore condition (ii) in Eq. (6.15) says that all separable states are situated on one side of

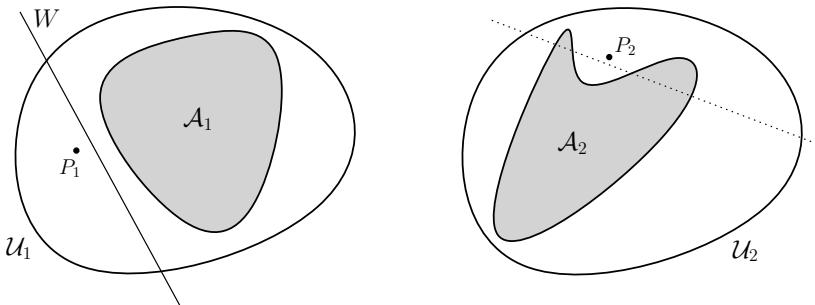


Fig. 6.1 Intuitive picture of separable states and of entanglement witnesses. The two shaded shapes represent a convex ( $\mathcal{A}_1$ ) and a non-convex ( $\mathcal{A}_2$ ) set living inside the universe ( $\mathcal{U}_1$  or  $\mathcal{U}_2$ ). The situation of the Hilbert-Schmidt space for density matrices is analogous to the left sketch: the set of separable states, here represented by  $\mathcal{A}_1$ , is convex; there exists a witness (the analogue of the line  $W$ ) to any entangled state (the analogue of the point  $P_1 \in \mathcal{U}_1 \setminus \mathcal{A}_1$ ), that separates it from the set of separable states. The situation in the right sketch is unphysical, indeed the point  $P_2$  can be obtained as a convex combination of two points inside  $\mathcal{A}_2$ , but lies outside it.

this hyperplane. Since the separable states form a convex set, there is a witness to any entangled state that detects it, just like there is a line to any point outside  $\mathcal{A}_1$  that separates it from the gray shaded area.

The main purpose of the geometric approach consists in finding a way to optimize the witness operator or to replace the hyperplane by a curved manifold, tangent to the set of separable states, as explained in Lewenstein *et al.* (2000).

#### 6.2.4 Positive maps and witnesses

The concept of positive maps and of entanglement witnesses as criteria to detect entanglement are closely related. Let us indeed consider a positive map  $\Theta_A$  which is not completely positive. This means that, for some state  $\rho_{AB}$ , the extension  $(\Theta_A \otimes I_B)[\rho_{AB}]$  has an eigenvector  $|\chi\rangle$  associated to a negative eigenvalue  $\lambda < 0$ . We now show that the observable

$$W = (\Theta_A^\dagger \otimes I_B)|\chi\rangle\langle\chi| \quad (6.22)$$

is an entanglement witness. Indeed, for any pure separable state  $|\psi^{\text{sep}}\rangle$  one has

$$\text{Tr}\left[\left((\Theta_A^\dagger \otimes I_B)|\chi\rangle\langle\chi|\right)|\psi^{\text{sep}}\rangle\langle\psi^{\text{sep}}|\right] = \text{Tr}\left[|\chi\rangle\langle\chi|\left((\Theta_A \otimes I_B)|\psi^{\text{sep}}\rangle\langle\psi^{\text{sep}}|\right)\right] \geq 0, \quad (6.23)$$

where the inequality follows from the positivity of  $\Theta_A$  and the fact that  $|\psi^{\text{sep}}\rangle$  is a separable state. Moreover the witness (6.22) detects  $\rho_{AB}$  to be entangled:

$$\text{Tr}\left[\left((\Theta_A^\dagger \otimes I_B)|\chi\rangle\langle\chi|\right)[\rho_{AB}]\right] = \text{Tr}\left[\left((\Theta_A \otimes I_B)[\rho_{AB}]\right)|\chi\rangle\langle\chi|\right] = \lambda < 0, \quad (6.24)$$

because of the above eigenvector relation.

We are now close to give a quantitative definition of entanglement for a generic bipartite quantum state, going beyond its basic properties outlined in Sec. 6.1. Our final goal is to answer the following questions: What is a measure of entanglement? How can entanglement be quantified in practice? Before going into the details of such topic, we first need to introduce two entropic measures that lie at the basis of the theory of information. Namely, the classical Shannon entropy and the quantum von Neumann entropy. The von Neumann entropy is the appropriate measure of quantum information, just as the Shannon entropy is for classical information. In particular, we will show that the von Neumann entropy is able to quantify the qualitative observation that an entangled state provides more information about the total system than about subsystems.

### 6.3 The Shannon entropy

The first basic task of classical information theory is to quantify the information contained in a message. This problem was solved by Shannon in (Shannon, 1948). A message is a string of letters chosen from an alphabet  $\mathcal{A} = \{a_1, a_2, \dots, a_k\}$ . We

assume that the letters in the message are statistically independent and that the letter  $a_i$  occurs with *a priori* probability  $p_i$ , where  $\sum_{i=1}^k p_i = 1$ . The assumption that the letters are statistically independent has been made to simplify the discussion. In practice, this is not the case in many important examples. For instance, there are strong correlations between consecutive letters in an English text. However, the ideas developed in this section can be extended to include more complicated situations with correlations. Thus, in what follows statistical independence of the letters will always be assumed and it should not be forgotten that the case of a real language (such as English) is somewhat different.

The Shannon entropy associated with the probability distribution  $\{p_1, p_2, \dots, p_k\}$  is defined by

$$H(p_1, p_2, \dots, p_k) \equiv - \sum_{i=1}^k p_i \log p_i. \quad (6.25)$$

Note that, here as in the rest of this book, all the logarithms are base 2 unless otherwise indicated. We shall show that the Shannon entropy quantifies how much information we gain, on average, when we learn the value of a letter of the message. Let us consider the special case  $k = 2$  and define  $p_1 = p$  (where  $0 \leq p \leq 1$ ). Since  $p_2 = 1 - p$ , the Shannon *binary* entropy is a function of  $p$  alone and we can write

$$H_{\text{bin}}(p) \equiv H(p_1, p_2) = -p \log p - (1-p) \log(1-p). \quad (6.26)$$

In the following we shall simply write  $H(p)$  instead of  $H_{\text{bin}}(p)$ . The Shannon binary entropy  $H(p)$  is plotted in Fig. 6.2: it is equal to zero when  $p = 0$  or  $p = 1$  and attains its maximum value  $H = 1$  when  $p = \frac{1}{2}$ . This is consistent with our interpretation of  $H(p)$  as the average information content of each letter in the message. Indeed, information is a measure of our *a priori ignorance*. If we already know that we shall receive the letter  $a_1$  with certainty ( $p = 1$ ), then no information is gained from the reception of this letter. The same conclusion holds when  $p = 0$  and we always receive  $a_2$ . If, on the other hand, both letters are equiprobable, our *a priori* ignorance is maximum and therefore when we receive a letter, we gain the maximum

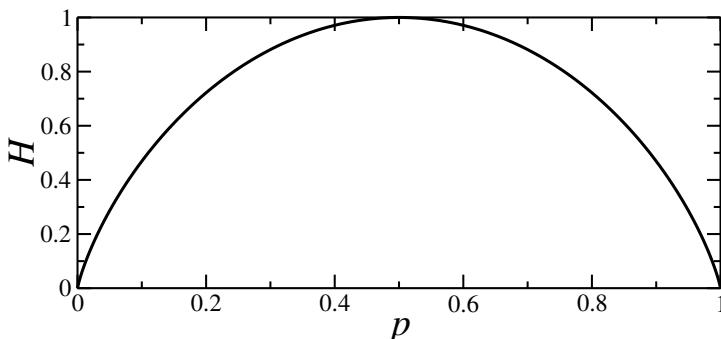


Fig. 6.2 The Shannon binary entropy  $H(p) = -p \log p - (1-p) \log(1-p)$ .

possible information  $H(\frac{1}{2}) = 1$ . In this case, we say that we have received one unit of information, known as a *bit*. Typically, we write the letters as binary digits; that is,  $a_1 = 0$  and  $a_2 = 1$ .

**Exercise 6.2** Show that the Shannon entropy  $H(p_1, \dots, p_k)$  is maximum when  $p_1 = \dots = p_k = 1/k$ .

### 6.3.1 Mutual information

As we will detail later in Chap. 8, Shannon entropy is a function of the information content of a given random variable  $X$  that takes the value  $x \in \mathcal{A} = \{a_1, \dots, a_k\}$  with probability distribution  $p(x) \in \{p_1, \dots, p_k\}$ . Hereafter we will adopt the compact notation  $H(X)$  for the corresponding function  $H(p_1, \dots, p_k)$ , which can be calculated according to the definition in Eq. (6.25). Let us now consider two random variables  $X$  and  $Y$ . It is possible to quantify how much information they have in common, by means of the so-called mutual information. This can be introduced through the following definitions.

#### Joint entropy

The joint entropy of a pair of random variables  $X$  and  $Y$  having values  $x$  and  $y$  with probabilities  $p(x)$  and  $p(y)$ , respectively, is defined by

$$H(X, Y) \equiv - \sum_{x,y} p(x, y) \log p(x, y), \quad (6.27)$$

where  $p(x, y)$  is the probability that  $X = x$  and  $Y = y$ . Thinking in terms of set theory, if we represent the information content of a given random variable with a set,  $H(X, Y)$  describes the union of the two sets corresponding to the two variables  $X$  and  $Y$  (the full colored area in Fig. 6.3).

#### Conditional entropy

The conditional entropy  $H(Y|X)$  is defined by

$$H(Y|X) \equiv H(X, Y) - H(X). \quad (6.28)$$

It is a measure of our residual ignorance about  $Y$ , provided we already know the value of  $X$  (in terms of set theory, this is represented by the subtraction between set  $X$ , and its intersection with set  $Y$ , as in Fig. 6.3). Similarly, we can define  $H(X|Y) \equiv H(X, Y) - H(Y)$ . It is easy to show that

$$H(Y|X) = - \sum_{x,y} p(x, y) \log p(y|x), \quad (6.29)$$

where  $p(y|x) = p(x, y)/p(x)$  defined the conditional probability that  $Y = y$ , provided  $X = x$  (this follows from Bayes rule for classical variables, see, for example, Stone, 2013). Indeed,

$$\begin{aligned}
H(X, Y) - H(X) &= - \sum_{x,y} p(x, y) \log p(x, y) + \sum_x p(x) \log p(x) \\
&= - \sum_{x,y} p(x, y) \log(p(x)p(y|x)) + \sum_{x,y} p(x, y) \log p(x) \\
&= - \sum_{x,y} p(x, y) \log p(y|x),
\end{aligned} \tag{6.30}$$

where we have used  $\sum_y p(x, y) = p(x)$ . Similarly, we obtain

$$H(X|Y) = - \sum_{x,y} p(x, y) \log p(x|y). \tag{6.31}$$

### Mutual information

The mutual information  $\mathcal{I}(X:Y)$  is defined as

$$\mathcal{I}(X:Y) \equiv H(X) + H(Y) - H(X, Y). \tag{6.32}$$

This quantity is a measure of how much information  $X$  and  $Y$  have in common (i.e., the intersection between the two sets in Fig. 6.3). Using again Bayes rule as in Eq. (6.30), it can be easily shown that

$$\mathcal{I}(X:Y) = - \sum_{x,y} p(x, y) \log \frac{p(x)p(y)}{p(x, y)}. \tag{6.33}$$

It thus immediately appears that the mutual information is related to the conditional entropy as follows:

$$\mathcal{I}(X:Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{6.34}$$

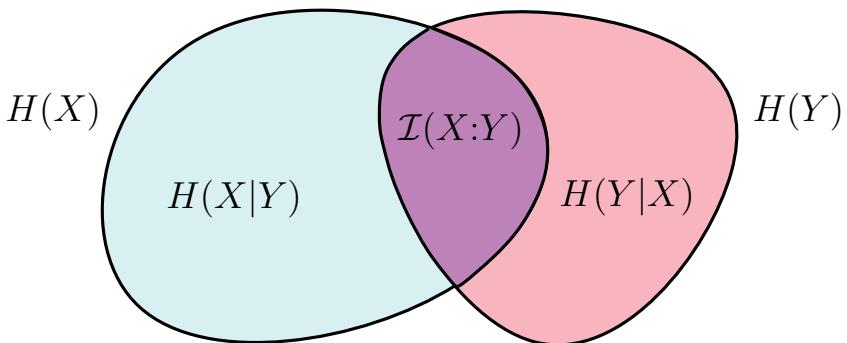


Fig. 6.3 Pictorial representation, in terms of set theory, of the entropies and the mutual information involving two random variables. The Shannon entropies,  $H(X)$  and  $H(Y)$ , associated to the information content of each variable  $X$  and  $Y$ , are depicted by two sets. The joint entropy  $H(X, Y)$  is simply the union of the two sets. The conditional entropy  $H(X|Y)$ , associated to the information content of variable  $X$ , once  $Y$  is known, is the subtraction between set  $X$  and the intersection among the two sets, which represents their mutual information  $\mathcal{I}(X:Y)$ . The converse occurs for  $H(Y|X)$ .

Note that, as is clear from its definition (6.32), the mutual information is symmetric:

$$\mathcal{I}(Y:X) = \mathcal{I}(X:Y). \quad (6.35)$$

From Eq. (6.33) it is also clear that, if  $X$  and  $Y$  are independent, namely  $p(x,y) = p(x)p(y)$ , then  $\mathcal{I}(X:Y) = 0$ .

We now assume that Alice wants to send a message to Bob, and we would like to know how much information can Bob gain on the message by performing measurements on the received state. If  $X$  and  $Y$  denote the random variables associated with the letters generated by Alice and with Bob's measurement outcomes, respectively, then the *accessible information* is defined as the maximum of the mutual information  $\mathcal{I}(X:Y)$  over all possible measurement schemes. We will show in Sec. 6.9 that, moving to the quantum scenario, the situation will be less trivial, since all classical probability distributions can be thought as special cases of density matrices, namely purely diagonal density matrices.

**Exercise 6.3** Given two probability distributions  $p(x)$  and  $q(x)$ , we define the *relative entropy* of  $p$  relative to  $q$  as

$$D(p||q) = \sum_x p(x)[\log p(x) - \log q(x)]. \quad (6.36)$$

Prove that  $D(p||q) \geq 0$ , with equality if and only if  $p(x) = q(x)$  for all  $x$  (*Hint*: use the inequality  $\log x \leq (x-1)/\ln 2$  for all positive  $x$ , with equality if and only if  $x = 1$ ).

**Exercise 6.4** Prove that  $H(X) \leq H(X, Y)$ .

**Exercise 6.5** Prove the *subadditivity* property of the Shannon entropy:  $H(X, Y) \leq H(X) + H(Y)$ .

## 6.4 The von Neumann entropy

Let us now define the quantum analogue of the Shannon entropy, that is the von Neumann entropy. If a quantum system is described by the density matrix  $\rho$ , its von Neumann entropy  $S(\rho)$  is defined as

$$S(\rho) \equiv -\text{Tr}(\rho \log \rho). \quad (6.37)$$

To see the analogy with the Shannon entropy, let us consider the following situation: Alice has at her disposal an alphabet  $\mathcal{A} = \{\rho_1, \rho_2, \dots, \rho_k\}$ , where the letters  $\rho_i$  are density matrices describing quantum states (pure or mixed). The letters are chosen at random with probabilities  $p_i$ , where  $\sum_{i=1}^k p_i = 1$ . Let us assume that Alice sends a letter (a quantum state) to Bob and that Bob only knows that the letter has been taken from the ensemble  $\{\rho_i, p_i\}$ . Thus, he describes this quantum system by means of the density matrix

$$\rho = \sum_{i=1}^k p_i \rho_i. \quad (6.38)$$

Therefore,

$$S(\rho) = -\text{Tr}(\rho \log \rho) = -\sum_{i=1}^k \lambda_i \log \lambda_i = H(\lambda_1, \dots, \lambda_k), \quad (6.39)$$

where the  $\lambda_i$  are the eigenvalues of the density matrix  $\rho$  and  $H(\lambda_1, \dots, \lambda_k)$  is the Shannon entropy associated with the ensemble  $\{\lambda_i\}$ .

The von Neumann entropy satisfies the following properties:

- (1) For a pure state,  $S(\rho) = 0$ . Indeed, in this case only one eigenvalue of  $\rho$  is different from zero, say  $\lambda_1 = 1$ , so that  $-\sum_i \lambda_i \log \lambda_i = -\lambda_1 \log \lambda_1 = 0$ .
- (2) The entropy is not modified by a unitary change of basis; that is,  $S(U\rho U^\dagger) = S(\rho)$ . Actually  $S(\rho)$  depends only on the eigenvalues of  $\rho$ , which are basis-independent. This property means that the von Neumann entropy is invariant under unitary temporal evolution.
- (3) If the density operator  $\rho$  acts on a  $N$ -dimensional Hilbert space, then  $0 \leq S(\rho) \leq \log N$ . It is easy to see that  $S(\rho) \geq 0$  since  $0 \leq \lambda_i \leq 1$  and therefore  $-\lambda_i \log \lambda_i \geq 0$ . To show that  $S(\rho) \leq \log N$ , we use  $S(\rho) = H(\lambda_1, \dots, \lambda_N)$  and remember that the Shannon entropy  $H(\lambda_1, \dots, \lambda_N)$  takes its maximum value  $\log N$  when  $\lambda_1 = \dots = \lambda_N = 1/N$  (see exercise 6.2). Hence,  $S_{\max} = -\frac{1}{N} \sum_{i=1}^N \log \frac{1}{N} = \log N$ .
- (4) For bipartite *pure* states  $\rho_{AB} = |\psi\rangle_{AB} \langle \psi|$ , we have  $S(\rho_A) = S(\rho_B)$ , where  $\rho_A$  and  $\rho_B$  are the reduced density matrices,  $\rho_A = \text{Tr}_B(\rho_{AB})$  and  $\rho_B = \text{Tr}_A(\rho_{AB})$ . This property follows from the fact that  $\rho_A$  and  $\rho_B$  have the same nonzero eigenvalues, as it can be readily seen from the Schmidt decomposition of  $|\psi\rangle_{AB}$ .
- (5) *Subadditivity*: for any state  $\rho_{AB}$  of a bipartite system  $AB$ ,

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B), \quad (6.40)$$

with  $\rho_A$  and  $\rho_B$  reduced density matrices of the subsystems  $A$  and  $B$ , respectively. The equality in Eq. (6.40) occurs if and only if  $\rho_{AB} = \rho_A \otimes \rho_B$ . For a proof of this property, see exercise 6.7.

- (6) *Concavity*: given probabilities  $p_i$ , namely  $p_i \geq 0$  for all  $i$  and  $\sum_i p_i = 1$ , we have

$$S\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i S(\rho_i). \quad (6.41)$$

This result can be intuitively understood as the ignorance about the overall state  $\sum_i p_i \rho_i$  is larger than the weighted ignorances of the single states  $\rho_i$ , i.e.  $\sum_i p_i S(\rho_i)$ , since there is also a contribution due to our ignorance about which state in the ensemble was prepared. For a proof of this property, see exercise 6.8. If in particular one assumes that the set of operators  $\rho_i$  has support<sup>1</sup> on orthogonal subspaces, then (see exercise 6.9)

$$S\left(\sum_i p_i \rho_i\right) = \sum_i p_i S(\rho_i) + H(p_1, p_2, \dots). \quad (6.42)$$

---

<sup>1</sup>The support of an operator is the subspace spanned by the set of eigenvectors with nonzero eigenvalues.

- (7) *Triangle inequality* (also known as Araki-Lieb inequality): For a bipartite system,

$$S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|. \quad (6.43)$$

For a proof, see exercise 6.10. Note that the triangle inequality allows  $S(\rho_{AB}) \leq S(\rho_A), S(\rho_B)$ , while for the Shannon entropy  $H(X, Y) \geq H(X), H(Y)$  (see exercise 6.4), namely the entropy of a part of a classical bipartite system cannot be larger than the entropy of the entire system. On the other hand, for an entangled pure quantum state  $|\psi\rangle_{AB}$ , we have  $S(\rho_{AB}) = S(|\psi\rangle_{AB} \langle \psi|) = 0$  while (see Sec. 6.5)  $S(\rho_A) = S(\rho_B) > 0$ . In this case we have complete knowledge of the overall state  $\rho_{AB}$ , since it is a pure state, while our knowledge of the (mixed) states of the parts  $A$  and  $B$  is incomplete. Information on the overall state is encoded in the nonclassical correlations between the parts of the system, which cannot be detected when we look at  $A$  or  $B$  separately.

- (8) *Strong subadditivity*: for any state  $\rho_{ABC}$  of a tripartite system  $ABC$ ,

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC}). \quad (6.44)$$

For a proof of this result see, e.g., Wilde (2013).

**Exercise 6.6** Given two density operators  $\rho$  and  $\sigma$ , the *quantum relative entropy* of  $\rho$  with respect to  $\sigma$  is defined by

$$D(\rho||\sigma) = \text{Tr}[\rho(\log \rho - \log \sigma)]. \quad (6.45)$$

Prove that  $D(\rho||\sigma) \geq 0$  with equality if and only if  $\rho = \sigma$  (*Hint*: use the spectral decompositions  $\rho = \sum_i p_i |i\rangle\langle i|$  and  $\sigma = \sum_\alpha q_\alpha |\alpha\rangle\langle \alpha|$  to write

$$D(\rho||\sigma) = \sum_i p_i \left( \log p_i - \sum_\alpha A_{i\alpha} \log q_\alpha \right), \quad (6.46)$$

where  $A_{i\alpha}$  is a *doubly stochastic matrix*, namely its matrix elements are non-negative real numbers, and each row and each column sums to one:  $\sum_i A_{i\alpha} = \sum_\alpha A_{i\alpha} = 1$ . Then use the fact that the logarithm is a concave function to show that  $D(\rho||\sigma)$  is lower bounded by a (classical) relative entropy, which we have seen in exercise 6.3 to be nonnegative).

**Exercise 6.7** Using the nonnegativity of the quantum relative entropy of  $\rho_{AB}$  with respect to  $\rho_A \otimes \rho_B$ , prove that the von Neumann entropy is subadditive.

**Exercise 6.8** Using subadditivity, prove the concavity of the von Neumann entropy (*Hint*: consider the state

$$\rho_{AB} = \sum_i p_i \rho_i \otimes |i\rangle\langle i|, \quad (6.47)$$

where  $\{|i\rangle\}$  is an orthonormal basis of an ancillary system  $B$ ).

**Exercise 6.9** Prove Eq. (6.42).

**Exercise 6.10** Prove the triangle inequality (*Hint*: consider a pure state  $|\psi\rangle_{ABC}$  which purifies  $\rho_{AB}$ , namely  $\rho_{AB} = \text{Tr}_C (|\psi\rangle_{ABC} \langle \psi|)$ ).

The following examples give a flavour of the similarities and the differences between the von Neumann entropy and the Shannon entropy.

### 6.4.1 Example 1: source of orthogonal pure states

In the simplest case, Alice has at her disposal a source of two orthogonal pure states for a qubit. These states constitute a basis for the single qubit Hilbert space and we call them  $|0\rangle$  and  $|1\rangle$ . The corresponding density matrices are  $\rho_0 = |0\rangle\langle 0|$  and  $\rho_1 = |1\rangle\langle 1|$ . We assume that the source generates the states  $|0\rangle$  or  $|1\rangle$  with the *a priori* probabilities  $p_0 = p$  and  $p_1 = 1 - p$ , respectively. Therefore, we can write

$$\rho = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1| = \begin{bmatrix} p_0 & 0 \\ 0 & p_1 \end{bmatrix}, \quad (6.48)$$

and the von Neumann entropy is given by

$$\begin{aligned} S(\rho) &= -\text{Tr}(\rho \log \rho) = -\text{Tr}\left(\begin{bmatrix} p_0 & 0 \\ 0 & p_1 \end{bmatrix} \begin{bmatrix} \log p_0 & 0 \\ 0 & \log p_1 \end{bmatrix}\right) \\ &= -p_0 \log p_0 - p_1 \log p_1 = H(p_0, p_1). \end{aligned} \quad (6.49)$$

Therefore, in this case, in which the letters of the alphabet correspond to orthogonal pure states, the von Neumann entropy coincides with the Shannon entropy. Thus, the situation is in practice classical, from the point of view of information theory. This is quite natural since orthogonal states are perfectly distinguishable.

### 6.4.2 Example 2: source of non-orthogonal pure states

Let us consider the case in which the pure states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  generated by a source are not orthogonal. It is always possible to choose an appropriate basis set  $\{|0\rangle, |1\rangle\}$  (see Fig. 6.4) so that

$$|\tilde{0}\rangle = \cos \theta |0\rangle + \sin \theta |1\rangle = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad |\tilde{1}\rangle = \sin \theta |0\rangle + \cos \theta |1\rangle = \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix}. \quad (6.50)$$

We consider, without any loss of generality,  $0 \leq \theta \leq \pi/4$ . Note that the inner product of these two states is in general non-zero and given by

$$\langle \tilde{0} | \tilde{1} \rangle = \sin 2\theta. \quad (6.51)$$

The density matrices corresponding to the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  read

$$\rho_0 = |\tilde{0}\rangle\langle \tilde{0}| = \begin{bmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{bmatrix}, \quad \rho_1 = |\tilde{1}\rangle\langle \tilde{1}| = \begin{bmatrix} \sin^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \cos^2 \theta \end{bmatrix}. \quad (6.52)$$

If the source generates the state  $|\tilde{0}\rangle$  with probability  $p$  and the state  $|\tilde{1}\rangle$  with probability  $(1 - p)$ , the corresponding density matrix is

$$\rho = p\rho_0 + (1 - p)\rho_1 = \begin{bmatrix} \sin^2 \theta + p \cos 2\theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \cos^2 \theta - p \cos 2\theta \end{bmatrix}. \quad (6.53)$$

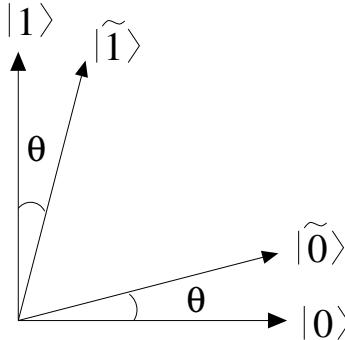


Fig. 6.4 A representation of two non-orthogonal quantum states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  in an appropriately chosen basis  $\{|0\rangle, |1\rangle\}$  for a qubit.

The eigenvalues of the density matrix are

$$\lambda_{\pm} = \frac{1}{2} \left( 1 \pm \sqrt{1 + 4p(p-1) \cos^2 2\theta} \right). \quad (6.54)$$

They are represented in the upper panel of Fig. 6.5, as a function of the probability  $p$  and for different values of  $\theta$ . We note that for  $\theta = 0$  the states are orthogonal and the eigenvalues of the density matrix are  $p$  and  $1 - p$ ; namely, we recover the classical case. For the other values of  $\theta$  the eigenvalues “repel” each other.

Starting from the eigenvalues of the density matrix (6.53), it is easy to compute the von Neumann entropy

$$S(\rho) = -\lambda_+ \log \lambda_+ - \lambda_- \log \lambda_-, \quad (6.55)$$

which is shown in the bottom panel of Fig. 6.5. At  $\theta = 0$ , we recover the classical results since in this case  $S(\rho) = H(p)$ . If  $\theta = \pi/4$ , then  $S(\rho) = 0$ . Indeed, since in this case the states are identical, there is no transmission of information. As can be seen in the figure,  $S(\rho) \leq H(p)$  and it is possible to prove that this inequality has general validity. A qualitative interpretation follows from our understanding of entropy as a measure of our ignorance about the system. If the states are non-orthogonal, their similarity increases with their inner product  $\langle \tilde{0} | \tilde{1} \rangle = \sin 2\theta$ . Therefore, Bob obtains less information from the reception of a state taken from the ensemble  $\{|\tilde{0}\rangle, |\tilde{1}\rangle\}$  since his *a priori* ignorance is smaller. In the limiting case  $\theta = \pi/4$  the superposition of the states of the ensemble is unity; that is, the states are identical and there is no *a priori* ignorance about the system. Therefore, no information is transmitted in this case.

## 6.5 Entanglement concentration

We can now work out the necessary formalism which is needed to quantify the operational definition of entanglement that we gave in Sec. 6.1.

As we have seen in Chap. 5, several important communication protocols, such as dense coding and quantum teleportation, strongly rely on the assumption of

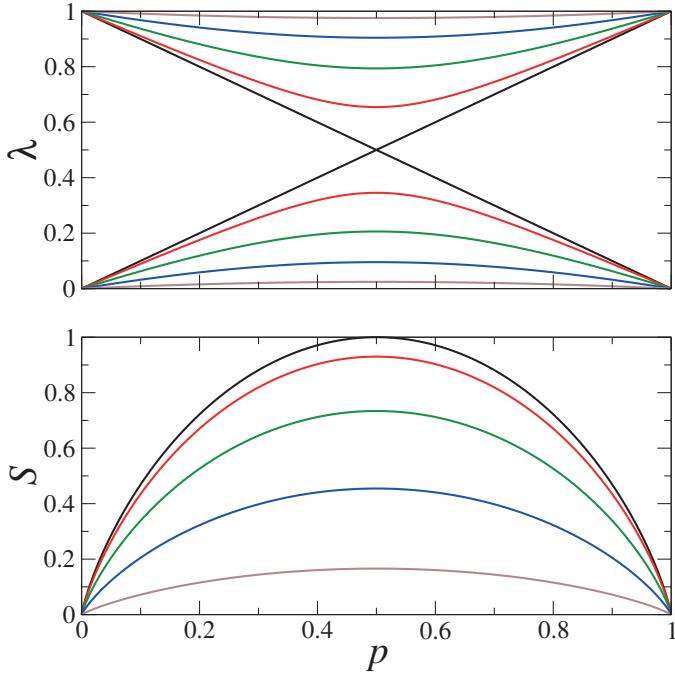


Fig. 6.5 Eigenvalues (upper panel) and von Neumann entropy (lower panel) of the density matrix (6.53) as a function of the probability  $p$ . The various colors are associated with the following values of the angle  $\theta$ :  $\theta = 0$  (black),  $\theta = 0.2 \times \frac{\pi}{4}$  (red),  $\theta = 0.4 \times \frac{\pi}{4}$  (green),  $\theta = 0.6 \times \frac{\pi}{4}$  (blue), and  $\theta = 0.8 \times \frac{\pi}{4}$  (brown). The value  $\theta = 0$  corresponds to orthogonal states.

having an initially entangled state. We wish to stress that teleportation is also interesting from the viewpoint of quantum computation since it is a powerful tool for transferring quantum states between different systems, as would be necessary in a quantum computer with several independent units. Since faithful teleportation requires that Alice and Bob share a maximally entangled EPR pair, it is important to devise methods to *distill* maximally entangled states starting from partially entangled pairs. These entanglement concentration techniques act on qubits that can be located very far away and therefore only rely on LOCC.

It is instructive to study in detail the following example of entanglement concentration, devised by Bandyopadhyay (2000). We assume that initially Alice and Bob share a pure entangled state. Taking into account the Schmidt decomposition described in Sec. 2.7, we can write this state as

$$|\psi\rangle_{AB} = \alpha|00\rangle_{AB} + \beta|11\rangle_{AB}, \quad (6.56)$$

where, without any loss of generality, we may assume  $\alpha, \beta$  to be real and positive, and  $\alpha \geq \beta$ . We assume that Alice knows the coefficients  $\alpha, \beta$  of the Schmidt decomposition in advance and prepares an ancillary qubit in the state

$$|\chi\rangle_A = \alpha|0\rangle_A + \beta|1\rangle_A. \quad (6.57)$$

Hence, the combined state of the three qubits is given by

$$\begin{aligned} |\Psi\rangle &= |\chi\rangle_A \otimes |\psi\rangle_{AB} = (\alpha|0\rangle_A + \beta|1\rangle_A) \otimes (\alpha|00\rangle_{AB} + \beta|11\rangle_{AB}) \\ &= \alpha^2|000\rangle_{A_1 A_2 B} + \alpha\beta|011\rangle_{A_1 A_2 B} + \alpha\beta|100\rangle_{A_1 A_2 B} + \beta^2|111\rangle_{A_1 A_2 B}. \end{aligned} \quad (6.58)$$

The first two qubits, denoted by  $A_1$  and  $A_2$ , belong to Alice and the third ( $B$ ) to Bob. Alice performs a CNOT gate on the two qubits in her possession,  $A_1$  being the control and  $A_2$  the target qubits. The resulting state is

$$|\Psi\rangle = \alpha^2|000\rangle_{A_1 A_2 B} + \alpha\beta|011\rangle_{A_1 A_2 B} + \alpha\beta|110\rangle_{A_1 A_2 B} + \beta^2|101\rangle_{A_1 A_2 B}. \quad (6.59)$$

After interchanging the position of the first two qubits and writing the wave-function normalization in an appropriate manner, we have

$$\begin{aligned} |\Psi\rangle &= \sqrt{\alpha^4 + \beta^4}|0\rangle_{A_2} \otimes \left( \frac{\alpha^2}{\sqrt{\alpha^4 + \beta^4}}|00\rangle_{A_1 B} + \frac{\beta^2}{\sqrt{\alpha^4 + \beta^4}}|11\rangle_{A_1 B} \right) \\ &\quad + \sqrt{2\alpha^2\beta^2}|1\rangle_{A_2} \otimes \frac{1}{\sqrt{2}}(|01\rangle_{A_1 B} + |10\rangle_{A_1 B}). \end{aligned} \quad (6.60)$$

Alice then performs a standard projective measurement of qubit  $A_2$  on the basis  $\{|0\rangle, |1\rangle\}$ . It is straightforward to see from Eq. (6.60) that she obtains outcome 0 with probability  $(\alpha^4 + \beta^4)$  or outcome 1 with probability  $(2\alpha^2\beta^2)$ . In the latter case, Alice and Bob realize an EPR pair (the qubits  $A_1$  and  $B$ ). In the first case, they obtain less entangled states. Thus, given  $N$  non-maximally entangled states (6.56), the above technique produces  $2\alpha^2\beta^2N$  maximally entangled states. As shown by Bandyopadhyay (2000), it is possible to iterate the procedure for the remaining  $N(1 - 2\alpha^2\beta^2)$  states to improve its efficiency (the efficiency being defined as the fraction of EPR pairs extracted). We note that classical communication is also required for Alice to transmit the results of her measurements to Bob, in order to select the successful cases.

It might appear paradoxical that local operations plus classical communication allow a concentration of entanglement, which is a purely quantum non-local property. However, there is no real surprise if we remember that quantum mechanics is a probabilistic theory and that a non-vanishing maximally entangled component is present in the state (6.56). This component is quantified by the fidelity

$$F = |{}_{AB}\langle \phi^+ | \psi \rangle_{AB}|^2 = \left| \frac{1}{\sqrt{2}}(\langle 00| + \langle 11|)(\alpha|00\rangle + \beta|11\rangle) \right|^2 = \frac{1}{2}(\alpha + \beta)^2, \quad (6.61)$$

where we have considered the EPR state  $|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . Therefore, the above entanglement concentration protocol selects this maximally entangled component.

The previous example naturally raises the following questions: What is the optimal entanglement concentration? Can we measure entanglement? Nowadays, we can answer these questions unambiguously but only for bipartite pure states. First of all, a few definitions are needed.

### Entanglement cost

Let us assume that Alice and Bob share many EPR pairs, say  $|\phi^+\rangle_{AB}$ , and that they wish to prepare a large number  $n$  of copies of a given bipartite pure state  $|\psi\rangle_{AB}$ , using only local operations and classical communication. If we call  $k_{\min}$  the minimum number of EPR pairs necessary to accomplish this task, we define the entanglement cost  $E_C(|\psi\rangle_{AB})$  as the limiting ratio  $k_{\min}/n$ , for  $n \rightarrow \infty$ . Formally, we can write:

$$E_C(|\psi\rangle_{AB}) = \inf_{\Gamma_{\text{LOCC}}} \lim_{n \rightarrow \infty} \frac{k_{\min}}{n}. \quad (6.62)$$

Here  $\Gamma_{\text{LOCC}}$  denotes all the set of possible LOCC between Alice and Bob.

### Distillable entanglement

Now consider the reverse process; that is, Alice and Bob share a large number  $n$  of copies of a pure state  $|\psi\rangle_{AB}$  and they wish to concentrate entanglement, again using only local operations supplemented by classical communication. If  $k'_{\max}$  denotes the maximum number of EPR pairs that can be obtained in this manner, we define the distillable entanglement as the ratio  $k'_{\max}/n$  in the limit  $n \rightarrow \infty$ . Formally, we can write:

$$E_D(|\psi\rangle_{AB}) = \sup_{\Gamma_{\text{LOCC}}} \lim_{n \rightarrow \infty} \frac{k'_{\max}}{n}. \quad (6.63)$$

It is clear that  $k'_{\max} \leq k_{\min}$ . Otherwise, we could employ local operations and classical communication to create entanglement, which is a non-local, purely quantum resource (it would be sufficient to prepare  $n$  states  $|\psi\rangle_{AB}$  from  $k_{\min}$  EPR pairs and then distill  $k'_{\max} > k_{\min}$  EPR states). Furthermore, it is possible to show that, asymptotically in  $n$ , the entanglement cost and the distillable entanglement coincide and that the ratios  $k_{\min}/n$  and  $k'_{\max}/n$  are given by the reduced single-qubit von Neumann entropies. Indeed, we have

$$E_C(|\psi\rangle_{AB}) = E_D(|\psi\rangle_{AB}) = S(\rho_A) = S(\rho_B), \quad (6.64)$$

where  $S(\rho_A)$  and  $S(\rho_B)$  are the von Neumann entropies of the reduced density matrices  $\rho_A = \text{Tr}_B (|\psi\rangle_{AB} \langle \psi|)$  and  $\rho_B = \text{Tr}_A (|\psi\rangle_{AB} \langle \psi|)$ , respectively. Therefore, the process that changes  $n$  copies of  $|\psi\rangle_{AB}$  into  $k$  copies of  $|\phi^+\rangle_{AB}$  is asymptotically reversible. Moreover, it is possible to show that it is faithful; namely, the change takes place with unit fidelity when  $n \rightarrow \infty$ . The proof of this result can be found in Bennett *et al.* (1996a). We can therefore univocally quantify the entanglement of a bipartite pure state  $|\psi\rangle_{AB}$  as

$$E(|\psi\rangle_{AB}) = S(\rho_A) = S(\rho_B). \quad (6.65)$$

In this context, the von Neumann entropy  $S(\rho_A)$  is known as the *entropy of entanglement*. It ranges from 0 for a separable state to 1 for maximally entangled two-qubit states (the EPR states). Hence, it is common practice to say that the

entanglement of an EPR pair is 1 *ebit*. More generally, a maximally entangled state of two subsystems has  $d$  equally weighted terms in its Schmidt decomposition ( $d$  is the dimension of the Hilbert space of the smaller subsystem) and therefore its entanglement content is  $\log d$  ebits.

Up to now we quantified entanglement from the perspective of LOCC in the asymptotic limit. As we have seen, this approach can be solved completely for pure states. Indeed it demonstrates that the manipulation of entanglement in the LOCC framework is reversible, thus imposing a unique order on pure entangled states via the entropy of entanglement. A natural extension of this discussion would consider bipartite mixed states, with  $\rho_{AB} = \sum_i p_i |\psi_i\rangle_{AB} \langle\psi_i|$ , instead of pure states. Unfortunately, for mixed states the situation is far more complicated and the above mentioned reversibility is lost. In general we may have that  $E_C(\rho_{AB}) \neq E_D(\rho_{AB})$ . The mixed-state entanglement is the focus of ongoing research (see the guide to the bibliography at the end of the chapter).

### 6.5.1 \* Entanglement of a random state

Before discussing entanglement measures applicable to the generic case of mixed states, it might be interesting to consider the entanglement content of a “generic” pure quantum state, that is, of a state with randomly drawn components (we will clarify below in this section the definition of a pure random state). A simple argument helps understanding why the bipartite entanglement content of a pure random state  $|\psi\rangle$  is almost maximal. In a given basis  $\{|i\rangle\}$  the density matrix for the state  $|\psi\rangle = \sum_i c_i |i\rangle$  is written as follows:

$$\rho_{ij} = \langle i|\psi\rangle\langle\psi|j\rangle = c_i c_j^*, \quad (6.66)$$

where  $c_i = \langle i|\psi\rangle$  are the components of the state  $|\psi\rangle$  in the  $\{|i\rangle\}$  basis. In the case of a random state we expect the components to be uniformly distributed, with amplitudes  $c_i \approx 1/\sqrt{N}$  and random phases. Here  $N$  is the Hilbert space dimension and the value  $1/\sqrt{N}$  of the amplitudes ensures that the wave vector  $|\psi\rangle$  is normalized. The density matrix can therefore be written as

$$\rho \approx \text{diag}\left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right) + \Omega, \quad (6.67)$$

where  $\Omega$  is a  $N \times N$  zero-diagonal matrix with random complex matrix elements of amplitude  $\approx 1/N$ . Suppose now that we partition the Hilbert space of the system into two parts,  $A$  and  $B$ , with dimensions  $N_A$  and  $N_B$ , where  $N_A N_B = N$ . Without loss of generality, we assume  $N_A \leq N_B$ . The reduced density matrix  $\rho_A$  is defined as follows:

$$\rho_A = \text{Tr}_B \rho = \sum_{i_B} c_{i_A i_B} c_{i'_A i_B}^* |i_A\rangle\langle i'_A|, \quad (6.68)$$

where  $|i\rangle = |i_A i_B\rangle$ . Using Eq. (6.67), we obtain

$$\rho_A \approx \text{diag}\left(\frac{1}{N_A}, \frac{1}{N_A}, \dots, \frac{1}{N_A}\right) + \Omega_A, \quad (6.69)$$

where  $\Omega_A$  is a zero-diagonal matrix with matrix elements of  $O(\sqrt{N_B}/N \ll 1/N_A)$  (sum of  $N_B \gg 1$  terms of order  $1/N$  with random phases). Neglecting  $\Omega_A$  in (6.69), the reduced von Neumann entropy of subsystem  $A$  is given by  $S(\rho_A) = \log(N_A)$ , that is, the maximum entropy that the subsystem  $A$  can have.

### Page's formula

The exact mean value  $\langle E_{AB}(|\psi\rangle)\rangle$  of the bipartite entanglement is given by Page's formula, obtained by considering the ensemble of random pure states drawn according to the Haar measure on  $U(N)$  (*i.e.*, on the group of  $N \times N$  unitary matrices). More precisely, we write an  $N$ -level random state in the form

$$|\psi\rangle = \sum_{k=0}^{N-1} r_k e^{i\phi_k} |k\rangle, \quad (6.70)$$

where  $\phi_k$  are independent random variables uniformly distributed in  $[0, 2\pi)$  and  $\mathbf{r} = (r_0, \dots, r_{N-1})$  is a random point uniformly distributed on the unit hypersphere  $\mathbb{S}^{N-1} = \{\mathbf{r} \in \mathbb{R}^N \mid \mathbf{r}^2 = 1\}$ , with distribution function

$$p(\mathbf{r}) = C_N \prod_{k=0}^{N-1} r_k \delta(\mathbf{r}^2 - 1), \quad (6.71)$$

with  $C_N$  a constant introduced to normalize the random state (the value of  $C_N$  is determined in the solution to exercise 6.11). Page's formula states that, for  $N_A \leq N_B$ ,

$$S_P \equiv \langle E_{AB}(|\psi\rangle)\rangle = \langle S(\rho_A) \rangle = \langle S(\rho_B) \rangle = \frac{1}{\ln 2} \left[ \sum_{k=N_B+1}^{N_A N_B} \frac{1}{k} - \frac{N_A - 1}{2N_B} \right], \quad (6.72)$$

where  $\langle \cdot \rangle$  denotes the (ensemble) average over the Haar measure. For  $1 \ll N_A \leq N_B$ , this formula simplifies as follows:

$$S_P \approx \log N_A - \frac{N_A}{2N_B \ln 2}, \quad (6.73)$$

and therefore  $\langle E_{AB} \rangle$  is close to its maximum value  $E_{AB}^{\max}(|\psi\rangle) = \log N_A \gg 1$ . Note that, if we fix  $N_A$  and let  $N_B \rightarrow \infty$ , then  $\langle E_{AB} \rangle$  tends to  $E_{AB}^{\max}$ .

Remarkably, if we consider the *thermodynamic limit*, that is, we fix  $N_A/N_B$  and let  $N_A \rightarrow \infty$ , then the reduced von Neumann entropy concentrates around its average value (6.72). This is a consequence of the so-called *concentration of measure* phenomenon: the uniform measure on the  $k$ -sphere  $\mathbb{S}^k$  in  $\mathbb{R}^{k+1}$  concentrates very strongly around the equator when  $k$  is large: Any polar cap smaller than a hemisphere has relative volume exponentially small in  $k$ . This observation implies, in particular, the concentration of the entropy of the reduced density matrix  $\rho_A$  around its average value, see Życzkowski and Sommers (2001); Hayden *et al.* (2006). This in turn implies that when the dimension  $N$  of the quantum system is large it is meaningful to apply statistical methods and discuss typical (entanglement) behaviour or random states, in the sense that almost all random states behave in essentially the same way.

### The purity and Lubkin's formula

The purity of state described by the density matrix  $\rho$  is defined as

$$P(\rho) = \text{Tr}(\rho^2). \quad (6.74)$$

Given a pure state  $|\psi\rangle \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$  of a bipartite quantum system, we can write (see Sec. 2.7) the Schmidt decomposition, that is, there exist orthonormal states  $\{|i\rangle_A\}$  for  $\mathcal{H}_A$  and  $\{|i'\rangle_B\}$  for  $\mathcal{H}_B$  such that  $|\psi\rangle = \sum_{i=1}^k \sqrt{p_i} |i\rangle_A |i'\rangle_B$ , and therefore

$$P(\rho_A) = P(\rho_B) = \sum_i p_i^2. \quad (6.75)$$

The purity is much easier to investigate analytically than the von Neumann entropy. Moreover, it provides the first non-trivial term in a Taylor series expansion of the von Neumann entropy about its maximum. Indeed, if we write  $p_i = \frac{1+\epsilon_i}{N_A}$ , with  $\epsilon_i \ll 1$  and  $\sum_i \epsilon_i = 0$ , we obtain

$$S(\rho_A) \approx \log N_A - \frac{N_A}{2 \ln 2} P(\rho_A). \quad (6.76)$$

The purity ranges from  $1/N_A$  for maximally entangled states to 1 for separable states.

For random states, the average value of the purity of the reduced density matrices  $\rho_A$  and  $\rho_B$  is given by Lubkin's formula (see exercise 6.11):

$$P_L \equiv \langle P(\rho_A) \rangle = \langle P(\rho_B) \rangle = \frac{N_A + N_B}{N_A N_B + 1}. \quad (6.77)$$

Note that, if we fix  $N_A$  and let  $N_B \rightarrow \infty$ , then  $P_L$  tends to its minimum value  $1/N_A$ . If we fix  $N_A/N_B$  and let  $N_A \rightarrow \infty$ , then  $P_L \rightarrow 0$ . For large  $N$ , the variance

$$\sigma_P^2 = \langle P^2 \rangle - P_L^2 \approx \frac{2}{N^2}, \quad (6.78)$$

so that the relative standard deviation

$$\frac{\sigma_P}{P_L} \approx \frac{\sqrt{2}}{N_A + N_B} \quad (6.79)$$

tends to zero in the thermodynamic limit  $N_A \rightarrow \infty$  (at fixed  $N_A/N_B$ ). For a *balanced bipartition*, corresponding to  $N_A = N_B = \sqrt{N}$ , we have

$$\frac{\sigma_P}{P_L} = O\left(\frac{1}{\sqrt{N}}\right). \quad (6.80)$$

Note that the fact that  $\sigma_P/P_L \rightarrow 0$  when  $N \rightarrow \infty$  is again a consequence of the concentration of measure phenomenon.

**Exercise 6.11** Using an  $N$ -level random state in the form (6.70), prove Eqs. (6.77) and (6.78).

## 6.6 Requirements for bipartite entanglement measures

The impossibility to order quantum (mixed) states in terms of rates of LOCC entanglement interconversions implies that, in general, an LOCC-based classification of entanglement could be very difficult to achieve. One can however proceed in a different fashion and try to build up alternative entanglement measures which satisfy the basic properties outlined in Sec. 6.1.1. A good entanglement measure  $E$  has to fulfill several requirements, which are listed below. However, it is still an open question whether all these conditions are necessary or not. In fact, some of them are not satisfied by many of the measurements proposed so far.

- (i) A bipartite entanglement measure  $E(\rho)$  is a map from density matrices to positive real numbers:

$$\rho \rightarrow E(\rho) \in \mathbb{R}^+, \quad (6.81)$$

where  $\rho$  is an arbitrary state of a bipartite quantum system. The measure is normalized in such a way that the maximally entangled state  $|\phi_d^+\rangle$  has  $E(|\phi_d^+\rangle) = \log d$ .

- (ii) If a quantum state  $\rho$  is separable, then  $E(\rho) = 0$ .
- (iii) Entanglement cannot increase under LOCC:

$$E[\Gamma_{\text{LOCC}}(\rho)] \leq E(\rho). \quad (6.82)$$

- (iv) *Continuity.* In the limit of vanishing distance between two density matrices, the difference between their entanglement should tend to zero:

$$\lim_{\|\rho - \sigma\| \rightarrow 0} [E(\rho) - E(\sigma)] = 0. \quad (6.83)$$

This implies that, for any pure state  $|\psi\rangle_{AB}$ , any entanglement measure  $E(|\psi\rangle_{AB})$  should reduce to the entropy of entanglement (6.65).

- (v) *Convexity.* Entanglement measures should be convex functions:

$$E\left(\sum_i p_i \rho_i\right) \leq \sum_i p_i E(\rho_i), \quad \text{where } \sum_i p_i = 1, \quad p_i > 0. \quad (6.84)$$

- (vi) *Additivity.* A given number  $n$  of identical copies of the state  $\rho$  should contain  $n$  times the entanglement of a single copy:

$$E(\rho^{\otimes n}) = n E(\rho). \quad (6.85)$$

Some significant examples of entanglement measures do not satisfy this condition. However there is a straightforward way to remove this deficiency, after defining the regularized, or asymptotic version:

$$E^\infty(\rho) = \lim_{n \rightarrow \infty} \frac{E(\rho^{\otimes n})}{n}, \quad (6.86)$$

which automatically satisfies Eq. (6.85).

- (vii) *Subadditivity.* For any pair of states  $\rho$  and  $\sigma$ , one should require that

$$E(\rho \otimes \sigma) \leq E(\rho) + E(\sigma). \quad (6.87)$$

Any function  $E(\rho)$  satisfying the first three conditions is called an *entanglement monotone*. The entanglement cost and the distillable entanglement are two examples of such monotones.

## 6.7 Other entanglement measures

Apart from the two entanglement measures defined in terms of rates of LOCC interconversions, namely the entanglement cost  $E_C$  of Eq. (6.62), and the distillable entanglement  $E_D$  of Eq. (6.63), there are other two important entanglement monotones.

### Entanglement of formation

A generic mixed state  $\rho$  can be always written as a statistical mixture of pure states, that is,  $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ , with  $\sum_i p_i = 1$ ,  $p_i > 0$ . The entanglement of formation is the minimal possible average entanglement over all the pure-state decompositions of  $\rho$ :

$$E_F(\rho) = \inf_{\text{dec}} \sum_i p_i E(|\psi_i\rangle), \quad (6.88)$$

where  $E(|\psi_i\rangle)$  is the entropy of entanglement of the pure state  $|\psi_i\rangle$ , quantified by the von Neumann entropy.

### Relative entropy of Entanglement

This measure is defined as

$$E_R(\rho) = \inf_{\sigma \in \text{sep}} \text{Tr}[\rho(\log \rho - \log \sigma)], \quad (6.89)$$

and quantifies a kind of “distance” (even if it is not a proper distance, from a mathematical point of view), or distinguishability, of the state  $\rho$  to the closest separable state  $\sigma$ .

The above defined measures are constrained by some inequalities. For example it is known that, for any entanglement measure  $E(\rho)$ , one has  $E_D(\rho) \leq E(\rho) \leq E_C(\rho)$ . Moreover the entanglement cost  $E_C(\rho)$  is identical to the regularized version of the entanglement of formation,  $E_F^\infty(\rho) = \lim_{n \rightarrow \infty} E_F(\rho^{\otimes n})/n$  (it is also conjectured, yet not proven, to be the same of the entanglement of formation itself).

Unfortunately all the entanglement monotones defined above are associated to variational problems which are extremely difficult to solve, in general. This restricts the quantification of an entanglement monotone, for bipartite systems, to few very specific cases. Specifically, while for pure states it can be shown that the four measurements coincide:  $E_C(|\psi\rangle_{AB}) = E_D(|\psi\rangle_{AB}) = E_F(|\psi\rangle_{AB}) = E_R(|\psi\rangle_{AB})$  and are given by the entropy of entanglement, for mixed states the situation is much more involved. One usually has to resort to numerical optimization techniques, while analytical methods are available only for special symmetric cases.

### 6.7.1 \* Concurrence

Generic bipartite states of two qubits represent a remarkable example where an exact closed-form solution for the entanglement of formation is known. This has

been found by Wootters (1998), and is given by

$$E_F(\rho) = H_{\text{bin}} \left( \frac{1 + \sqrt{1 + C(\rho)^2}}{2} \right), \quad (6.90)$$

where  $H_{\text{bin}}(x)$  is the Shannon binary entropy of Eq. (6.26). The quantity  $C(\rho)$  is called the *concurrence* of the state  $\rho$ , and is defined as

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \quad (6.91)$$

where  $\lambda_i$  are the square roots of the eigenvalues, in descending order, of the Hermitian matrix  $R = \sqrt{\rho} \tilde{\rho} \sqrt{\rho}$  (or equivalently of the non-Hermitian matrix  $R' = \rho \tilde{\rho}$ ). Here  $\tilde{\rho}$  denotes the spin flipped state

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y) \rho^* (\sigma_y \otimes \sigma_y), \quad (6.92)$$

where the complex conjugation is taken in the computational basis of the two qubits.

For pure states, it can be shown that the concurrence (6.91) reduces to a very simple formula, given by

$$C(|\psi\rangle) = |\langle\psi|\tilde{\psi}\rangle|, \quad (6.93)$$

where  $|\tilde{\psi}\rangle = (\sigma_y \otimes \sigma_y)|\psi^*\rangle$  denotes the analogue of Eq. (6.92) for a pure state. As before,  $|\psi^*\rangle$  is the complex conjugate of  $|\psi\rangle$ , where the complex conjugation is expressed in the standard eigenbasis of  $\sigma_z$ . We can write this overlap explicitly, starting from the expression for a generic pure superposition of two qubits,

$$|\psi\rangle = \alpha_1|00\rangle + \alpha_2|01\rangle + \alpha_3|10\rangle + \alpha_4|11\rangle, \quad \text{with} \quad \sum_i |\alpha_i|^2 = 1, \quad (6.94)$$

which leads to the reduced density matrix of the first qubit:

$$\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \begin{bmatrix} |\alpha_1|^2 + |\alpha_2|^2 & \alpha_1^* \alpha_3 + \alpha_2^* \alpha_4 \\ \alpha_1 \alpha_3^* + \alpha_2 \alpha_4^* & |\alpha_3|^2 + |\alpha_4|^2 \end{bmatrix}. \quad (6.95)$$

Inserting Eq. (6.94) into the definition (6.93), it is thus easy to see that

$$C(|\psi\rangle) = 2|\alpha_1\alpha_4 - \alpha_2\alpha_3| = 2\sqrt{\det\rho_A}. \quad (6.96)$$

Note that the concurrence is a monotonic function of the entanglement of formation. For this reason it is very frequently used to characterize the entanglement, rather than of  $E_F(\rho)$ . However it should be stressed that, strictly speaking, the concurrence is not an entanglement measure, and that it acquires its meaning only because of its relation to the entanglement of formation.

### 6.7.2 \* Negativity

The properties of entangled states under partial transposition lead to an operative measure of entanglement, which is very simple to calculate. Specifically, in Sec. 6.2.1 we have seen that the partial transpose of a given mixed state  $\rho_{AB}$  is a very helpful

tool to decide on its separability: if at least one of the eigenvalues of  $\rho_{AB}^{T_A}$  is negative, then  $\rho_{AB}$  is entangled. This leads to the following definition of *negativity*:

$$N(\rho_{AB}) = - \sum_{\lambda_i < 0} \lambda_i, \quad (6.97)$$

where  $\lambda_i$  are the eigenvalues of  $\rho_{AB}^{T_A}$ . This quantity has been introduced in Życzkowski *et al.* (1998), where it was also shown that the set of separable states has a positive volume. From the definition of Eq. (6.97) it follows that, if  $\rho_{AB}^{T_A}$  is positive definite, then  $N = 0$ . Conversely,  $N$  becomes greater than zero if  $\rho_{AB}^{T_A}$  has one or more negative eigenvalues.

While being a convex entanglement monotone, the negativity defined in Eq. (6.97) is not additive. A more suitable choice for an entanglement monotone is the so-called *logarithmic negativity*, defined as

$$E_N(\rho_{AB}) = \log \left[ \text{Tr} \sqrt{(\rho_{AB}^{T_A})^\dagger (\rho_{AB}^{T_A})} \right]. \quad (6.98)$$

The major practical advantage of  $E_N$  is that it can be calculated very easily, since it is an algebraic function of the spectrum of  $\rho_{AB}^{T_A}$ . In addition it also has various operational interpretations, since it has been proven to be an upper bound for the distillable entanglement.

Unfortunately the negativity assigns non-vanishing entanglement only to those states that are detected via their negative partial transpose. Therefore, as for the Peres separability criterion, negativity is fully reliable only for bipartite  $2 \times 2$  and  $2 \times 3$  systems.

## 6.8 \* Multipartite entanglement

Although the bipartite scenario provides a variety of interesting and non-trivial situations for quantum entanglement, the multipartite setting allows to explore a much wider range of circumstances. We will see that, already at the level of basic properties, there is a crucial difference with respect to the bipartite setting. This generally makes the quantification of multipartite entanglement a very difficult problem.

In order to grasp such differences, we start from a simple example and consider three qubits in a pure state. This system may exhibit different forms of entanglement, involving two or three qubits; these are exemplified by the states

$$|\text{GHZ}\rangle_3 = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle) \quad (6.99)$$

$$|\text{W}\rangle_3 = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle), \quad (6.100)$$

representing the so-called GHZ-state and the W-state for the tripartite scenario (we adopt the usual computational basis of the three qubits). The essential difference between these states becomes clear after performing a measurement along the  $z$ -axis of one of the three qubits. In the case of the GHZ-state (6.99), measuring any one of

the qubits completely destroys the entanglement between the other two qubits (e.g., if we measure one qubit and obtain the result 0, the state of the other two qubits collapses into  $|00\rangle$ ). We call this an “essential three-way entanglement”. Conversely, if we measure any one qubit of the W-state (6.100) and get the result 0, the other two qubits will collapse into the maximally entangled EPR pair  $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ . This type of entanglement is called “pairwise entanglement”.

Let us now go back to the basic properties of entanglement and discuss the possibility to define, for multipartite states, the equivalent of a bipartite maximally entangled state. We focus again on qubit systems where, in the bipartite setting, maximally entangled states correspond to an EPR pair like  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . A natural choice for a  $n$ -partite state with this property would be precisely the GHZ-state for  $n$  qubits:

$$|GHZ\rangle_n = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}). \quad (6.101)$$

This state has the appealing property that its entropy of entanglement across any bipartite cut assumes the largest possible value of 1 ebit. However, there are entangled states that cannot be obtained from the GHZ-state using LOCC alone. It can be demonstrated that a W-state represents one of such cases (Dür *et al.*, 2000). This argument shows that it is not possible to establish a generic notion of a maximally entangled state, and more involved processes need to be invoked. Even in the asymptotic setting of arbitrarily many identically prepared states, it can be shown that this fact forbids the establishment of a reversible interconversion of quantum states in terms of LOCC.

Another fundamental issue about multipartite states concerns their separability. The most natural definition would be the one given in Eq. (6.3); however there are states which are not separable according to such definition, but still do not exhibit genuine multipartite entanglement. For example, in a tripartite scenario, the state  $\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) \otimes |0\rangle_C$  exhibits maximal bipartite entanglement between the parties  $A$  and  $B$ , while the third party  $C$  is uncorrelated from the other two. For a system with  $n$  parties, it is thus possible to define  $k$ -entangled states. For example, in a tripartite system, one could define the set of 2-entangled states as composed of any  $\rho_{ABC}$  that may be written according to

$$\rho_{ABC} = \sum_k p_k \rho_A^{(k)} \otimes \rho_{BC}^{(k)} + \sum_k q_k \rho_B^{(k)} \otimes \rho_{CA}^{(k)} + \sum_k r_k \rho_C^{(k)} \otimes \rho_{AB}^{(k)}, \quad (6.102)$$

with real positive numbers  $p_k, q_k, r_k$ .

All the problems outlined above clearly indicate that, in the multipartite setting, it is generally very difficult to find an unambiguous definition of entanglement measures. In the bipartite setting, such definition was univocal only for pure states; here the situation is complex even in that case. Below we just give a flavour of the issues that may rise, focusing on pure states.

### 6.8.1 \* Monogamy of entanglement and tangle measures

One of the most striking properties of entanglement is the so-called *monogamy* which, in its simplest statement, can be expressed as follows. Let us consider three parties (hereafter labelled as  $A$ ,  $B$ , and  $C$ ), having the same Hilbert space dimension. If two parties, say  $A$  and  $B$ , are very entangled, then a third party, say  $C$ , can only be weakly entangled with either  $A$  or  $B$ . For example, if each of the three parties is a qubit and two of them are in an EPR state, then they cannot be entangled with the other qubit at all. Note that this property is genuinely quantum: in the classical world, if two bits are perfectly correlated, there is no reason to have constraints on the correlations between them and another bit.

The above idea has been formalized by Coffman *et al.* (2000), using the following argument. Given a pure state of three qubits, we want to understand how the concurrence  $C_{AB}$  between  $A$  and  $B$  is related to the concurrence  $C_{AC}$  between  $A$  and  $C$ . In this case, the general formula for the concurrence, Eq. (6.91), simplifies: since each pair of qubits may be entangled with only another qubit in a joint pure state, the reduced density matrix of the pair has at most two nonzero eigenvalues. It follows that the matrix  $R'_{AB} = \rho_{AB} \tilde{\rho}_{AB}$  also has only two nonzero eigenvalues, say  $\lambda_1$  and  $\lambda_2$ . Therefore the following inequality for the concurrence  $C_{AB}$  holds:

$$C_{AB}^2 = (\lambda_1 - \lambda_2)^2 = (\lambda_1^2 + \lambda_2^2) - 2\lambda_1\lambda_2 \leq \text{Tr}(\rho_{AB} \tilde{\rho}_{AB}), \quad (6.103)$$

where the last step follows from the fact that  $\text{Tr}(\rho_{AB} \tilde{\rho}_{AB}) = \lambda_1^2 + \lambda_2^2$ . An analogous inequality can be proven for  $C_{AC}^2$ , so that we obtain

$$C_{AB}^2 + C_{AC}^2 \leq \text{Tr}(\rho_{AB} \tilde{\rho}_{AB}) + \text{Tr}(\rho_{AC} \tilde{\rho}_{AC}). \quad (6.104)$$

The right hand side of the latter expression can be evaluated explicitly, starting from the generic pure state superposition of the three-qubit system in the computational basis, and working out the spin flip operation on such state. We do not detail all the steps (they are found in Coffman *et al.*, 2000), and give only the result which, after some algebra, can be cast according to

$$\text{Tr}(\rho_{AB} \tilde{\rho}_{AB}) + \text{Tr}(\rho_{AC} \tilde{\rho}_{AC}) = 4 \det \rho_A. \quad (6.105)$$

This expression can be interpreted as follows. Identifying the pair  $BC$  as a single object, it is possible to introduce a notion of concurrence  $C_{A(BC)}$  between the qubit  $A$  and the pair  $BC$ , analogously to the discussion leading to Eq. (6.91). Indeed, as stated above, we recall that the reduced state  $\rho_{BC}$  has at most two non-zero eigenvalues, since  $A$  is only a qubit and the state of the whole system  $ABC$  is pure. In this context, one may thus effectively treat  $A$  and  $BC$  as a pair of qubits in a pure state. As a consequence, from Eq. (6.96) it follows that the corresponding bipartite concurrence is  $C_{A(BC)} = 2\sqrt{\det \rho_A}$ .

Combining the two expressions (6.104) and (6.105), and using the definition of  $C_{A(BC)}$ , we eventually get the following inequality:

$$C_{AB}^2 + C_{AC}^2 \leq C_{A(BC)}^2, \quad (6.106)$$

which quantifies the concept of entanglement monogamy. This constraint also allows to define the following entanglement quantifier known as the *three-tangle* (or the *residual tangle*).

$$\tau_{ABC} = C_{A(BC)}^2 - C_{AB}^2 - C_{AC}^2. \quad (6.107)$$

Such quantity is a measure of the purely tripartite entanglement. We note that the inequality (6.106) can be extended to  $n$  qubits, thus admitting a generalization of the concept of tangle for an arbitrary  $n$ -partite scenario.

## 6.9 Quantum discord

We have seen before that an operational way to define the genuine quantum character of a physical system is through testing Bell's inequalities. Specifically, any violation of such inequalities excludes the possibility of a local and realistic description of natural phenomena. Bell's inequalities require the presence of entanglement in order to exceed the classically determined limit for correlations. As a consequence, entangled states are correlated in a way that is inaccessible to classical objects. Following what we have discussed previously, one may regard as classical those states which can be prepared with the help of LOCC. According to this notion, the set of classical states coincides with that of separable (not entangled) quantum states; that is, for the case of bipartite systems,

$$\rho_{AB}^{\text{sep}} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}, \quad (6.108)$$

with  $p_k \geq 0$  and  $\sum_k p_k = 1$ , is separable. Conversely, states which cannot be written as  $\rho_{AB}^{\text{sep}}$  possess quantum correlations which correspond to entanglement.

However, it is possible to identify some separable states as quantum correlated, yet not entangled. Broadly speaking, differently from the pure classical world, we can define various classes of bipartite states having a different degree of correlations. States having *classical-classical* types of correlations have the form

$$\rho_{AB}^{\text{c-c}} = \sum_{i,j} p_{i,j} (|i\rangle_{AA}\langle i|) \otimes (|j\rangle_{BB}\langle j|) \quad (\text{classical-classical}), \quad (6.109)$$

where  $|i\rangle_A$  and  $|j\rangle_B$  are orthonormal bases in the Hilbert spaces of  $A$  and  $B$ , respectively. Analogously, states having *quantum-classical* or *classical-quantum* types of correlations are defined by

$$\rho_{AB}^{\text{q-c}} = \sum_j p_j \rho_A^{(j)} \otimes (|j\rangle_{BB}\langle j|) \quad (\text{quantum-classical}), \quad (6.110)$$

$$\rho_{AB}^{\text{c-q}} = \sum_i p_i (|i\rangle_{AA}\langle i|) \otimes \rho_B^{(i)} \quad (\text{classical-quantum}). \quad (6.111)$$

Following the alternative notion of classicality proposed above, which goes beyond the LOCC framework, we can infer that classical states would be given by a subset of  $\rho_{AB}^{\text{sep}}$  which, as we shall see later, can be written as  $\rho_{AB}^{\text{c-q}}$ ,  $\rho_{AB}^{\text{q-c}}$  or  $\rho_{AB}^{\text{c-c}}$ ,

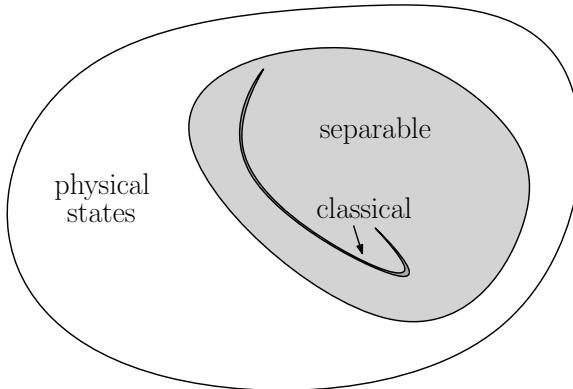


Fig. 6.6 Pictorial sketch of generic bipartite quantum states in a Hilbert space. The set of classical (zero-discord) states turns out to be a non-convex subset (with a smaller dimensionality) of the set of separable (non-entangled) states, which is in turn a subset of the full Hilbert space of the system.

depending on which subsystem(s) is being measured. We note that the subset of classical states is non-convex, since mixing two classically correlated states one may obtain a state which is not classically correlated anymore, and that it is of null measure, being nowhere dense within the subset of separable states (Ferraro *et al.*, 2010). The typical scenario is depicted in Fig. 6.6. For the purpose of identifying separable states which are not classically correlated, we are now going to construct the so-called *quantum discord*, namely a measure which generalizes the concept of quantum correlations beyond entanglement, since it takes into account any non-classical source of correlations.

### 6.9.1 Definition

Quantum discord, which was introduced by Ollivier and Zurek (2001) and by Henderson and Vedral (2001), relies upon the fact that two systems are correlated if, once they are put together, they contain more information than taken separately. As we have seen in Sec. 6.3.1, classically this is captured by the mutual information. In the quantum context, it is crucial to observe that it is not obvious to generalize the concept of conditional entropy measuring the residual ignorance about a given subsystem, once the other subsystem has been already measured. Indeed it is not sufficient to replace the Shannon entropy with the von Neumann entropy of the corresponding density matrices: one also requires to specify the basis over which the (quantum) measurement of the subsystem is performed.

Given the two classical random variables  $X$  and  $Y$  associated with Alice's and Bob's readout of their states, their mutual information is given by

$$\mathcal{I}(X:Y) = H(X) + H(Y) - H(X, Y), \quad (6.112)$$

$H(X)$  being the Shannon entropy. Using Bayes rule for classical variables, a conditional probability can be defined in such a way that the mutual information between

$X$  and  $Y$  can be equivalently given by Eq. (6.34), that is,

$$\mathcal{J}(X:Y) = H(X) - H(X|Y), \quad (6.113)$$

where  $H(X|Y)$  denotes the conditional entropy (6.29). Obviously  $\mathcal{I}(X:Y) = \mathcal{J}(X:Y)$ . Classical correlations can thus be interpreted as information gain about one subsystem as a result of the measurement on the other (see Sec. 6.3). If we now try to export  $\mathcal{J}(X:Y)$  to the quantum context, we will soon realize that quantum variables do not have a well defined value, since the latter substantially depends on the measurement process. Indeed different measurements can be performed on the quantum system, and measurements generally disturb the quantum state.

### Quantum mutual information

Let us proceed step by step and consider a bipartite quantum system in the Hilbert space  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ , described by the density matrix  $\rho_{AB}$  (as we shall see later, the pure case trivially reduces to the common notion of entanglement). The mutual information (6.112) is straightforwardly generalized by replacing the Shannon entropy of the classical probability distributions by the von Neumann entropy of the corresponding appropriate density matrices. We thus obtain the *quantum mutual information*:

$$\mathcal{I}(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}), \quad (6.114)$$

where  $\rho_A$  and  $\rho_B$  respectively denote Alice's and Bob's reduced density matrices, obtained after partially tracing the global state  $\rho_{AB}$ . In this formula,  $S(\rho_A) + S(\rho_B)$  represents the uncertainty of Alice and Bob subsystems as treated separately, while  $S(\rho_{AB})$  is the uncertainty of the composite system. Due to subadditivity of the von Neumann entropy, the mutual information (6.114) is positive, while it vanishes for states that can be factorized in a tensor product. Non-zero values of  $\mathcal{I}(A:B)$  are hence due to the presence of general correlations within the considered state, being them entanglement-like or just classical. Note also that, by definition, the formula (6.114) is symmetric:  $\mathcal{I}(A:B) = \mathcal{I}(B:A)$ .

### Classical correlations

The quantum generalization of Eq. (6.113) is not automatic, since the conditional entropy  $H(A|B)$  requires us to specify the state of Alice, given the state of Bob after a measurement has been performed. Such statement in quantum theory is ambiguous, until Bob's measurement is specified. We restrict, for the moment, to one-dimensional orthogonal projectors on Bob's subsystem, so that we can specify a measurement basis on  $\mathcal{H}$  by means of the following operators:  $\{I_A \otimes [\Pi_j]_B\}_{j=1\dots k}$ , where  $[\Pi_j]_B = |j\rangle_B \langle j|$  and  $\{|j\rangle\}$  defines an orthonormal basis in  $\mathcal{H}_B$ . Taking the one-dimensional projectors  $\{[\Pi_j]_B\}$ , the initial state  $\rho_{AB}$  after Bob's (projective) measurement is given by

$$\rho_{AB}|_{[\Pi_j]_B} = \frac{\{I_A \otimes [\Pi_j]_B\} \rho_{AB} \{I_A \otimes [\Pi_j]_B\}}{\text{Tr} [\{I_A \otimes [\Pi_j]_B\} \rho_{AB}]}, \quad (6.115)$$

where Bob's projection over the state  $|j\rangle_B$  is obtained with a given probability  $p_j = \text{Tr}[\{I_A \otimes [\Pi_j]_B\} \rho_{AB}]$ . The conditional state of Alice is thus

$$\rho_{A|\Pi_j} = \text{Tr}_B[\rho_{AB}|\Pi_j], \quad (6.116)$$

where we have omitted the index  $B$  over the measure. This allows us to define the classical-quantum version of the conditional entropy:

$$\mathcal{J}(A:B)|_{\{\Pi_j\}} = S(\rho_A) - S(\rho_{AB}|\{\Pi_j\}) = S(\rho_A) - \sum_j p_j S(\rho_{A|\Pi_j}), \quad (6.117)$$

which evidently depends of the choice of the projectors  $\{\Pi_j\}$ .

Expression (6.117) defines the *classical correlations* of the state  $\rho_{AB}$  and represents the information gained about Alice's subsystem ( $A$ ) as a result of the measurement  $\{[\Pi_j]_B\}$  on Bob's subsystem ( $B$ ). Conversely, one can define  $\mathcal{J}(B:A)|_{\{\Pi_i\}}$  as the information gained about Bob's subsystem ( $B$ ) as a result of the measurement  $\{[\Pi_i]_A\}$  on Alice's subsystem ( $A$ ). Notice that in general  $\mathcal{J}(A:B) \neq \mathcal{J}(B:A)$ , namely  $\mathcal{J}$  depends on which subsystem is measured.

### Quantum discord

The quantum discord of a state  $\rho_{AB}$  under Bob's projective measurements is defined as the difference between the quantum mutual information and the optimal classical correlations maximized over all the (projective) measures:

$$\mathcal{D}(A:B) = \mathcal{I}(A:B) - \mathcal{J}(A:B), \quad (6.118)$$

where

$$\mathcal{J}(A:B) = S(\rho_A) - \min_{\{\Pi_j\}} S(\rho_{AB}|\{\Pi_j\}). \quad (6.119)$$

As already pointed out above, since  $\mathcal{I}(A:B) \geq 0$  is saturated to the null value only for tensor product states, the mutual information can be considered a measure of the global amount of correlations between the parts of the system. This interpretation is confirmed by the intuitive idea that the evaluation of such quantity does not alter the considered state. On the contrary,  $\mathcal{J}(A:B)$  only captures the classical fractions of such correlations, which can be acquired via a measurement on  $B$ . Such measure modifies the quantum system, destroying some of the initial correlations that in this way cannot be measured. This fact leads to an interpretation of the quantum discord as a measure of *purely quantum correlations*.

We remark that the definition of quantum discord can be straightforwardly generalized to generic positive operator-valued measurements (POVMs) with elements  $F_j = M_j^\dagger M_j$ , where  $M_j$  is the measurement operator associated to the  $j$ -th outcome (see Sec. 2.9.1). Now suppose that Bob performs the above POVM on the initial state  $\rho_{AB}$ . As a consequence to that, he will observe outcome  $j$  with probability  $p_j = \text{Tr}[\{I_A \otimes [F_j]_B\} \rho_{AB}]$  and the conditional state of Alice will be

$$\rho_{A|[F_j]_B} = \frac{1}{p_j} \text{Tr}_B[\{I_A \otimes [F_j]_B\} \rho_{AB}]. \quad (6.120)$$

In analogy to Eqs. (6.117) and (6.119), by summing over all possible outcomes, this allows to define the classical correlations  $\mathcal{J}(A:B)$  and eventually the quantum discord of  $\rho_{AB}$  under Bob's POVM. However it has to be stressed that, in most cases, the minimization leading to optimal classical correlations can be approximated very well by considering projective measurements. More precisely, it can be rigorously shown that it is sufficient to consider rank-one POVMs, that is, POVMs with only rank-one elements which are proportional to projectors (but need not to be orthogonal) and which are linearly independent (Datta, 2008). For this reason, from now on we will only address the quantum discord under projective measurements.

### 6.9.2 Basic properties

The quantum discord (6.118) possesses the following properties:

- (i) It is not symmetric:

$$\mathcal{D}(A:B) \neq \mathcal{D}(B:A). \quad (6.121)$$

This property follows from the fact that the conditional entropy  $\mathcal{J}(A:B)$  itself is not symmetric, since it involves a measurement on one end (in our case Bob, allowing the observer to infer the state of Alice).

- (ii) It is a non-negative quantity:

$$\mathcal{D}(A:B) \geq 0. \quad (6.122)$$

This is a consequence of the concavity of conditional entropy.

- (iii) It vanishes if and only if the state is quantum-classical (or classical-quantum, depending on which subsystem is measured):

$$\mathcal{D}(A:B) = 0 \iff \rho_{AB} \in \rho_{AB}^{q-c}, \quad (6.123)$$

$$\mathcal{D}(B:A) = 0 \iff \rho_{AB} \in \rho_{AB}^{c-q}. \quad (6.124)$$

This condition implies separability if the state considered is pure.

**Proof.** Let us focus on the first statement. The equality  $\mathcal{D}(A:B) = 0$  means that, in order to detect some non-classicality by measuring subsystem  $B$ , the state has to be written as a superposition of some orthogonal projectors in  $BA$ . However, the conditional states  $\rho_{A|\Pi_j}$  left in  $A$  at the end of the measuring process could still present quantum features, e.g., non commutativity. We are thus left with a quantum-classical state  $\rho_{AB}^{q-c}$ , see Eq. (6.110). For this reason, having a zero  $\mathcal{D}(A:B)$ , gives information only about classicality in the measured subsystem. Analogously, if a measurement on subsystem  $A$  is considered, a state with null discord  $\mathcal{D}(B:A)$  will have the form of a classical-quantum state  $\rho_{AB}^{c-q}$  — see Eq. (6.111).  $\square$

- (iv) It is bounded from above by the entropy:

$$\mathcal{D}(A:B) \leq S(A). \quad (6.125)$$

- (v) It remains unchanged by performing local unitary transformations  $U_A \otimes U_B$ , with  $U_A$  and  $U_B$  arbitrary unitaries acting on systems  $A$  and  $B$ . This property follows from the fact that for the von Neumann entropy  $S(\rho_{AB}) = S[(U_A \otimes U_B)\rho_{AB}(U_A \otimes U_B)^\dagger]$ .
- (vi) For pure states, it is symmetric and equivalent to the unique measure of entanglement, that is, the entropy of entanglement. This confirms that, to find quantum features that cannot be described by the mere entanglement, mixed states have to be taken into account.

**Proof.** Let us consider a pure state  $\rho_{AB} = |\psi\rangle\langle\psi|$ . Once a projective measurement  $\Pi_j$  is performed on  $B$ , this collapses into

$$\rho_{AB|\Pi_j} = \frac{\Pi_j \rho_{AB} \Pi_j}{\text{Tr}(\Pi_j \rho_{AB})} = \frac{\Pi_j |\psi\rangle\langle\psi| \Pi_j}{\langle\psi|\Pi_j|\psi\rangle}, \quad (6.126)$$

where we omitted to indicate the identity operator on  $A$ . From this, using the fact that  $(\Pi_j)^2 = \Pi_j$ , it follows that

$$(\rho_{AB|\Pi_j})^2 = \frac{\Pi_j |\psi\rangle\langle\psi| \Pi_j \Pi_j |\psi\rangle\langle\psi| \Pi_j}{\langle\psi|\Pi_j|\psi\rangle \langle\psi|\Pi_j|\psi\rangle} = \frac{\Pi_j |\psi\rangle\langle\psi| \Pi_j}{\langle\psi|\Pi_j|\psi\rangle} = \rho_{AB|\Pi_j}. \quad (6.127)$$

Therefore  $\text{Tr}_{AB}(\rho_{AB|\Pi_j})^2 = \text{Tr}_{AB}(\rho_{AB|\Pi_j}) = 1$ , and thus  $\rho_{AB|\Pi_j}$  is a pure state. The expression for the quantum discord becomes:

$$D(A:B) = S(\rho_B) - S(\rho_{AB}) + \min_{\{\Pi_j\}} \sum_j p_j S(\rho_{A|\Pi_j}) = S(\rho_B), \quad (6.128)$$

which coincides with the entanglement entropy for a pure state.  $\square$

### 6.9.3 Examples

The optimization procedure involved in computing the quantum discord (6.118) makes it a challenge to evaluate it for general states, similar to the problem of calculating the entanglement of formation. Analytical results are known only for few specific cases, and typically studies of discord heavily rely on numerical optimizations for determining the measurement basis that maximizes the classical correlations. Hereafter we will only focus on specific cases of two-qubit systems.

We start with an instructive example where the difference between separability and vanishing discord can be illustrated. Let us consider the Werner state  $(\rho_W)_{AB}$  as defined in Eq. (6.7), such that  $(\rho_W)_A = (\rho_W)_B = \text{Tr}_B[(\rho_W)_{AB}] = \frac{1}{2}I$ . The eigenvalues of the density matrix  $(\rho_W)_{AB}$ , written in the usual computational basis as in Eq. (6.8), are  $\lambda_1 = \lambda_2 = \lambda_3 = q_-/2$  and  $\lambda_4 = (q_+ + q)/2$ , where we defined  $q_\pm = \frac{1}{2}(1 \pm q)$ . Therefore the mutual information of this state is  $I(A:B) = 2 - 3h(q_-/2) - h(q_+/2 + q/2)$ , with  $h(x) = -x \log x$ . The classical correlations do not depend on the choice of the basis for the measurements of  $B$ , since both the identity  $I$  and the state  $|\psi^-\rangle$  are invariant under local rotations. The minimization in  $J(A:B)$  is thus straightforward to be computed. Let us consider the measurement

of  $B$  along the computational basis, that is  $\Pi_a = |a\rangle\langle a|$  ( $a = 0, 1$ ), so that the conditional states of Alice are:

$$\rho_{A|\Pi_a} = \frac{1}{4}(1-q)I + \frac{q}{2}|1-a\rangle\langle 1-a|, \quad (6.129)$$

with  $p_a = \frac{1}{2}$  for  $a = 0, 1$ . The classical correlations are thus given by  $\mathcal{J}(A:B) = 1 - h(q_+) - h(q_-)$  and thus the quantum discord is

$$\mathcal{D}(A:B) = 1 + h(q_+) + h(q_-) - 3h(q_-/2) - h(q_+/2 + q/2), \quad (6.130)$$

which is greater than 0 each time  $q > 0$ , as shown in Fig. 6.7. This has to be contrasted with the separability of such states when  $q < \frac{1}{3}$  (see exercise 6.1).

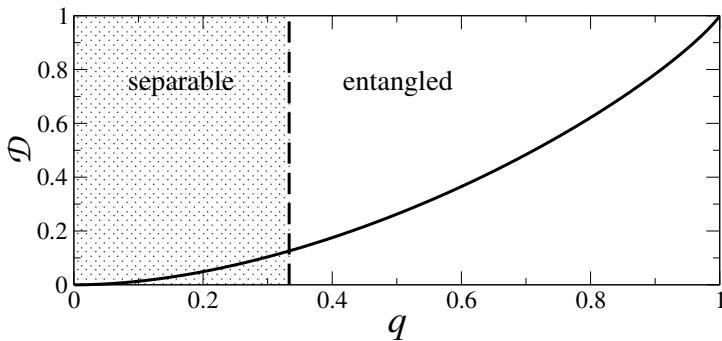


Fig. 6.7 Quantum discord for the Werner states  $(\rho_W)_{AB}$  of Eq. (6.7). The vertical dashed line distinguishes between separable states, for  $0 \leq q \leq \frac{1}{3}$ , and entangled states, for  $\frac{1}{3} < q \leq 1$ .

**Exercise 6.12** Calculate the quantum discord for a system of two qubits in the following state:  $\rho_{AB}(\xi) = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) + \frac{\xi}{2}(|00\rangle\langle 11| + |11\rangle\langle 00|)$ , where  $0 \leq \xi \leq 1$ .

#### 6.9.4 \* Other measures of quantum correlations

Apart from the discord (6.118), several other measures of purely quantum correlations have been introduced, in an effort to gain a better operational understanding on the subject (a process which closely reminds what happened in the characterization of entanglement). This proliferation also stems from the difficulty of identifying a measure which is at the same time well defined and easily computable (as we have seen in Sec. 6.9.3, the feasibility of practical computation may become challenging even for bipartite two-qubit systems).

The fact that the discord vanishes on classical-quantum states has been taken as a starting point to define geometric measures of correlations, characterizing non-classical features of a state  $\rho_{AB}$  by its distance from the set of classical-quantum states, in analogy with the geometric measures of entanglement defined in terms of distances from the set of separable states (see Sec. 6.2.3). There are also alternative non-geometric approaches to quantum correlations, as those which consider

perturbations induced by local von Neumann measurements on non-classically correlated states, or which investigate the role of correlations in work extraction from a heat bath. Here we will not enter the details of these alternative approaches. A list of definitions is discussed in Modi *et al.* (2012) and in Adesso *et al.* (2016). Here we just mention that, to be considered a valid measure of quantum correlations, a given functional of the bipartite density matrix,  $\mathcal{M}(\rho_{AB})$ , should satisfy the following properties:

- (i) it must vanish if and only if  $\rho_{AB}$  possesses a kind of classical correlations:

$$\mathcal{M}(\rho_{AB}) = 0 \iff \rho_{AB} \in \rho_{AB}^{\text{c-q}}, \rho_{AB}^{\text{q-c}}, \rho_{AB}^{\text{c-c}}, \quad (6.131)$$

depending on whether classicality has to be evaluated for a single subsystem  $A$  or  $B$ , or for both of them (faithfulness criterion);

- (ii) it must be invariant under local unitary transformations:

$$\mathcal{M}(\rho_{AB}) = \mathcal{M}[(U_A \otimes U_B)\rho_{AB}(U_A \otimes U_B)^\dagger]; \quad (6.132)$$

- (iii) if  $\rho_{AB}$  is a pure state, it must reduce to an entanglement monotone; that is,

$$|\psi\rangle_{AB} \xrightarrow{\text{LOCC}} |\phi\rangle_{AB} \implies \mathcal{M}(|\phi\rangle_{AB}\langle\phi|) \leq \mathcal{M}(|\psi\rangle_{AB}\langle\psi|). \quad (6.133)$$

This property is required because on pure states each measure should quantify the entanglement (there cannot be correlations of a different kind without classical mixing), which is decreasingly monotone under LOCC;

- (iv) if only the quantumness of subsystem  $A$  is being probed, a desirable property is that  $\mathcal{M}(\rho_{AB})$  should be non-increasing under any local operation on the unmeasured subsystem  $B$ . Conversely, a local operation on  $A$  can in general increase its quantumness, since orthogonal states can be mapped into non-orthogonal ones.

### Geometric discord

As a paradigmatic example of the geometric approach to correlations, let us discuss the *geometric discord*, which has been first proposed by Dakić *et al.* (2010). The advantage of this measure is that the minimization involved in the definition can be performed much more easily than that of the original quantum discord (6.118). Unfortunately this comes at the price of violating the above property (iv), so that more involved geometric alternatives need to be invoked to overcome this hindrance. The geometric discord is defined as:

$$\mathcal{D}_G(\rho) = \min_{\sigma \in CQ} \|\rho - \sigma\|_2^2, \quad (6.134)$$

where  $\|O\|_2 = \sqrt{\text{Tr}[O^\dagger O]}$  is the operatorial norm, and  $CQ$  indicates the set of quantum-classical states, Eq. (6.110).

It turns out that, for pure states, one has:

$$\mathcal{D}_G(|\psi\rangle) = 1 - \text{Tr}(\rho_A^2) = \frac{1}{2} [C(|\psi\rangle)]^2, \quad (6.135)$$

where  $C(|\psi\rangle)$  denotes the concurrence and last equality follows from Eq. (6.96). For pure states the geometric distance is thus an entanglement monotone. Remarkably, it is even possible to compute  $\mathcal{D}_G$  for a generic two-qubit state,

$$\rho_{AB} = \frac{1}{4} \sum_{\mu=0}^3 \sum_{\nu=0}^3 T_{\mu\nu} (\sigma_\mu \otimes \sigma_\nu), \quad (6.136)$$

where  $\sigma_\mu = \{I, \sigma_x, \sigma_y, \sigma_z\}$  are the identity and the Pauli matrices on one qubit, and the prefactors  $T_{\mu\nu} = \text{Tr}[\rho_{AB}(\sigma_\mu \otimes \sigma_\nu)]$  can be chosen to be real numbers, with  $T_{\mu\nu} \in [-1, 1]$ . The geometric discord can be shown to be

$$\mathcal{D}_G = \frac{1}{4} \left( \sum_{j=1}^3 \sum_{\nu=0}^3 T_{j\nu}^2 - \lambda_{\max} \right), \quad (6.137)$$

where  $\lambda_{\max}$  is the largest eigenvalue of the matrix

$$L = \vec{x}_A \vec{x}_A^T + \vec{T} \vec{T}^T, \quad (6.138)$$

built from the local Bloch vector  $\vec{x}_A = (T_{10}, T_{20}, T_{30})^T$  and the three-dimensional correlation matrix  $[\vec{T}]_{jk} = T_{jk}$ , with  $j, k = 1, 2, 3$ .

**Exercise 6.13** Calculate the geometric discord for the two-qubit Werner state of Eq. (6.7).

## 6.10 \* Quantum discord in continuous systems

In discrete quantum systems, finding states that possess a finite discord is relatively simple, since such kind of systems (such as, for example, two qubits) are very far from any classical description. The situation however may drastically change if we consider continuous variables. For the sake of simplicity, in this section we will focus on Gaussian states and measures (for generic states in the Hilbert space, the theoretical framework is more complicated, and there are several issues which are not well understood yet). We will show that even in the case of Gaussian states, which display features that are much closer to classical states than those of discrete systems, it is possible to highlight the existence of inherently quantum properties such as a non-zero discord. Before addressing this issue, we first need to evaluate the von Neumann entropy of an arbitrary Gaussian state.

### 6.10.1 \* Entropy of a Gaussian state

As we have seen in Sec. 5.6.1, Williamson's theorem enables us to decompose a generic  $N$ -mode Gaussian state  $\rho$  into a direct sum of single-mode thermal states, by means of a suitable symplectic transformation:

$$\rho = U_S \left[ \bigotimes_{j=1}^N \rho_{\text{th}}(\nu_j) \right] U_S^\dagger, \quad (6.139)$$

which is induced by the transformation of the corresponding covariance matrix in Eq. (5.171). In order to evaluate  $S(\rho)$ , it is useful to remind that the entropy is invariant under unitary transformations, so that we can first calculate the entropies  $S(\rho_{\text{th}})$  of each of the single-mode thermal states and eventually sum over all of them. Moreover, since a thermal state  $\rho_{\text{th}}(\bar{n}_j)$  is diagonal in its number-state representation [see Eq. (5.172)], the calculation drastically simplifies into:

$$S(\rho_{\text{th}}) = -\text{Tr}(\rho_{\text{th}} \log \rho_{\text{th}}) = -\sum_{n=0}^{+\infty} \frac{(\bar{n}_j)^n}{(\bar{n}_j + 1)^{n+1}} [n \log \bar{n}_j - (n+1) \log(\bar{n}_j + 1)]. \quad (6.140)$$

Using the expression for the geometric series  $\sum_{j=0}^{+\infty} z^j = (1-z)^{-1}$  (for  $|z| < 1$ ) and taking its derivative with respect to  $z$ , we can easily perform the summation over  $n$  and get

$$S(\rho_{\text{th}}) = -\bar{n}_j \log \bar{n}_j + (\bar{n}_j + 1) \log(\bar{n}_j + 1). \quad (6.141)$$

Eventually, writing it in terms of the symplectic eigenvalue  $\nu_j = 2\bar{n}_j + 1$ , we have:

$$S(\rho_{\text{th}}) = -\frac{\nu_j - 1}{2} \log\left(\frac{\nu_j - 1}{2}\right) + \frac{\nu_j + 1}{2} \log\left(\frac{\nu_j + 1}{2}\right) \equiv f(\nu_j), \quad (6.142)$$

where we defined the function  $f(\nu)$ , which will turn out to be useful in the following. The physical meaning of a thermal state  $\rho_{\text{th}}(\bar{n})$  is related to the fact that it is the state which maximizes the von Neumann entropy, for fixed number of bosons  $\bar{n}$ .

### 6.10.2 \* Discord of a Gaussian state

Hereafter we will restrict to a two-mode Gaussian state; the scenario depicted below can be easily generalized in a straightforward manner. Let us thus consider a bipartite Gaussian state into the modes  $A$  and  $B$  (with creation and annihilation operators  $\{a^\dagger, a\}$  and  $\{b^\dagger, b\}$  respectively, which can be defined by the first moments  $\vec{d}_A$ ,  $\vec{d}_B$  and the variances  $\Sigma_A$ ,  $\Sigma_B$ ). We also suppose that the measurement is performed on system  $B$  and is of Gaussian type, that is, when applied to Gaussian states it gives an outcome which is Gaussian distributed. Any Gaussian measurement of this kind can be accomplished using homodyne detection, linear optics, and Gaussian ancilla modes.

The Gaussian quantum discord is defined as the usual discord, where the conditional entropy is obtained by means of a generalized POVM on subsystem  $B$ . If we restrict to linear optics measures, the quantum discord is thus defined as

$$\mathcal{D}(A:B) = S(\rho_B) - S(\rho_{AB}) + \inf_{\{\Pi_\eta\}} \int d\eta p_\eta S(\rho_{A|\Pi_\eta}), \quad (6.143)$$

where  $\Pi_\eta$  is a Gaussian measurement on subsystem  $B$  and  $p_\eta$  is the probability to obtain the outcome  $\eta$  after the measure, while  $\rho_{A|\Pi_\eta}$  is the state of subsystem  $A$  after the Gaussian measurement on  $B$  has produced the outcome  $\eta$ . As we have seen in the previous section, the crucial point which characterizes the entropy of

a Gaussian state is that it depends only on its covariance matrix  $\Sigma$ , and  $\Sigma$  does not depend on the measurement outcome. Therefore, in Eq. (6.143) the entropy factorizes trivially, and the integral gives 1.

The projector of the Gaussian measure  $\Pi_\eta$  can be generally written as:

$$\Pi_\eta = \frac{1}{\pi} \exp(\eta b^\dagger - \eta^* b) \Pi_B^0 \exp(\eta^* b - \eta b^\dagger), \quad (6.144)$$

where  $\exp(\eta^* b - \eta b^\dagger)$  is the displacement operator on subsystem  $B$ , while  $\Pi_B^0$  is the density matrix of a generic single-mode Gaussian state (sometimes referred to as the “seed” of the measure). Denoting with  $\Sigma_B^0$  the covariance matrix of the seed  $\Pi_B^0$ , it is possible to show that the conditional state of subsystem  $A$  after measuring subsystem  $B$  has a covariance matrix (see Weedbrook *et al.*, 2012, and references therein):

$$\Sigma_{A|\Pi_\eta} = \Sigma_A - \Upsilon_{AB} (\Sigma_B + \Sigma_B^0)^{-1} (\Upsilon_{AB})^T, \quad (6.145)$$

where  $\Sigma_A$  and  $\Sigma_B$  are the covariance matrices of subsystems  $A$  and  $B$ , while  $\Upsilon_{AB}$  is the correlation matrix between subsystems  $A$  and  $B$  [see Eq. (5.173)]. Regarding the seed of the measure, here we limit ourselves to pure states. As a matter of fact, we are not interested in the most general case, but we only want to show that even for Gaussian states, it is possible to have non-zero discord.

The entropy of a generic Gaussian state can be calculated following the lines described in the previous section, that is, using Eq. (6.142) and then summing over all the symplectic eigenvalues. For a single-mode Gaussian state with covariance matrix  $\Sigma_A$ , the expression simplifies into

$$S(\rho_A) = f[\sqrt{\det(\Sigma_A)}], \quad (6.146)$$

where we remind the reader that the determinant of the covariance matrix is a symplectic invariant. Therefore, the discord (6.143) for a two-mode Gaussian state takes the form

$$\mathcal{D}(A:B) = f[\sqrt{\det(\Sigma_B)}] - f(\nu_+) - f(\nu_-) + \inf_{\Sigma_B^0} f\left[\sqrt{\det(\Sigma_{A|\Pi_\eta})}\right], \quad (6.147)$$

where  $S(\rho_{AB}) = f(\nu_+) + f(\nu_-)$ . In conclusion, the only optimization procedure to be performed is over the parametrization of the seed of the measure. Here we do not provide further details of the optimization; we just note that at this stage it is clear that, in general, we have a result which demonstrates the possibility to have a non-zero discord for a two-mode Gaussian state. The reader who is interested in the full calculation of  $\mathcal{D}(A:B)$ , with all the details, is referred to Adesso and Datta (2010) and to Giorda and Paris (2010).

## 6.11 \* Entropies in physics

The concept of entropy is very closely connected to those of energy, information and chaos. It is a fundamental concept in both information science and physics. There exist many entropy-like quantities. Here, we shall briefly describe those we

consider to be the most significant in physics (at least in relation to this book) while endeavouring to elucidate the possible links between the different definitions of entropy. We shall also discuss the relation between these entropies and the Shannon entropy.

### 6.11.1 \* Thermodynamic entropy

In order to define the thermodynamic entropy, consider first the integral

$$\int_A^B \frac{\delta Q}{T}, \quad (6.148)$$

extended over a *reversible transformation* from  $A$  to  $B$ , where  $A$  and  $B$  are two equilibrium states of a given system and  $\delta Q$  is the amount of heat absorbed *reversibly* by the system at temperature  $T$ . It can be proved that the above integral depends only on the initial and final states of the transformation, and not on the transformation itself; that is, it is the same for all reversible paths (transformations) joining  $A$  to  $B$ .

This property enables us to define a state function  $S(A)$ , known as the thermodynamic entropy.<sup>2</sup> The entropy  $S(A)$  of any equilibrium state  $A$  of the system is defined by

$$S(A) = \int_O^A \frac{\delta Q}{T}, \quad (6.149)$$

where the integration path is any reversible transformation from  $O$  to  $A$  and  $O$  is some chosen reference equilibrium state. Note that the entropy  $S(A)$  is only defined up to an arbitrary additive constant. Indeed, if we choose a different reference state  $O'$  instead of  $O$  and define  $S'(A) = \int_{O'}^A \frac{\delta Q}{T}$ , then  $S'(A) = S(A) + S'(O)$ . Therefore, the additive constant  $S'(O)$  is independent of the state  $A$ . The difference in the entropy of two states is, on the other hand, completely defined. We have

$$S(B) - S(A) = \int_A^B \frac{\delta Q}{T}. \quad (6.150)$$

It follows that, for any infinitesimal reversible transformation, the change in entropy is

$$dS = \frac{\delta Q}{T}. \quad (6.151)$$

Note that, in contrast to differently from  $\delta Q$ ,  $dS$  is an exact differential.

Nernst's theorem, also referred to as the third law of thermodynamics, allows us to determine the additive constant appearing in the definition of entropy. This theorem states that *the entropy of every system at absolute zero can always be taken equal to zero* (note that here we assume that the ground state of the system is non-degenerate). It is therefore convenient to choose the state of the system at  $T = 0$  as

---

<sup>2</sup>The thermodynamic entropy was introduced by Clausius in 1865.

the reference state in (6.149), so that its entropy is set equal to zero. The entropy of any equilibrium state  $A$  is now defined as follows:

$$S(A) = \int_{T=0}^A \frac{\delta Q}{T}. \quad (6.152)$$

Note that formula (6.152) is restricted to equilibrium states. However, for systems composed of several parts, it is possible to define the entropy even for non-equilibrium states, in the case in which each part is itself in an equilibrium state  $A_i$  with corresponding entropy  $S_i$ . The global entropy of the system is then given by the sum of the entropies of all the parts:  $S = \sum_i S_i$ .

An important property of entropy arises from Eq. (6.150). For a thermally isolated system (that is,  $\delta Q = 0$ ) reversible transformations do not change the entropy of the system:  $S(B) = S(A)$ . On the other hand, it is possible to show that for *irreversible transformations* we have

$$S(B) - S(A) \geq \int_A^B \frac{\delta Q}{T}. \quad (6.153)$$

Therefore, for  $\delta Q = 0$  we find

$$S(B) \geq S(A), \quad (6.154)$$

that is, for any transformation occurring in a thermally isolated system, the entropy of the final state can never be less than that of the initial state. Thus, the state of maximum entropy is the most stable state for an isolated system.

Let us consider a transformation from an initial state  $A$  to a final state  $B$  of a system in contact with an environment that is maintained at a constant temperature  $T$ . Applying Eq. (6.153), we obtain

$$Q = \int_A^B \delta Q \leq T[S(B) - S(A)]. \quad (6.155)$$

The first law of thermodynamics, see Eq. (1.36), tells us that the work  $W$  performed by the system is given by

$$W = -\Delta E + Q, \quad (6.156)$$

where  $\Delta E = E(B) - E(A)$  is the variation of the internal energy of the system. From Eqs. (6.155) and (6.156) we obtain

$$W \leq E(A) - E(B) + T[(S(B) - S(A))]. \quad (6.157)$$

This inequality sets an upper limit on the amount of work that can be extracted from the transformation  $A \rightarrow B$ . If such a transformation is reversible, then the equality sign holds and the work performed saturates the upper limit. It is useful to define the function

$$F = E - TS. \quad (6.158)$$

Then Eq. (6.157) becomes

$$W \leq F(A) - F(B) = -\Delta F. \quad (6.159)$$

For a reversible transformation we have  $W = -\Delta F$ . Therefore, the quantity  $F$ , known as the *free energy* of the system, plays a role analogous to that of the internal energy  $E$  in a purely mechanical system (indeed, in such a case,  $Q = 0$ , so that  $W = -\Delta E$ ).

### 6.11.2 \* Statistical entropy

One of the principal purposes of equilibrium statistical mechanics is to explain the laws of thermodynamics starting from the laws of molecular dynamics. The question is: given the laws of motion and the interactions between the molecules, what are the macroscopic properties of matter composed of these molecules?

In the second half of the nineteenth century, Boltzmann and Clausius tried to derive the second law of thermodynamics from mechanics. This followed the line of Maxwell, who had already put forward the idea that “the second law of thermodynamics has only a statistical certainty”.

As we discussed in the previous subsection, in a thermally isolated system entropy can never decrease. Thus, the system evolution is such that it never becomes more ordered. A familiar demonstration of this principle is the flow of heat from hot to cold bodies until a uniform temperature is reached.

Boltzmann related the notion of entropy to the logarithm of the number of possible different microscopic states compatible with a given macroscopic state. For instance, let us consider  $N/2$  white molecules and  $N/2$  black molecules ( $N \gg 1$ ) inside a single vessel and distinguish the microscopic state of each molecule by the fact that it is located in the left or right half of the vessel. It is clear that there is a single microscopic state corresponding to the macroscopic state “all white molecules in the left half and all black molecules in the right half of the vessel” while there are many more microscopic states corresponding to the macroscopic state “the white and black molecules are equally distributed between the left and the right halves of the vessel”. Therefore, the entropy, or “disorder”, is much larger in the latter macroscopic state (see too the discussion on Maxwell’s demon in Sec. 1.5.1). More precisely, Boltzmann defined the thermodynamic entropy as

$$S(E) = k_B \ln \omega(E), \quad (6.160)$$

where  $k_B$  is the Boltzmann constant and  $\omega(E)$  is the measure of the energy surface  $H(q, p) = E$ , where  $H$  is the system Hamiltonian and  $(q, p)$  denotes the phase-space coordinates and momenta of the  $N$  molecules.

We point out that Boltzmann’s definition of entropy (6.160) assumes that when a system is in thermodynamic equilibrium, all microscopic states satisfying the macroscopic conditions of the system are equiprobable. This implies that in equilibrium the density  $\rho(q, p)$  of points in phase space is described by the *microcanonical ensemble*:

$$\rho(q, p) = \frac{1}{\omega(E)} \delta(H(q, p) - E). \quad (6.161)$$

Although Eq. (6.160) is a bridge between the microscopic and the macroscopic descriptions of matter, it only refers to states in thermodynamic equilibrium. In order to obtain a definition of entropy that is also applicable out of equilibrium, it is convenient to consider the *canonical ensemble*, which is appropriate for the description of systems in contact with a heat reservoir at temperature  $T$ . In this

case, taking into account the first and the second principles of thermodynamics, one obtains (see, for instance, Toda *et al.*, 1983)

$$S(T) = k_B(\ln Z + \beta \bar{E}), \quad (6.162)$$

where  $\bar{E}$  is the average energy of the system,  $\beta = \frac{1}{k_B T}$  and

$$Z = \int dqdp e^{-\beta H(q,p)} \quad (6.163)$$

is the partition function. Taking into account that

$$\bar{E} = \int dqdp \rho(q,p) H(q,p), \quad (6.164)$$

with

$$\rho(q,p) = \frac{1}{Z} e^{-\beta H(q,p)}, \quad (6.165)$$

we obtain the following from (6.162):

$$\begin{aligned} S &= k_B \left( \ln Z + \beta \frac{1}{Z} \int dqdp e^{-\beta H(q,p)} H(q,p) \right) \\ &= k_B \left( \ln Z - \frac{1}{Z} \int dqdp e^{-\beta H(q,p)} \ln(\rho Z) \right) \\ &= -k_B \int dqdp \rho(q,p) \ln \rho(q,p). \end{aligned} \quad (6.166)$$

Note that the statistical entropy  $S = -k_B \int dqdp \rho(q,p) \ln \rho(q,p)$  can be directly defined as the average value of  $-\ln \rho(q,p)$ . In particular, in the case of the microcanonical ensemble,  $\rho(q,p)$  is given by (6.161) and therefore the statistical entropy (6.166) reduces to the thermodynamic entropy (6.160).

It can be shown that if on the energy surface we consider a distribution  $\rho'(q,p)$  different from the microcanonical distribution (6.161), then the entropy  $S = -k_B \int dqdp \rho'(q,p) \ln \rho'(q,p)$  is smaller than the thermodynamic entropy (6.160). We therefore conclude that the microcanonical distribution maximizes the statistical entropy.

It is interesting that the expression  $-\int dqdp \rho \ln \rho$  can be viewed as a measure of the “degree of uncertainty” associated with the measure  $d\mu = \rho dqdp$ . Such uncertainty is small when  $\mu$  is peaked and large when  $\rho$  is spread over the energy surface. We therefore obtain a simple statistical interpretation of the entropy of, say, a gas: it has the meaning of the degree of uncertainty in the microscopic state of the gas corresponding to a given macroscopic state. Hence, we can exploit the fact that the microcanonical ensemble maximizes the expression  $-\int dqdp \rho \ln \rho$  to justify its use in statistical physics.

Finally, we wish to point out that the expression  $-\int dqdp \rho \ln \rho$  is the analogue, for continuous variables, of the Shannon entropy  $-\sum_i p_i \ln p_i$ .

### 6.11.3 \* Dynamical Kolmogorov–Sinai entropy

The Kolmogorov–Sinai (KS) entropy refers to the dynamical behaviour of a system: it characterizes its dynamical stability and provides a measure of the rate at which memory of the initial conditions is lost. We shall not be concerned here with rigorous mathematical details and, instead, we give below a simple operative definition for computation of the KS entropy.

Let us consider a partition  $Q$  of the energy surface into  $N$  cells and attach an index  $j$  to each cell ( $j = 1, \dots, N$ ). We then follow the evolution of an orbit at discrete times  $t_0 = 0$ ,  $t_1 = T$ ,  $t_2 = 2T$  and so on. We associate the sequence of symbols (or letters)  $i_0, i_1, i_2, \dots$  with the orbit if the orbit resides in cell  $i_0$  at time  $t_0$ , in  $i_1$  at  $t_1$ , in  $i_2$  at  $t_2$  and so on. Given a sequence (word) of  $m$  symbols  $s_1, \dots, s_m$ , we call  $p_{s_1, \dots, s_m}^{(Q)}$  the probability that such a sequence appears in the orbit. The probability  $p_{s_1, \dots, s_m}^{(Q)}$  can be computed in practice by following the orbit up to very long times and counting the number of recurrences of the sequence  $s_1, \dots, s_m$  in the sequence  $i_0, i_1, i_2, \dots$  associated with the orbit. After repeating the same calculation for all  $m$ -letter words, we obtain the quantity

$$K^{(Q)}(m) = - \sum_{s_1, \dots, s_m=1}^N p_{s_1, \dots, s_m}^{(Q)} \ln p_{s_1, \dots, s_m}^{(Q)}. \quad (6.167)$$

The KS entropy  $h$  is finally defined as

$$h = \sup_Q \lim_{m \rightarrow \infty} \frac{K^{(Q)}(m)}{m}. \quad (6.168)$$

In practice, the entropy of a dynamical system is computed numerically starting from a regular partition of the energy surface into  $D$ -dimensional hypercubes of volume  $\epsilon^D$ . In this manner, the entropy of the  $\epsilon$ -partition can be computed as

$$h(\epsilon) = \lim_{m \rightarrow \infty} \frac{K^{(\epsilon)}(m)}{m}. \quad (6.169)$$

The dynamical entropy is then obtained as

$$h = \lim_{\epsilon \rightarrow 0} h(\epsilon). \quad (6.170)$$

The practical advantage of this procedure is evident: it is not necessary to consider all possible partitions, but simply take hypercubes of small enough volume. At any rate, the numerical calculation of the quantity  $h(\epsilon)$  is very difficult. Indeed, the Shannon–McMillan theorem states that the number of “typical”  $m$ -words increases as  $\exp(h(\epsilon)m)$ . Thus, for chaotic systems ( $h(\epsilon) > 0$ ) the number of typical  $m$ -words grows exponentially with  $m$ , so that it is very hard to compute  $K^{(\epsilon)}(m)$ . Intuitively, such an exponential proliferation is related to the exponential instability of orbits. This intuition is made rigorous by Pesin’s theorem, which states that

$$h = \sum_{\lambda_i > 0} \lambda_i, \quad (6.171)$$

where the sum extends over all positive Lyapunov exponents.<sup>3</sup> This result is also useful for computing the KS entropy in a simple manner.

It is interesting to investigate what relation (if any) exists between the Boltzmann–Gibbs statistical entropy

$$S(t) = -k_B \int dq dp \rho(q, p; t) \ln \rho(q, p; t) \quad (6.172)$$

and the Kolmogorov–Sinai entropy. First of all we observe that, according to Liouville’s theorem, the phase-space volume occupied by the distribution  $\rho(q, p; t)$  is conserved. This implies that the entropy  $S(t)$  does not vary with time at all. However, the shape of the volume becomes increasingly complicated, due to the chaotic dynamics. Thus, after smoothing of the probability distribution, the volume occupied increases. The simplest manner to perform such *coarse graining* is to divide the energy surface into  $N$  cells, each of the same extension. Let  $p_i(t)$  denote then the probability that the state of the system falls inside the cell  $i$  at time  $t$ . The coarse-grained statistical entropy is now defined as

$$S_c(t) = - \sum_i p_i(t) \ln p_i(t). \quad (6.173)$$

For a chaotic system the coarse-grained distribution converges to the microcanonical distribution; that is,  $\lim_{t \rightarrow \infty} p_i(t) = \frac{1}{N}$ . Therefore, the equilibrium value of the coarse-grained entropy (6.173) is  $S_c(\infty) = \lim_{t \rightarrow \infty} S_c(t) = \ln N$ . Let us assume that the initial, far-from-equilibrium distribution  $\rho(q, p; 0)$  is strongly peaked in phase space, for instance it is localized in a single cell. In this case, the coarse-grained entropy  $S_c(t)$  evolves from the initial value  $S_c(0) = 0$  to the equilibrium value  $S_c(\infty) = \ln N$ . There are no rigorous mathematical results connecting the KS entropy to the coarse-grained statistical entropy. Nevertheless, qualitative analytical arguments as well as numerical results (see Latora and Baranger, 1999) show that, for systems characterized by uniform (in phase space) exponential instability, after an initial transient stage  $S_c(t)$  increases linearly with a slope given by the Kolmogorov–Sinai entropy. A simple example illustrating the connection between the growth rate of the coarse-grained statistical entropy  $S_c$  and the KS entropy  $h$  is shown in Fig. 6.8.

## 6.12 A guide to the bibliography

There are several good reviews on the general subject of entanglement in the quantum information theory context; among them, we quote Bruß (2002), Alber *et al.* (2001), Plenio and Virmani (2007), and R. Horodecki *et al.* (2009). These also provide an introduction to the open problems of the quantification of mixed-state entanglement and of multipartite entanglement. The behaviour of entanglement

---

<sup>3</sup>For a dynamical system evolving in an  $n$ -dimensional phase space, there are  $n$  Lyapunov exponents; for their definition see, *e.g.*, Ott (2002). The largest Lyapunov exponent is defined in Sec. 1.4.2.

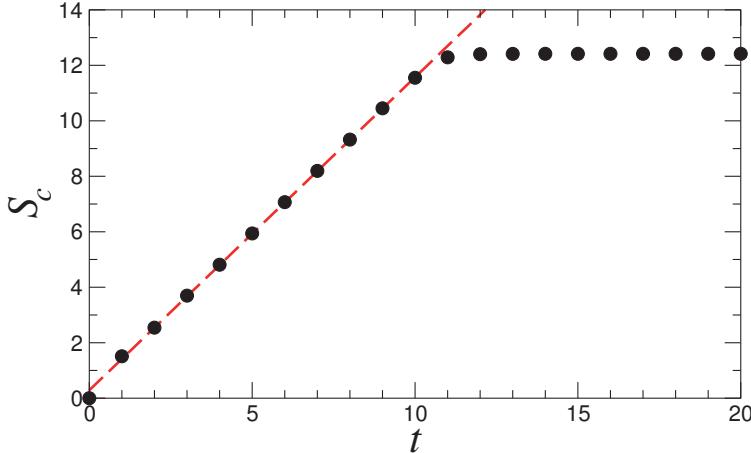


Fig. 6.8 Time evolution of the coarse-grained entropy  $S_c$  (circles) for the classical sawtooth map  $\bar{J} = J + K(\theta - \pi)$ ,  $\bar{\theta} = \theta + \bar{J}$  on the torus  $0 \leq \theta < 2\pi$ ,  $-\pi \leq J < \pi$ , with  $K = \sqrt{2}$  (this map corresponds to the classical limit of the quantum sawtooth map, described in Sec. 4.7.3). The coarse graining is obtained by dividing the torus into  $N = 2.5 \times 10^5$  square cells. The initial density distribution uniformly covers a single cell centred at  $(\theta, J) = (\frac{\pi}{5}, \frac{3\pi}{5})$ . The dashed line has slope given by the Kolmogorov–Sinai entropy  $h \approx 1.13$  (note that the sawtooth map is a conservative chaotic system and there is a single positive Lyapunov exponent  $\lambda$ , so that, according to Pesin’s theorem,  $h = \lambda$ ).

across a quantum phase transition in spin systems and other many-body systems has recently attracted much interest, see e.g. Amico *et al.* (2008). Many of such studies are based on entanglement estimators introduced in Wootters (1998) and Coffman *et al.* (2000). The Peres criterion was found by Peres (1996); see also M. Horodecki *et al.* (1996) and Alber *et al.* (2001).

The formula for the mean bipartite entanglement of a random state was conjectured by Page (1993); for proofs of this formula see Foong and Kanno (1994); Sánchez-Ruiz (1995); Sen (1996). The formula for the average purity of the reduced density matrices of a random state is due to Lubkin (1978).

A textbook containing a mathematically rigorous presentation of the geometry of quantum state spaces and quantum entanglement is Bengtsson and Życzkowski (2017).

The quantum discord is a relatively new concept of the quantum information, so that it is still difficult to provide a univocal characterization of the problem. Three recent reviews on the quantum discord and related measures of quantum correlations can be however found in Modi *et al.* (2012), Adesso *et al.* (2016), and Bera *et al.* (2018).

Thermodynamic and statistical entropies are discussed in statistical mechanics textbooks, such as Huang (1987), Toda *et al.* (1983), and Pathria and Beale (2011). An introduction to the Kolmogorov–Sinai entropy is Kornfeld *et al.* (1982).

# Chapter 7

## Decoherence

In practice, any quantum system is *open*; namely, it is never perfectly isolated from the *environment*. The word decoherence, used in its broader meaning, denotes any quantum-noise process due to the unavoidable coupling of the system to the environment. Decoherence theory has a fundamental interest beyond quantum information science since it provides explanations of the emergence of classicality in a world governed by the laws of quantum mechanics. The core of the problem is the superposition principle, according to which any superposition of quantum states is an acceptable quantum state. This entails consequences that are absurd according to classical intuition, such as the superposition of “live cat” and “dead cat” considered in Schrödinger’s well-known cat paradox. The interaction with the environment can destroy the coherence between the states appearing in a superposition (for instance, the “live-cat” and “dead-cat” states).

In quantum information processing, decoherence is a threat to the actual implementation of any quantum computation or communication protocol. Indeed, decoherence invalidates the quantum superposition principle, which lies at the heart of the potential power of any quantum algorithm. On the other hand, decoherence is also an essential ingredient for quantum information processing, which must end up with a measurement by converting quantum states into classical outcomes. We shall see that decoherence plays a key role in the quantum measurement process.

In this chapter, we shall describe decoherence using various tools, from the quantum-operation formalism to the master-equation and the quantum-trajectory approaches. General results will be illustrated by means of concrete examples, including single-qubit noise models and the damped quantum harmonic oscillator.

### 7.1 The Kraus representation

Let us consider a bipartite system  $1 + 2$ . The system undergoes a unitary evolution and we wish to describe the evolution of subsystem 1 alone. We assume that initially the two subsystems are not entangled (we shall see later in this section that there is no lack of generality in this assumption) and described by the density matrix

$$\rho_{12} = \rho_1 \otimes |0\rangle_2 \langle 0|. \quad (7.1)$$

Namely, subsystem 2 is in a pure state, which we call  $|0\rangle_2$ . There is no loss of generality in the assumption that subsystem 2 is initially in a pure state. As we saw in Sec. 2.8, if this is not the case, we can enlarge the Hilbert space of subsystem 2 in order to purify it. The temporal evolution of the total system is governed by the unitary time-evolution operator  $U$ , which leads to the new density matrix

$$\rho'_{12} = U \rho_{12} U^\dagger = U(\rho_1 \otimes |0\rangle_2 \langle 0|) U^\dagger. \quad (7.2)$$

The quantum circuit implementing this transformation is shown in Fig. 7.1.

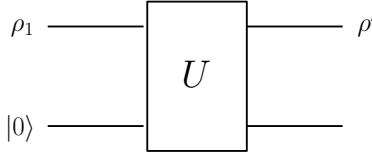


Fig. 7.1 A quantum circuit implementing the transformation (7.2). The state  $\rho'_1$  is obtained after partial tracing over the other subsystem (lower line in the figure) the overall density operator  $U(\rho_1 \otimes |0\rangle \langle 0|)U^\dagger$ .

As explained in Sec. 2.6.1, since we are interested in the new density matrix  $\rho'_1$  describing subsystem 1, we must trace over the second subsystem:

$$\rho'_1 = \text{Tr}_2(\rho'_{12}) = \text{Tr}_2\left[U(\rho_1 \otimes |0\rangle_2 \langle 0|)U^\dagger\right] = \sum_k {}_2\langle k|U|0\rangle_2 \rho_1 {}_2\langle 0|U^\dagger|k\rangle_2, \quad (7.3)$$

where  $\{|k\rangle_2\}$  is a basis set for the Hilbert space  $\mathcal{H}_2$  associated with subsystem 2 and  ${}_2\langle k|U|0\rangle_2$  is an operator acting on the Hilbert space  $\mathcal{H}_1$  associated with subsystem 1. If we define the *Kraus operators*

$$E_k \equiv {}_2\langle k|U|0\rangle_2, \quad (7.4)$$

then we can rewrite Eq. (7.3) as

$$\rho'_1 = \sum_k E_k \rho_1 E_k^\dagger. \quad (7.5)$$

Since  $U$  is unitary, the operators  $E_k$  satisfy the property

$$\sum_k E_k^\dagger E_k = \sum_k {}_2\langle 0|U^\dagger|k\rangle_2 {}_2\langle k|U|0\rangle_2 = {}_2\langle 0|U^\dagger U|0\rangle_2 = I_1, \quad (7.6)$$

where  $I_1$  denotes the identity operator in the Hilbert space  $\mathcal{H}_1$ . Note that we have used the completeness relation  $\sum_k |k\rangle_2 \langle k| = I_2$ . Equation (7.5) defines a linear map  $\mathbb{S}$  from linear operators to linear operators:

$$\mathbb{S} : \rho_1 \rightarrow \rho'_1 = \sum_k E_k \rho_1 E_k^\dagger. \quad (7.7)$$

If the completeness relation (7.6) is satisfied, map  $\mathbb{S}$  is known as a *quantum operation* or a *superoperator* and Eq. (7.5) is known as the *Kraus representation* (or the *operator-sum representation*) of the superoperator  $\mathbb{S}$ . Note that, if  $U(t)$  denotes

the time-evolution operator from time 0 to time  $t$ , then  $E_k$  depends on time and Eq. (7.7) can be written as  $\mathbb{S}(t): \rho_1(0) \rightarrow \rho_1(t) = \sum_k E_k(t) \rho_1(0) E_k^\dagger(t)$ , where  $\rho_1(t)$  is the density matrix describing subsystem 1 at time  $t$ .

A superoperator maps density operators to density operators, since:

- (1)  $\rho'_1$  is Hermitian if  $\rho_1$  is Hermitian:

$$(\rho'_1)^\dagger = \left( \sum_k E_k \rho_1 E_k^\dagger \right)^\dagger = \sum_k (E_k^\dagger)^\dagger \rho_1^\dagger E_k^\dagger = \sum_k E_k \rho_1 E_k^\dagger = \rho'_1; \quad (7.8)$$

- (2)  $\rho'_1$  has unit trace if  $\rho_1$  has unit trace:

$$\text{Tr}(\rho'_1) = \text{Tr}\left(\sum_k E_k \rho_1 E_k^\dagger\right) = \sum_k \text{Tr}(\rho_1 E_k^\dagger E_k) = \text{Tr}\left(\rho_1 \sum_k E_k^\dagger E_k\right) = 1; \quad (7.9)$$

- (3)  $\rho'_1$  is non-negative if  $\rho_1$  is non-negative:

$${}_1\langle \psi | \rho'_1 | \psi \rangle_1 = \sum_k {}_1\langle \psi | E_k \rho_1 E_k^\dagger | \psi \rangle_1 = \sum_k {}_1\langle \varphi_k | \rho_1 | \varphi_k \rangle_1 \geq 0, \quad (7.10)$$

where  $|\psi\rangle_1$  is any vector in  $\mathcal{H}_1$  and  $|\varphi_k\rangle_1 \equiv E_k^\dagger |\psi\rangle_1$ .

### Unitary representation

So far, we have shown that the unitary evolution of a composite system naturally gives rise to an operator-sum representation describing the evolution of a subsystem. We now tackle the converse problem: given a Kraus representation for the evolution of system 1, we shall show that it is possible to introduce an auxiliary system 2 so that the evolution of the total system  $1+2$  is unitary. In this manner, we construct the unitary representation corresponding to a given superoperator. We define an operator  $U$ , acting as follows on states of the form  $|\psi\rangle_1|0\rangle_2$ :

$$U|\psi\rangle_1|0\rangle_2 \equiv \sum_k E_k |\psi\rangle_1 |k\rangle_2, \quad (7.11)$$

where  $\{|k\rangle_2\}$  is an orthonormal basis for subsystem 2, whose dimension is given by the number of Kraus operators appearing in the operator-sum representation. The operator  $U$  preserves the inner product. Indeed, for arbitrary states  $|\psi\rangle_1$  and  $|\phi\rangle_1$  we have

$$\begin{aligned} {}_1\langle \psi | {}_2\langle 0 | U^\dagger U | \phi \rangle_1 | 0 \rangle_2 &= \left( \sum_k {}_1\langle \psi | E_k^\dagger {}_2\langle k | \right) \left( \sum_l E_l | \phi \rangle_1 | l \rangle_2 \right) \\ &= \sum_k {}_1\langle \psi | E_k^\dagger E_k | \phi \rangle_1 = {}_1\langle \psi | \phi \rangle_1, \end{aligned} \quad (7.12)$$

where we have used the orthonormality relation  ${}_2\langle k | l \rangle_2 = \delta_{kl}$  and the completeness relation (7.6). As the operator  $U$  preserves the inner product when acting on the subspace whose states are of the form  $|\psi\rangle_1|0\rangle_2$ , then it can be extended to a unitary operator acting on the entire Hilbert space  $\mathcal{H}_1 \otimes \mathcal{H}_2$  of the joint system.<sup>1</sup>

<sup>1</sup>It is easy to verify that the unitary transformation (7.11) really induces an operator-sum representation on subsystem 1. Indeed, the evolution of a pure state  $\rho_1 = |\psi\rangle_1 \langle \psi|$  is as follows:

$$\rho_1 \rightarrow \rho'_1 = \text{Tr}_2(U|\psi\rangle_1|0\rangle_2 {}_1\langle \psi | {}_2\langle 0 | U^\dagger) = \sum_k E_k |\psi\rangle_1 {}_1\langle \psi | E_k^\dagger = \sum_k E_k \rho_1 E_k^\dagger.$$

### Composing superoperators

Two superoperators  $\mathbb{S}_A$  and  $\mathbb{S}_B$  can be composed to give a new superoperator  $\mathbb{S} = \mathbb{S}_B \mathbb{S}_A$ , defined by  $\mathbb{S}(\rho_1) = \mathbb{S}_B(\mathbb{S}_A(\rho_1))$ . If  $\mathbb{S}_A$  describes the evolution of the density matrix for system 1 from time  $t_0$  to time  $t_1$  and  $\mathbb{S}_B$  from  $t_1$  to  $t_2$ , then  $\mathbb{S} = \mathbb{S}_B \mathbb{S}_A$  describes the evolution from  $t_0$  to  $t_2$ . It can be shown that a superoperator is invertible if and only if it is unitary. Thus, superoperators are a *semigroup* instead of a group. Physically, this means that an arrow of time has been introduced for subsystem 1. We can describe the evolution from  $t_0$  to  $t_1 > t_0$  by means of a superoperator but not from  $t_1$  to  $t_0$ . There is a loss of information from system 1 to system 2 (known as the *environment*) and we cannot run the evolution of system 1 backward if we know its state but ignore the state of the environment. This phenomenon is known as *decoherence* and we shall show that superoperators provide a very general theoretical framework for its description.

There is a freedom in the operator-sum representation; that is, different representations can give rise to the same superoperator. The following theorem holds: two superoperators  $\mathbb{S}(\rho_1) = \sum_k E_k \rho_1 E_k^\dagger$  and  $\mathbb{S}'(\rho_1) = \sum_k F_k \rho_1 F_k^\dagger$  coincide if and only if there exists a unitary matrix  $W$  such that  $F_i = \sum_j W_{ij} E_j$ . The proof of this theorem can be found in Schumacher (1996).

We note that, in order to build a unitary representation of a superoperator, the dimension of the Hilbert space  $\mathcal{H}_2$  must be at least as large as the number of operators  $E_k$  appearing in the Kraus representation. If the Hilbert space  $\mathcal{H}_1$  has dimension  $N$ , it is possible to prove that all superoperators  $\mathbb{S}(\rho_1)$  can be generated by an operator-sum representation containing at most  $N^2$  operators  $E_k$  (see, e.g., Preskill, 1998a). Therefore, it will be sufficient to consider a Hilbert space  $\mathcal{H}_2$  of dimension  $N^2$ .

We may also ask how many real parameters are required to parametrize a generic superoperator  $\mathbb{S} : \rho_1 \rightarrow \rho'_1$  on a Hilbert space of dimension  $N$ . A superoperator maps a density operator into another density operator; that is, an  $N \times N$  Hermitian matrix to another  $N \times N$  Hermitian matrix. A basis for the space  $\mathcal{M}_{\mathcal{H}}$  of the  $N \times N$  Hermitian matrices has  $N^2$  elements. Therefore, the most general linear transformation acting on the space  $\mathcal{M}_{\mathcal{H}}$  has  $(N^2)^2 = N^4$  free real parameters. We should then take into account the completeness relation  $\sum_k E_k^\dagger E_k = I_1$ , which gives  $N^2$  constraints for these parameters (note that we have  $N^2$  constraints and not  $2N^2$  since the terms  $E_k E_k^\dagger$  are Hermitian; in other words, the Hermitian conjugate of the completeness relation is again the same completeness relation). Hence, a generic superoperator is parametrized by  $N^4 - N^2$  real parameters. For instance, in the case of a single qubit ( $N = 2$ ) we need 12 real parameters, while in the two-qubit case ( $N = 4$ ) we need 240 real parameters.

We may now state the following fundamental theorem (for a proof see Schumacher, 1996):

---

And since a generic density matrix can be expressed as an ensemble of pure states,  $\rho_1 = \sum_i p_i |\psi_i\rangle_1 \langle \psi_i|$ , we recover the Kraus representation (7.5) for arbitrary  $\rho_1$ .

**Theorem 7.1** The Kraus representation theorem (Kraus, 1983): *A map  $\mathbb{S} : \rho_1 \rightarrow \rho'_1$  satisfying the following requirements: it*

(1) *is linear; that is,*

$$\mathbb{S}(p_1\rho_1 + p_2\rho_2) = p_1\mathbb{S}(\rho_1) + p_2\mathbb{S}(\rho_2), \quad (7.13)$$

(2) *preserves hermiticity,*

(3) *preserves trace,*

(4) *is completely positive,*

*has an operator-sum (Kraus) representation (7.5) and a unitary representation (7.11) on a larger Hilbert space, this latter generally known as Stinespring representation for  $\mathbb{S}$ .*

We say that  $\mathbb{S}$  is positive if, for a non-negative  $\rho_1$ ,  $\rho'_1 = \mathbb{S}(\rho_1)$  is also non-negative. The complete positivity of  $\mathbb{S}$  is a stronger requirement. It means that, for any extension of the Hilbert space  $\mathcal{H}_1$  to  $\mathcal{H}_1 \otimes \mathcal{H}_E$ , the superoperator  $\mathbb{S} \otimes \mathbb{I}_E$  is positive. That is, if we add any system  $E$  that has a trivial dynamics (the identity  $\mathbb{I}_E$  means that no state of  $E$  is changed), independently of the dynamics of system 1, the resulting superoperator  $\mathbb{S} \otimes \mathbb{I}_E$  must be positive. This requirement is physically motivated since, in general, it cannot be excluded that the two systems are initially entangled. If this is the case and we call  $\rho_{1E}$  the density matrix corresponding to the initially entangled state, then  $\rho'_{1E} \equiv (\mathbb{S} \otimes \mathbb{I}_E)(\rho_{1E})$  must be a valid density matrix. This implies the positivity of  $\mathbb{S} \otimes \mathbb{I}_E$  for any  $E$ , namely the complete positivity of  $\mathbb{S}$ .

**Exercise 7.1** Consider the state

$$|\psi\rangle_{1E} = \frac{1}{\sqrt{2}}(|0\rangle_1|1\rangle_E + |1\rangle_1|0\rangle_E). \quad (7.14)$$

Let  $\rho_1$  denote the density operator for the first qubit and show that the transposition operator

$$\mathbb{T}(\rho_1) \equiv \rho_1^T \quad (7.15)$$

is positive but not completely positive. For this purpose, it will be sufficient to show that  $\mathbb{T} \otimes \mathbb{I}_E$  is not positive.

The Kraus representation theorem tells us that, if the evolution  $\rho'_1 = \mathbb{S}(\rho_1)$  of the density matrix  $\rho_1$  preserves hermiticity and trace, is linear and completely positive, then this evolution can be realized by the unitary transformation (7.11), acting on a larger Hilbert space  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . Note that in (7.11) subsystems 1 and 2 are not initially entangled. Thus, if we are only interested in the evolution of the density matrix  $\rho_1$ , there is no lack of generality in assuming that subsystem 1 is not initially entangled with subsystem 2.

### Examples

As a simple example illustrating the Kraus representation, we now consider two single-qubit subsystems 1 and 2, with initial density operator given by  $\rho_{12} = |0\rangle_2\langle 0| \otimes \rho_1$ , whose matrix representation is

$$\rho_{12} = \begin{bmatrix} \rho_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (7.16)$$

where  $\rho_1$  and  $\mathbf{0}$  are  $2 \times 2$  submatrices and all elements of  $\mathbf{0}$  are zero. Note that we have taken the qubit in the state  $|0\rangle_2$  as the most significant qubit to obtain a simple block representation for the matrix (7.16). The evolution of the global system  $1+2$  is governed by a unitary  $4 \times 4$  matrix  $U$  and we obtain the new reduced density matrix  $\rho'_1$  after tracing over the degrees of freedom of subsystem 2 (see the quantum circuit implementing this transformation in Fig. 7.1). We have

$$\rho'_1 = \text{Tr}_2(U \rho_{12} U^\dagger) = \text{Tr}_2 \left( \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} A^\dagger & C^\dagger \\ B^\dagger & D^\dagger \end{bmatrix} \right) = A \rho_1 A^\dagger + C \rho_1 C^\dagger, \quad (7.17)$$

where the unitary matrix  $U$  has been expressed in terms of the  $2 \times 2$  submatrices  $A$ ,  $B$ ,  $C$  and  $D$ . If we require that the transformation maps density matrices to density matrices, then  $\text{Tr}(\rho'_1) = \text{Tr}(\rho_1(A^\dagger A + C^\dagger C)) = \text{Tr}(\rho_1)$  for any  $\rho_1$ . It follows that we must have

$$A^\dagger A + C^\dagger C = I_1. \quad (7.18)$$

Thus, we have explicitly constructed a Kraus representation for a single qubit, with the Kraus operators  $A$  and  $C$  satisfying property (7.6).

We point out that, in the most general case for a single qubit, the number of Kraus operators appearing in the operator-sum representation is  $N^2 = 4$  ( $N = 2$  being the dimension of the Hilbert space for a single qubit), corresponding to two qubits for subsystem 2. For instance, if we consider the unitary evolution

$$U|\psi\rangle_1|0\rangle_2 = \sqrt{1-p}I_1|\psi\rangle_1|0\rangle_2 + \sqrt{\frac{1}{3}p}\left[\sigma_1^x|\psi\rangle_1|1\rangle_2 + \sigma_1^y|\psi\rangle_1|2\rangle_2 + \sigma_1^z|\psi\rangle_1|3\rangle_2\right], \quad (7.19)$$

with  $0 \leq p \leq 1$ , then the four Kraus operators are given by  $E_k = {}_2\langle k|U|0\rangle_2$ . We readily obtain

$$E_0 = \sqrt{1-p}I_1, \quad E_1 = \sqrt{\frac{1}{3}p}\sigma_1^x, \quad E_2 = \sqrt{\frac{1}{3}p}\sigma_1^y, \quad E_3 = \sqrt{\frac{1}{3}p}\sigma_1^z, \quad (7.20)$$

and it is easy to check that the operators  $E_k$  satisfy the normalization condition  $\sum_k E_k^\dagger E_k = I_1$ . The evolution of the reduced density matrix  $\rho_1$ , corresponding to the unitary evolution (7.19), is given by

$$\rho_1 \rightarrow \rho'_1 = \sum_{k=0}^3 E_k \rho_1 E_k^\dagger = (1-p)I_1 + \frac{1}{3}p\left[\sigma_1^x \rho_1 \sigma_1^x + \sigma_1^y \rho_1 \sigma_1^y + \sigma_1^z \rho_1 \sigma_1^z\right]. \quad (7.21)$$

This example corresponds to the so-called depolarizing channel and will be discussed, together with several other examples of Kraus representations for a single qubit, in the next section.

## 7.2 Decoherence models for a single qubit

In this section, we shall study quantum-noise (decoherence) processes that can act on a single qubit. A general formulation of the problem, in terms of Kraus operators, will be given in Sec. 7.2.1. Before doing so, it is instructive to consider a very simple decoherence model, drawn in Fig. 7.2. Here the environment consists of a single qubit and the system–environment interaction is represented by a CNOT gate. Let us assume that initially the system is in a pure state,  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , corresponding to the density matrix  $\rho = |\psi\rangle\langle\psi|$ , whose matrix representation in the  $\{|0\rangle, |1\rangle\}$  basis is given by

$$\rho = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}. \quad (7.22)$$

The diagonal terms of  $\rho$  are known as *populations* (see Sec. 2.6), and give the probabilities to obtain, from a polarization measurement along the  $z$ -axis, outcomes 0 or 1, respectively. The off-diagonal terms, known as *coherences*, appear when the state  $|\psi\rangle$  is a superposition of the states  $|0\rangle$  and  $|1\rangle$ . They are completely destroyed by the decoherence process drawn in Fig. 7.2. Indeed, this quantum circuit changes the initial global system–environment state,

$$|\Psi\rangle = |\psi\rangle \otimes |0\rangle = (\alpha|0\rangle + \beta|1\rangle)|0\rangle, \quad (7.23)$$

into the final state

$$|\Psi'\rangle = \alpha|00\rangle + \beta|11\rangle. \quad (7.24)$$

Note that the CNOT interaction has entangled the qubit with the environment, as the state  $|\Psi'\rangle$  is non-separable. The final density matrix  $\rho'$  of the system is obtained after tracing over the environment:

$$\rho' = \text{Tr}_{\text{env}} [|\Psi'\rangle\langle\Psi'|] = \begin{bmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{bmatrix}. \quad (7.25)$$

This decoherence process has a particularly appealing interpretation: it is evident from Eq. (7.24) that the environment has learnt, through the CNOT interaction, what the state of the system is. Indeed, if the state of the system is  $|0\rangle$ , the state of the environment remains  $|0\rangle$ ; on the other hand, if the state of the system is  $|1\rangle$ , the state of the environment is flipped and becomes  $|1\rangle$ . Therefore, the CNOT gate is basically a premeasurement (see Sec. 7.7) performed by the environment on the system. The information on the relative phases of the coefficients  $\alpha$  and  $\beta$  appearing in the initial state  $|\psi\rangle$  is now hidden in the system–environment quantum correlations. Since we do not keep records of the state of the environment, this information is lost for us. In short, information leaks from the system into the external world.

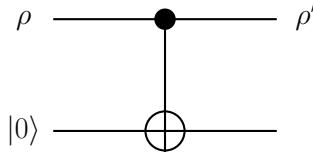


Fig. 7.2 Quantum circuit modelling complete decoherence.

### 7.2.1 The quantum black box

Let us consider a two-level system (qubit) interacting with a generic physical system. This system is known as a quantum *black box*, and its action on the qubit is described in terms of a quantum operation  $\mathbb{S}$ :

$$\rho \rightarrow \rho' = \mathbb{S}(\rho) = \sum_k E_k \rho E_k^\dagger, \quad \text{with} \quad \sum_k E_k^\dagger E_k = I, \quad (7.26)$$

where the Kraus operators are denoted by  $E_k$ . It is convenient to write the states  $\rho$  and  $\rho'$  in the Bloch-sphere representation (3.11):

$$\rho = \frac{1}{2}(I + \mathbf{r} \cdot \boldsymbol{\sigma}) \quad \text{and} \quad \rho' = \frac{1}{2}(I + \mathbf{r}' \cdot \boldsymbol{\sigma}), \quad (7.27)$$

where the Bloch vectors  $\mathbf{r} = (x, y, z)$  and  $\mathbf{r}' = (x', y', z')$  are such that  $|\mathbf{r}|, |\mathbf{r}'| \in [0, 1]$ . The transformation

$$\mathbf{r} \rightarrow \mathbf{r}' = M\mathbf{r} + \mathbf{c} \quad (7.28)$$

is known as an *affine map*. To find the matrix  $M$  and the vector  $\mathbf{c}$  as functions of the Kraus operators  $E_k$ , it is convenient to expand the Kraus operators over the basis  $\{I, \sigma_1 \equiv \sigma_x, \sigma_2 \equiv \sigma_y, \sigma_3 \equiv \sigma_z\}$ :

$$E_k = \gamma_k I + \sum_{l=1}^3 a_{kl} \sigma_l. \quad (7.29)$$

After a lengthy but straightforward calculation (see exercise 7.2), we obtain

$$M_{jk} = \sum_{l=1}^3 \left[ a_{lj} a_{lk}^* + a_{lj}^* a_{lk} + \left( |\gamma_l|^2 - \sum_{p=1}^3 |a_{lp}|^2 \right) \delta_{jk} + i \sum_{p=1}^3 \epsilon_{jkl} \left( \gamma_l a_{lp}^* - \gamma_l^* a_{lp} \right) \right], \quad (7.30)$$

$$c_j = 2i \sum_{k,l,m=1}^3 \epsilon_{jlm} a_{kl} a_{km}^*, \quad (7.31)$$

where  $\epsilon_{jlm}$  is the Levi-Civita antisymmetric tensor.

**Exercise 7.2** Check Eqs. (7.30) and (7.31).

In order to clarify the meaning of the affine map (7.28), we may take advantage of the polar decomposition  $M = OS$ , where  $S$  is a symmetric, non-negative matrix and  $O$  an orthogonal matrix (see App. A.1.6.3). Hence, in the affine map  $S$  deforms the Bloch sphere into an ellipsoid, while  $O$  rotates it and  $\mathbf{c}$  displaces its centre.

We need to determine 12 parameters to describe the action of a generic quantum black box on a two-level system: 6 parameters to determine the symmetric  $3 \times 3$  matrix  $S$ , 3 for the orthogonal matrix  $O$  and 3 for the displacement  $\mathbf{c}$ . Note that the values taken by these parameters must be such that  $\mathbf{r}'$  is still a Bloch vector; that is,  $\rho'$  is still a density matrix.

The number of independent parameters in the single-qubit case is in agreement with the general result of Sec. 7.1: we need  $N^4 - N^2$  independent real parameters to characterize a quantum operation acting on an  $N$ -level quantum system. In particular, we have 12 parameters in the single-qubit case ( $N = 2$ ).

### 7.2.2 Measuring a quantum operation acting on a qubit

We wish to measure the 12 parameters that determine the quantum operation  $\mathbb{S}$  mapping the single-qubit density matrix  $\rho$  to  $\rho' = \mathbb{S}(\rho)$ . For this purpose, we consider the experiment drawn schematically in Fig. 7.3. The source  $S$  emits a large number of qubits, whose states are described by the density matrix  $\rho$ . The qubits enter the quantum black box and come out in states described by a different density matrix,  $\rho'$ . A detector  $D$  measures the density matrix  $\rho'$ , following the procedure described in Sec. 3.2.2.

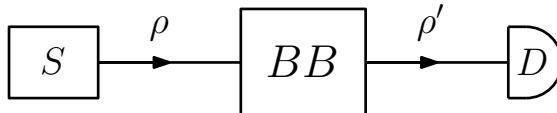


Fig. 7.3 A schematic diagram of the measurement procedure used to determine the effect of a quantum black box (BB) on a qubit.

The affine map (7.28) reads

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix}, \quad (7.32)$$

where we assume the parameters  $M_{ij}$  and  $c_i$  to be time-independent; namely, the quantum black box always acts in the same manner on every two-level system. We wish to determine these parameters. To this end it is sufficient to consider pure initial states  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ . The corresponding density matrix  $\rho = |\psi\rangle\langle\psi|$  is given by Eq. (7.22). Note that in the Bloch-sphere representation

$$\rho = \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix}, \quad (7.33)$$

and so the coordinates  $(x, y, z)$  are related to  $\alpha$  and  $\beta$  as follows:

$$\alpha\beta^* = \frac{1}{2}(x - iy), \quad |\alpha|^2 = \frac{1}{2}(1 + z), \quad |\beta|^2 = \frac{1}{2}(1 - z). \quad (7.34)$$

To determine the 12 parameters  $M_{ij}$  and  $c_i$ , we need to prepare different, appropriate, initial states, for instance:

- (i)  $|\psi_1\rangle = |0\rangle$  ( $\alpha = 1, \beta = 0, x = y = 0, z = 1$ ). As described in Sec. 3.2.2, if we have at our disposal a large number of identically prepared qubits in the state  $|\psi_1\rangle$  and entering the quantum black box, we can measure the final density matrix  $\rho'_1$  and determine its Bloch coordinates  $x'_1, y'_1$  and  $z'_1$ , up to statistical errors. From Eq. (7.32), we obtain

$$x'_1 = M_{13} + c_x, \quad y'_1 = M_{23} + c_y, \quad z'_1 = M_{33} + c_z. \quad (7.35)$$

- (ii)  $|\psi_2\rangle = |1\rangle$  ( $\alpha = 0, \beta = 1, x = y = 0, z = -1$ ). In this case,

$$x'_2 = -M_{13} + c_x, \quad y'_2 = -M_{23} + c_y, \quad z'_2 = -M_{33} + c_z. \quad (7.36)$$

We can now determine the 6 parameters  $M_{i3}$  and  $c_i$  from Eqs. (7.35) and (7.36).

- (iii)  $|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  ( $\alpha = \frac{1}{\sqrt{2}}, \beta = \frac{1}{\sqrt{2}}, x = 1, y = z = 0$ ).

$$x'_3 = M_{11} + c_x, \quad y'_3 = M_{21} + c_y, \quad z'_3 = M_{31} + c_z. \quad (7.37)$$

- (iv)  $|\psi_4\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$  ( $\alpha = \frac{1}{\sqrt{2}}, \beta = \frac{i}{\sqrt{2}}, y = 1, x = z = 0$ ).

$$x'_4 = M_{12} + c_x, \quad y'_4 = M_{22} + c_y, \quad z'_4 = M_{32} + c_z. \quad (7.38)$$

The remaining 6 unknown parameters  $M_{i1}$  and  $M_{i2}$  are computed using Eqs. (7.37) and (7.38).

In principle, the method described in this section can be extended to quantum black boxes acting on many-qubit systems; already though with two qubits (a Hilbert space of dimension  $N = 4$ ), there are  $N^4 - N^2 = 240$  real parameters to be determined.

### 7.2.3 Quantum circuits simulating noise channels

A useful representation of quantum operations is obtained using quantum circuits, in which the environment is represented by ancillary qubits.

Let us consider the circuit drawn in Fig. 7.4. We have a single-qubit system plus an environment with two ancillary qubits. We assume that initially the system is described by the density matrix  $\rho$ , while the ancillary qubits are in the pure state

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle, \quad (7.39)$$

with the normalization condition  $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$ . The initial total density matrix (system plus environment) is given by

$$\rho_{\text{in}}^{(\text{tot})} = |\psi\rangle\langle\psi| \otimes \rho = \begin{bmatrix} |\alpha|^2\rho & .. & .. & .. \\ .. & |\beta|^2\rho & .. & .. \\ .. & .. & |\gamma|^2\rho & .. \\ .. & .. & .. & |\delta|^2\rho \end{bmatrix}, \quad (7.40)$$

where, to simplify the expression, we have denoted by .. the matrix blocks whose expressions are not needed in subsequent calculations.

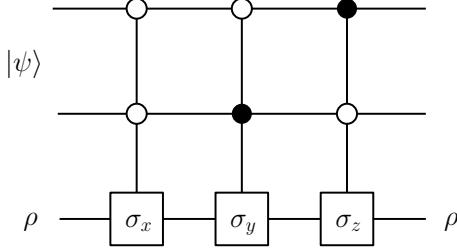


Fig. 7.4 A quantum circuit implementing the deformation of the Bloch sphere into an ellipsoid centred at the origin of the Bloch sphere with axes directed along  $x$ ,  $y$  and  $z$ .

The quantum circuit in Fig. 7.4 implements the unitary transformation

$$U = \begin{bmatrix} \sigma_x & 0 & 0 & 0 \\ 0 & \sigma_y & 0 & 0 \\ 0 & 0 & \sigma_z & 0 \\ 0 & 0 & 0 & I \end{bmatrix}, \quad (7.41)$$

where  $I$  is the  $2 \times 2$  identity matrix. This means that, as shown in Fig. 7.4,  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$ , or  $I$  are applied to the bottom qubit if the two upper qubits are in the states  $|00\rangle$ ,  $|01\rangle$ ,  $|10\rangle$ , or  $|11\rangle$ . The final three-qubit state (system plus environment) is described by the density matrix

$$\rho_{\text{fin}}^{(\text{tot})} = U \rho_{\text{in}}^{(\text{tot})} U^\dagger. \quad (7.42)$$

Note that the system is now, in general, entangled with the environment. After tracing over environmental qubits, we obtain the final state of the system:

$$\rho' = \text{Tr}_{\text{env}} [\rho_{\text{fin}}^{(\text{tot})}] = \mathbb{S}(\rho) = |\alpha|^2 \sigma_x \rho \sigma_x^\dagger + |\beta|^2 \sigma_y \rho \sigma_y^\dagger + |\gamma|^2 \sigma_z \rho \sigma_z^\dagger + |\delta|^2 \rho. \quad (7.43)$$

If we introduce the Kraus operators

$$\begin{aligned} E_1 &= |\alpha| \sigma_x, & E_2 &= |\beta| \sigma_y, \\ E_3 &= |\gamma| \sigma_z, & E_0 &= |\delta| I, \end{aligned} \quad (7.44)$$

we have  $\rho' = \sum_{i=0}^3 E_i \rho E_i^\dagger$ . The transformation induced by the Kraus operators (7.44) can be clearly visualized in the Bloch-sphere representation. Let us consider the Bloch vectors  $\mathbf{r}$  and  $\mathbf{r}'$  associated with the density matrices  $\rho$  and  $\rho'$ , respectively (see Eq. (7.27)). It is easy to check by direct computation that

$$\sigma_x \rho \sigma_x^\dagger = \sigma_x \rho \sigma_x = \frac{1}{2} \begin{bmatrix} 1-z & x+iy \\ x-iy & 1+z \end{bmatrix}, \quad (7.45a)$$

$$\sigma_y \rho \sigma_y^\dagger = \sigma_y \rho \sigma_y = \frac{1}{2} \begin{bmatrix} 1-z & -(x+iy) \\ -(x-iy) & 1+z \end{bmatrix}, \quad (7.45b)$$

$$\sigma_z \rho \sigma_z^\dagger = \sigma_z \rho \sigma_z = \frac{1}{2} \begin{bmatrix} 1+z & -(x-iy) \\ -(x+iy) & 1-z \end{bmatrix}. \quad (7.45c)$$

Taking into account that  $|\delta|^2 = 1 - |\alpha|^2 - |\beta|^2 - |\gamma|^2$ , we obtain

$$\begin{aligned} x' &= [1 - 2(|\beta|^2 + |\gamma|^2)]x, \\ y' &= [1 - 2(|\gamma|^2 + |\alpha|^2)]y, \\ z' &= [1 - 2(|\alpha|^2 + |\beta|^2)]z. \end{aligned} \quad (7.46)$$

These expressions tell us that the Bloch sphere is deformed into an ellipsoid, centred at the origin of the Bloch sphere and whose axes are directed along  $x$ ,  $y$  and  $z$ . As we shall see in the following, depending on the choice of the parameters  $|\alpha|$ ,  $|\beta|$  and  $|\gamma|$ , many interesting noise channels can be obtained.

We note that, as can be clearly seen from Eq. (7.46), the state  $\rho'$  only depends on the amplitudes of the coefficients  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  in the state (7.39) and not on their phases. This implies that, for any density matrix describing the initial state of the two ancillary qubits and having diagonal terms equal to  $|\alpha|^2$ ,  $|\beta|^2$ ,  $|\gamma|^2$  and  $|\delta|^2$ , the circuit of Fig. 7.4 would implement the quantum operation (7.43).

We shall show in the following subsections that commonly investigated noise channels such as the bit-flip or the phase-flip channel can be obtained as special cases of the circuit in Fig. 7.4.

#### 7.2.4 The bit-flip channel

The bit-flip channel is obtained by taking  $\beta = \gamma = 0$  in the state (7.39). In this case Eq. (7.43) reduces to

$$\rho' = \mathbb{S}(\rho) = |\alpha|^2 \sigma_x \rho \sigma_x^\dagger + (1 - |\alpha|^2) \rho, \quad (7.47)$$

and the transformation  $\rho \rightarrow \rho' = \sum_k E_k \rho E_k^\dagger$  can be implemented by means of the Kraus operators

$$E_0 = \sqrt{1 - |\alpha|^2} I, \quad E_1 = |\alpha| \sigma_x. \quad (7.48)$$

We point out that this noise channel flips the state of a qubit (from  $|0\rangle$  to  $|1\rangle$  and *vice versa*) with probability  $|\alpha|^2$ . Indeed, the state  $\rho = |0\rangle\langle 0|$  is mapped into  $\rho' = |\alpha|^2|1\rangle\langle 1| + (1 - |\alpha|^2)|0\rangle\langle 0|$ , while  $\rho = |1\rangle\langle 1|$  becomes  $\rho' = |\alpha|^2|0\rangle\langle 0| + (1 - |\alpha|^2)|1\rangle\langle 1|$ .

The deformation of the Bloch-sphere coordinates, given by Eq. (7.46), simplifies as follows:

$$x' = x, \quad y' = (1 - 2|\alpha|^2)y, \quad z' = (1 - 2|\alpha|^2)z. \quad (7.49)$$

Hence, the Bloch sphere is mapped into an ellipsoid with  $x$  as the symmetry axis. Note that the  $x$  component is not modified by the bit-flip channel since the eigenstates  $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$  of the Kraus operator  $E_1$  are directed along the  $x$ -axis of the Bloch sphere and  $\mathbb{S}(|\pm\rangle\langle \pm|) = |\pm\rangle\langle \pm|$ .

A quantum circuit implementing the bit-flip channel is shown in Fig. 7.5. Note that a single auxiliary qubit is sufficient to describe such a quantum operation (in this quantum circuit we take  $|\psi\rangle = \alpha|0\rangle + \delta|1\rangle$  with  $|\delta| = \sqrt{1 - |\alpha|^2}$ ).

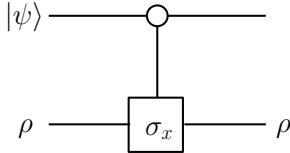


Fig. 7.5 A quantum circuit implementing the bit-flip channel.

### 7.2.5 The phase-flip channel

The phase-flip channel is obtained by taking  $\alpha = \beta = 0$  in the state (7.39). In this case Eq. (7.43) reduces to

$$\rho' = \mathbb{S}(\rho) = |\gamma|^2 \sigma_z \rho \sigma_z^\dagger + (1 - |\gamma|^2) \rho, \quad (7.50)$$

and the superoperator  $\rho \rightarrow \rho' = \mathbb{S}(\rho)$  can be realized by means of the Kraus operators

$$E_0 = \sqrt{1 - |\gamma|^2} I, \quad E_1 = |\gamma| \sigma_z. \quad (7.51)$$

This noise channel introduces a phase error with probability  $|\gamma|^2$ . Indeed, if we apply the superoperator  $\mathbb{S}$  to a pure state  $|\varphi_+\rangle = \mu|0\rangle + \nu|1\rangle$ , described by the density matrix  $\rho = |\varphi_+\rangle\langle\varphi_+|$ , we obtain

$$\rho' = \mathbb{S}(\rho) = |\gamma|^2 |\varphi_-\rangle\langle\varphi_-| + (1 - |\gamma|^2) |\varphi_+\rangle\langle\varphi_+|, \quad (7.52)$$

where  $|\varphi_-\rangle = \mu|0\rangle - \nu|1\rangle$ . Note that  $|\varphi_-\rangle$  differ from  $|\varphi_+\rangle$  only in the relative sign of the coefficients in front of the basis states  $|0\rangle$  and  $|1\rangle$ . Therefore, this noise channel has no classical analogue.

The deformation of the Bloch-sphere coordinates, given by Eq. (7.46), simplifies as follows:

$$x' = (1 - 2|\gamma|^2) x, \quad y' = (1 - 2|\gamma|^2) y, \quad z' = z. \quad (7.53)$$

Hence, the Bloch sphere is mapped into an ellipsoid with  $z$  as symmetry axis. Note that the component  $z$  is not modified by the phase-flip channel since the eigenstates of the Kraus operator  $E_1$  ( $|0\rangle$  and  $|1\rangle$ ) are directed along the  $z$ -axis of the Bloch sphere, and we have  $\mathbb{S}(|0\rangle\langle 0|) = |0\rangle\langle 0|$ ,  $\mathbb{S}(|1\rangle\langle 1|) = |1\rangle\langle 1|$ .

A quantum circuit implementing the phase-flip channel is shown in Fig. 7.6 (in this quantum circuit, the initial state of the auxiliary qubit is  $|\psi\rangle = \gamma|0\rangle + \delta|1\rangle$ , with  $|\delta| = \sqrt{1 - |\gamma|^2}$ ).

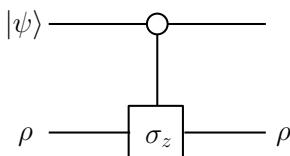


Fig. 7.6 A quantum circuit implementing the phase-flip channel.

As we saw in Sec. 7.1, there is a freedom in the Kraus representation; namely, we can choose different sets of Kraus operators giving rise to the same quantum operation. An example is shown in Fig. 7.7: this circuit leads (see exercise 7.3) to the Kraus operators

$$F_0 = \begin{bmatrix} 1 & 0 \\ 0 & \cos \theta \end{bmatrix}, \quad F_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sin \theta \end{bmatrix} \quad (7.54)$$

and we have

$$\rho' = \mathbb{S}(\rho) = \sum_{i=0}^1 E_i \rho E_i^\dagger = \sum_{i=0}^1 F_i \rho F_i^\dagger, \quad (7.55)$$

provided  $\cos \theta = 1 - 2|\gamma|^2$ . It is easy to check that  $\{E_0, E_1\}$  and  $\{F_0, F_1\}$  are connected by a unitary transformation, as it must be the case for different sets of Kraus operators representing the same superoperator. Indeed, we have  $F_i = \sum_{j=0}^1 W_{ij} E_j$ , where the unitary matrix  $W$  reads

$$W = \begin{bmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{bmatrix}. \quad (7.56)$$

**Exercise 7.3** Show that the quantum circuit of Fig. 7.7 induces a quantum operation  $\rho \rightarrow \rho' = F_0 \rho F_0^\dagger + F_1 \rho F_1^\dagger$ , where  $F_0$  and  $F_1$  are the Kraus operators written in Eq. (7.54).

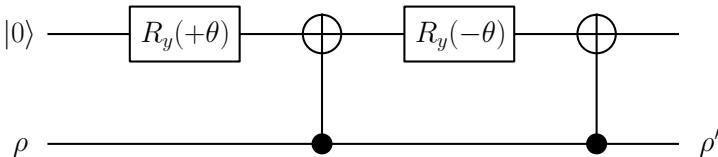


Fig. 7.7 A second circuit implementing the phase-flip channel. The gates labelled  $\pm\theta/2$  stand for the rotation matrices  $R_y(\mp\theta)$  (see Eq. (5.21)).

### 7.2.6 The bit-phase-flip channel

The bit-phase-flip channel is defined by setting  $\alpha = \gamma = 0$  in the state (7.39), so that Eq. (7.43) reduces to

$$\rho' = \mathbb{S}(\rho) = |\beta|^2 \sigma_y \rho \sigma_y^\dagger + (1 - |\beta|^2) \rho, \quad (7.57)$$

and the superoperator  $\rho \rightarrow \rho' = \mathbb{S}(\rho)$  can be expressed in terms of the Kraus operators

$$E_0 = \sqrt{1 - |\beta|^2} I, \quad E_1 = |\beta| \sigma_y. \quad (7.58)$$

This channel induces both bit flip and phase flip. Indeed, it maps the state  $\mu|0\rangle + \nu|1\rangle$  into  $\mu|1\rangle - \nu|0\rangle$  with probability  $|\beta|^2$ .

The transformation of the Bloch-sphere coordinates is given by:

$$x' = (1 - 2|\beta|^2)x, \quad y' = y, \quad z' = (1 - 2|\beta|^2)z. \quad (7.59)$$

Hence, the Bloch sphere is mapped into an ellipsoid symmetric about the  $y$ -axis.

A quantum circuit implementing the bit-phase-flip channel is shown in Fig. 7.8 (in this quantum circuit, the initial state of the auxiliary qubit is  $|\psi\rangle = \beta|0\rangle + \delta|1\rangle$  with  $|\delta| = \sqrt{1 - |\beta|^2}$ ).

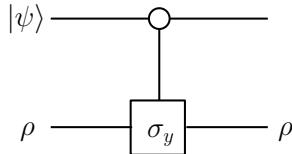


Fig. 7.8 A quantum circuit implementing the bit-phase-flip channel.

**Exercise 7.4** Study the quantum-noise operation implemented by the circuit of Fig. 7.9, where  $U$  is a generic  $2 \times 2$  unitary matrix and  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  a generic single-qubit pure state.

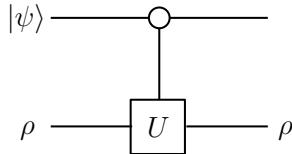


Fig. 7.9 A quantum circuit implementing a single-qubit quantum-noise operation.

### 7.2.7 The depolarizing channel

The depolarizing channel is defined by setting  $|\alpha|^2 = |\beta|^2 = |\gamma|^2 = p/3$  in the state (7.39). We can apply the results of Sec. 7.2.3 to this special case and obtain

$$\rho' = \frac{1}{3}p [\sigma_x \rho \sigma_x^\dagger + \sigma_y \rho \sigma_y^\dagger + \sigma_z \rho \sigma_z^\dagger] + (1 - p)\rho. \quad (7.60)$$

In the Bloch-sphere representation,

$$\mathbf{r}' = \left(1 - \frac{4}{3}p\right) \mathbf{r}. \quad (7.61)$$

Therefore, the Bloch vector is contracted by a factor  $(1 - \frac{4}{3}p)$ , independently of its direction. The centre of the Bloch sphere,  $\mathbf{r} = (0, 0, 0)$ , is the fixed point of this noise channel. If  $p = \frac{3}{4}$ , then  $\mathbf{r}' = (0, 0, 0)$  for any  $\mathbf{r}$ . This corresponds to complete depolarization since, as we saw in Sec. 3.1.2, for this state the qubit polarization along any direction is equal to zero.

The quantum operation (7.60) can be implemented by means of the Kraus operators

$$E_0 = \sqrt{1-p} I, \quad E_1 = \sqrt{\frac{1}{3}p} \sigma_x, \quad E_2 = \sqrt{\frac{1}{3}p} \sigma_y, \quad E_3 = \sqrt{\frac{1}{3}p} \sigma_z. \quad (7.62)$$

### 7.2.8 Amplitude damping

The amplitude-damping channel is defined by

$$\rho' = \mathbb{S}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger, \quad (7.63)$$

where the two Kraus operators  $E_0$  and  $E_1$  read

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{bmatrix}. \quad (7.64)$$

Using this definition, it is straightforward to obtain

$$\rho' = \begin{bmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{10} & (1-p)\rho_{11} \end{bmatrix}, \quad (7.65)$$

where  $\rho_{ij}$  are the matrix elements of the density operator  $\rho$  in the basis  $\{|0\rangle, |1\rangle\}$ . Equation (7.65) implies that the Bloch-sphere coordinates change as follows:

$$x' = \sqrt{1-p} x, \quad y' = \sqrt{1-p} y, \quad z' = p + (1-p)z. \quad (7.66)$$

Therefore, the Bloch sphere is deformed into an ellipsoid, with axes directed along  $x$ ,  $y$  and  $z$  and centre at  $(0, 0, p)$ . It is clearly seen from Eq. (7.66) that  $p$  represents the probability that the state  $|1\rangle$  decays to the state  $|0\rangle$  (*damping probability*). Indeed, if we start from  $\rho = |1\rangle\langle 1|$ ; that is,  $\mathbf{r} = (0, 0, -1)$ , we obtain  $\mathbf{r}' = (0, 0, p - (1-p))$ , corresponding to  $\rho' = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ .

It is instructive to consider the case in which the amplitude-damping channel is applied repeatedly. In this case, we obtain

$$\rho_{11}^{(n)} = (1-p)^n \rho_{11} = e^{n \ln(1-p)} \rho_{11}, \quad (7.67)$$

where  $n$  denotes the number of applications of the channel. Therefore, the probability  $p_1^{(n)}$  to find the qubit in the state  $|1\rangle$  drops exponentially with the number  $n$  of channel iterations:

$$p_1^{(n)} = (1-p)^n p_1^{(0)} = e^{n \ln(1-p)} p_1^{(0)}. \quad (7.68)$$

This means that, for  $n \rightarrow \infty$ , the system is driven to  $\rho^{(\infty)} = |0\rangle\langle 0|$ . We should stress that, even though quantum noise generally transforms pure states into mixed states, in this case, whatever the initial state is (pure or mixed), we always end up with the pure state  $|0\rangle$ .

Of course, it is possible to give a continuous time version of this result. If  $p = \Gamma(\Delta t)$ ,  $t = n(\Delta t)$  is time and we let  $\Delta t \rightarrow 0$ , then

$$p_1(t) = \lim_{\Delta t \rightarrow 0} (1 - \Gamma \Delta t)^{t/\Delta t} p_1(0) = e^{-\Gamma t} p_1(0). \quad (7.69)$$

Therefore,  $\Gamma$  represents the transition rate for  $|1\rangle \rightarrow |0\rangle$ .

**Exercise 7.5** Show that the amplitude-damping channel can be modelled by means of the circuit in Fig. 7.10, where the gates labelled  $\pm\theta/2$  stand for the rotation matrices  $R_y(\mp\theta)$  (see Eq. (5.21)), with  $\cos \theta = \sqrt{p}$ ,  $\sin \theta = \sqrt{1-p}$ .

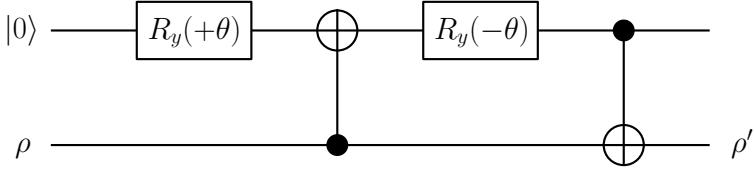


Fig. 7.10 A quantum circuit implementing the amplitude-damping channel.

### 7.2.9 Phase damping

The most general single-qubit density matrix can be written as

$$\rho = \begin{bmatrix} p & \alpha \\ \alpha^* & 1-p \end{bmatrix}, \quad (7.70)$$

where the diagonal, real elements  $p$  and  $1 - p$  ( $0 \leq p \leq 1$ ) represent the probabilities of finding the qubit in the state  $|0\rangle$  or  $|1\rangle$ , respectively. The off-diagonal elements (*quantum coherences*) have no classical analogue. Note that we have  $|\alpha| \leq \sqrt{p(1-p)}$ . As we shall see in the following, the effect of the phase-damping channel is to induce a decay of the off-diagonal terms, a process known as *decoherence*.<sup>2</sup> Therefore, as we shall discuss later in this chapter, the phase-damping channel plays a central role in the transition from the quantum to the classical world.

Two phenomenological models leading to decoherence are the simple example discussed at the beginning of Sec. 7.2 and the quantum circuits implementing the phase-flip channel, introduced in Sec. 7.2.5. Of course, these models are phenomenological and do not represent the physical mechanisms inducing decoherence any better than a resistance in an electric circuit represents the scattering processes undergone by conduction electrons. It is therefore useful to justify decoherence by means of a simple model, leaving a more complete and formal development for the subsequent sections. Our qubit is described by the density matrix (7.70), and we assume that quantum coherences are initially non-zero ( $\alpha \neq 0$ ). We model the effect of the interaction with the environment as a rotation (*phase kick*) through an angle  $\theta$  about the  $z$ -axis of the Bloch sphere. This rotation is described, as we saw in Sec. 3.4.1, by the matrix

$$R_z(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}. \quad (7.71)$$

We assume that the rotation angle is drawn from the random distribution

$$p(\theta) = \frac{1}{\sqrt{4\pi\lambda}} e^{-\frac{\theta^2}{4\lambda}}. \quad (7.72)$$

---

<sup>2</sup>Here it is useful to remind the reader that, more generally, the word decoherence is used to refer to any quantum-noise process due to coupling of the system with the environment.

Therefore, the new density matrix  $\rho'$ , obtained after averaging over  $\theta$ , is given by

$$\rho' = \int_{-\infty}^{+\infty} d\theta p(\theta) R_z(\theta) \rho R_z^\dagger(\theta) = \begin{bmatrix} p & \alpha e^{-\lambda} \\ \alpha^* e^{-\lambda} & 1-p \end{bmatrix}. \quad (7.73)$$

This means that the Bloch-sphere coordinates are mapped by the phase-damping channel as follows:

$$x' = e^{-\lambda} x, \quad y' = e^{-\lambda} y, \quad z' = z. \quad (7.74)$$

Since these transformations coincide with those of Eq. (7.53) (provided we set  $1 - 2|\gamma|^2 = e^{-\lambda}$  in that equation), the phase-damping channel is the same as the phase-flip channel.

**Exercise 7.6** Check Eqs. (7.73) and (7.74).

Notice that, in the case in which the phase-damping channel is applied repeatedly, coherences drop to zero exponentially:  $\alpha_n = e^{-\lambda n} \alpha$ , where  $n$  denotes the number of applications of the channel. Similarly to what was discussed for the amplitude-damping channel, it is possible to give a continuous time version of the coherences decay. If  $\lambda = \Gamma(\Delta t)$ ,  $t = n(\Delta t)$  is time variable and we let  $\Delta t \rightarrow 0$ , then  $\alpha(t) = e^{-\Gamma t} \alpha(0)$ . Therefore,  $\Gamma$  represents the *decoherence rate* associated with this noise channel.

**Exercise 7.7** Study the transformation of the Bloch sphere induced by the circuit of Fig. 7.11, where

$$D = \begin{bmatrix} C_0 & 0 & -S_0 & 0 \\ 0 & C_1 & 0 & -S_1 \\ S_0 & 0 & C_0 & 0 \\ 0 & S_1 & 0 & C_1 \end{bmatrix}, \quad (7.75)$$

with  $C_i \equiv \cos \theta_i$ ,  $S_i \equiv \sin \theta_i$ , ( $i = 0, 1$ ), and

$$U = \exp\left[-i\frac{\xi}{2}(\mathbf{n} \cdot \boldsymbol{\sigma})\right], \quad (7.76)$$

where  $\mathbf{n}$  is a unit vector and  $\xi$  a real number.

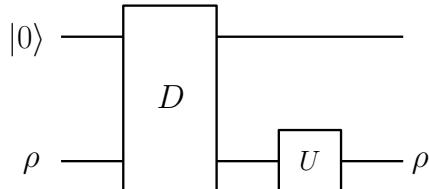


Fig. 7.11 A quantum circuit modelling the noise channel described in exercise 7.7.

**Exercise 7.8** Determine how many quantum-noise operations, characterized by  $4 \times 4$  unitary matrices  $U_1, U_2, \dots$  (see Fig. 7.12) do we need to generate a generic affine map  $\rho \rightarrow \rho' = \mathbb{S}(\rho)$ ?

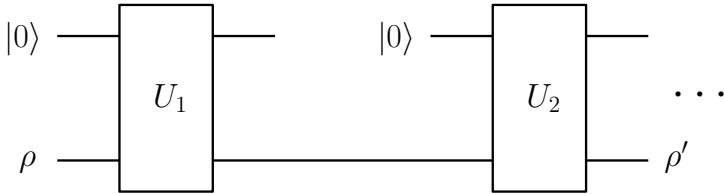


Fig. 7.12 A quantum circuit modelling the noise channel described in exercise 7.8.

**Exercise 7.9** Show that the final density matrix  $\rho'$  in the quantum circuit of Fig. 7.13 is independent of the unitary matrix  $V$  (in this circuit,  $U$  is a generic  $4 \times 4$  unitary matrix).

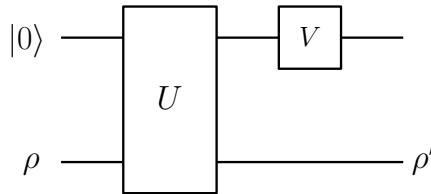


Fig. 7.13 A quantum circuit modelling the noise channel described in exercise 7.9.

**Exercise 7.10** Study the quantum-noise operation implemented by the quantum circuit of Fig. 7.14, in the cases in which the unitary matrix  $U$  is given by (i)  $U = \sigma_x$ , (ii)  $U = \frac{1}{\sqrt{2}}(I \pm i\sigma_j)$  ( $j = x, y, z$ ).

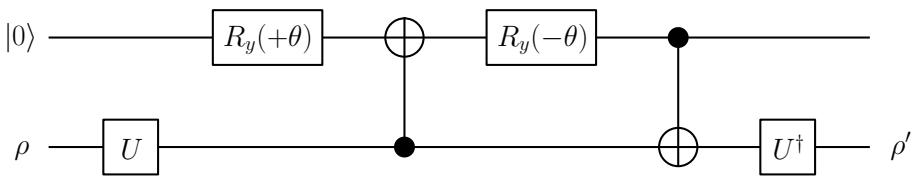


Fig. 7.14 A quantum circuit modelling the noise channel described in exercise 7.10.

### 7.2.10 De-entanglement

Entanglement is arguably the most peculiar feature of quantum systems, with no analogue in classical mechanics. Furthermore, it is an important physical resource for quantum communication and computation. It is therefore important, both for the problem of quantum to classical correspondence and for quantum information science, to investigate the problem of the stability of entanglement in the presence of decoherence effects.

In this section, we shall consider the model drawn in Fig. 7.15. In this quantum circuit, the two most significant qubits are initially prepared in a maximally entangled state, say the Bell state

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \quad (7.77)$$

This corresponds to the density matrix

$$\rho = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| + |01\rangle\langle 10| + |10\rangle\langle 01|), \quad (7.78)$$

whose matrix representation in the basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  is

$$\rho = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (7.79)$$

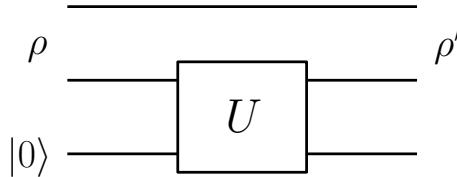


Fig. 7.15 A quantum circuit modelling de-entanglement (loss of the entanglement between the two upper qubits due to the coupling with the third qubit, which represents the environment).

To be concrete, we consider the case in which the interaction with the environment (*i.e.*, the third qubit) is modelled by the phase-flip channel described in Sec. 7.2.5. We have

$$\rho' = \sum_{i=0}^1 F_i^{(2)} \rho (F_i^{(2)})^\dagger, \quad (7.80)$$

where the Kraus operators  $F_0^{(2)}$  and  $F_1^{(2)}$  are defined as

$$F_0^{(2)} = I \otimes F_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cos \theta \end{bmatrix}, \quad (7.81)$$

$$F_1^{(2)} = I \otimes F_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & \sin \theta \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sin \theta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin \theta \end{bmatrix},$$

$F_0$  and  $F_1$  being the Kraus operators (7.54) introduced in Sec. 7.2.5 for the phase-flip channel. It can be seen by direct calculation that

$$\rho' = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \cos \theta & 0 \\ 0 & \cos \theta & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (7.82)$$

For  $\cos \theta = 1$ , we have  $\rho' = \rho$ ; that is, the final state is identical to the initial, maximally entangled, Bell state. For  $\cos \theta = 0$ , the final state is separable. Indeed, we have

$$\rho' = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|), \quad (7.83)$$

and this density matrix corresponds to the statistical mixture of the separable states  $|01\rangle$  and  $|10\rangle$ , taken with equal probabilities. In the intermediate case  $0 < \cos \theta < 1$ , we have partial loss of entanglement, corresponding to the partial loss of quantum coherence discussed in Sec. 7.2.5.

**Exercise 7.11** Study the errors introduced in the teleportation and dense-coding protocols when the partially entangled state (7.82) is used instead of a Bell state.

It is interesting that, if we only consider a member of the Bell pair, its state is preserved with unit fidelity by the phase-flip channel. Indeed, the reduced density matrix describing this state is  $\frac{1}{2}I$ , which is not modified by this noise channel. On the other hand, as seen above, the Bell pair (7.79) is corrupted, even though the phase-flip noise acts only on a member of the pair (the second line in Fig. 7.15). We can intuitively understand this result by saying that it is more difficult to preserve both the state of a system and the entanglement of the system with the outside world (here, the other member of the Bell pair) than just the state of the system. This intuition can be formalized, see Schumacher (1996) and the concept of entanglement fidelity discussed in Sec. 8.4.2.

### 7.3 \* The Bloch-Fano representation

The Bloch-Fano representation (also known as Fano form) is a particularly simple and physically appealing description of quantum operations, in that it directly provides the evolution of the expectation values of the system's polarization measurements.

#### 7.3.1 \* Bloch-Fano representation of a state

To simplify writing, we discuss the Bloch-Fano representation only for qubits, even though the obtained results can be readily extended to qudit systems (that is, to

$d$ -level quantum systems). Any  $n$ -qubit state  $\rho$  can be written in the Bloch-Fano representation as follows:

$$\rho = \frac{1}{N} \sum_{\alpha_1, \dots, \alpha_n = x, y, z, I} r_{\alpha_1 \dots \alpha_n} \sigma_{\alpha_1} \otimes \dots \otimes \sigma_{\alpha_n}, \quad (7.84)$$

where  $N = 2^n$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  are the Pauli matrices, and  $\sigma_I \equiv I$ , and

$$r_{\alpha_1 \dots \alpha_n} = \text{Tr}(\sigma_{\alpha_1} \otimes \dots \otimes \sigma_{\alpha_n} \rho). \quad (7.85)$$

Note that the normalization condition  $\text{Tr}(\rho) = 1$  implies  $r_{I \dots I} = 1$ . Moreover, the generalized Bloch vector  $\mathbf{r} = \{r_\alpha\}_{\alpha=1, \dots, N^2-1}$  is real due to the hermiticity of  $\rho$ . Here  $r_\alpha \equiv r_{\alpha_1 \dots \alpha_n}$ , with  $\alpha \equiv \sum_{k=1}^n i_k 4^{n-k}$ , where we have defined  $i_k = 1, 2, 3, 4$  in correspondence to  $\alpha_k = x, y, z, I$ . Note that from 1 to  $n$  qubits run from the most significant to the least significant. For a single qubit ( $n = 1$ ) we recover the usual Bloch vector,  $\mathbf{r} = (r_1, r_2, r_3) \equiv (x, y, z)$ , while for two qubits ( $n = 2$ ) the  $N^2 - 1 = 15$  components of vector  $\mathbf{r}$  are ordered as follows:

$$\begin{aligned} \mathbf{r}^T &= (r_1, r_2, \dots, r_{15}) \\ &= (r_{xx}, r_{xy}, r_{xz}, r_{xI}, r_{yI}, r_{yy}, r_{yz}, r_{yI}, r_{zx}, r_{zy}, r_{zz}, r_{zI}, r_{Ix}, r_{Iy}, r_{Iz}). \end{aligned} \quad (7.86)$$

If we write the vector representation of the density matrix,

$$\mathbf{v} = [\rho_{11}, \rho_{12}, \dots, \rho_{1N}, \rho_{21}, \dots, \rho_{2N}, \dots, \rho_{N1}, \dots, \rho_{NN}]^T, \quad (7.87)$$

we can perform the transition from the vector representation to the Bloch-Fano representation by means of a  $N \times N$  matrix  $\mathcal{F}$ :

$$\begin{bmatrix} \mathbf{r} \\ 1 \end{bmatrix} = \mathcal{F} \mathbf{v}. \quad (7.88)$$

**Exercise 7.12** Determine  $\mathcal{F}$  and its inverse  $\mathcal{F}^{-1}$  for  $n = 1$  and  $n = 2$  qubits.

### 7.3.2 \* Bloch-Fano representation of a quantum operation

Due to the linearity of quantum mechanics any quantum operation  $\rho \rightarrow \rho' = \mathbb{S}(\rho)$  is represented in the Fano basis  $\{\sigma_{\alpha_1} \otimes \dots \otimes \sigma_{\alpha_n}\}$  by an affine map:

$$\begin{bmatrix} \mathbf{r}' \\ 1 \end{bmatrix} = \mathcal{M} \begin{bmatrix} \mathbf{r} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{c} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ 1 \end{bmatrix}, \quad (7.89)$$

where  $\mathbf{M}$  is a  $(N^2 - 1) \times (N^2 - 1)$  matrix,  $\mathbf{c}$  a column vector of dimension  $N^2 - 1$  and  $\mathbf{0}$  the null vector of the same dimension. All information about the quantum operation  $\mathbb{S}$  is contained in the  $N^4 - N^2$  free elements of matrix  $\mathcal{M}$ , namely in the matrix

$$\chi_F = \begin{bmatrix} \mathbf{M} & \mathbf{c} \end{bmatrix}. \quad (7.90)$$

*Quantum process tomography* is the process of reconstructing un unknown quantum operation from experimental data, that is, of determining the above introduced matrix  $\chi_F$ , which we shall call quantum process matrix. For such purpose, one needs

to prepare  $N^2$  linearly independent initial states  $\{\rho_i\}$ , let them evolve according to the quantum operation  $\mathbb{S}$  and then measure the resulting states  $\{\rho'_i = \mathbb{S}(\rho_i)\}$ . If we call  $\mathcal{R}$  the  $N^2 \times N^2$  matrix whose columns are given by the Fano representation of states  $\{\rho_i\}$  and  $\mathcal{R}'$  the corresponding matrix constructed from states  $\{\rho'_i\}$ , we have

$$\mathcal{R}' = \mathcal{M}\mathcal{R}, \quad (7.91)$$

and therefore

$$\mathcal{M} = \mathcal{R}'\mathcal{R}^{-1}. \quad (7.92)$$

As it is well known (see for instance Nielsen and Chuang, 2000), the standard quantum process tomography can be performed with initial states being product states and local measurements of the final states. As initial states  $\{\rho_i\}$  we can choose the  $4^n$  tensor-product states of the 4 single-qubit states

$$|0\rangle, \quad |1\rangle, \quad \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle). \quad (7.93)$$

To estimate  $\mathcal{R}'$ , one needs to prepare many copies of each initial state  $\rho_i$ , let them evolve according to the quantum operation  $\mathbb{S}$  and then measure observables  $\sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_n}$ . Of course, such measurements can be performed on the computational basis  $\{|0\rangle, |1\rangle\}^{\otimes n}$ , provided each measurement is preceded by suitable single-qubit rotations. Note that in the case of a single qubit we can easily recover the results of Sec. 7.2.2.

For a single-qubit system, the matrices  $\mathcal{R}$  and  $\mathcal{R}^{-1}$  corresponding to basis (7.93) read

$$\mathcal{R} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathcal{R}^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (7.94)$$

As we have discussed in Sec. 7.2, we need  $N^4 - N^2 = 12$  parameters to characterize a generic quantum operation acting on a single qubit. Each parameter describes a particular noise channel (like bit flip, phase flip, amplitude damping, ...) and can be most conveniently visualized as associated with rotations, deformations and displacements of the Bloch ball. Here we point out that these noise channels lead to specific patterns in the quantum process matrix  $\chi_F$ .

For instance, for the phase-damping channel (7.73) we have

$$\mathcal{R}' = \begin{bmatrix} 0 & 0 & e^{-\lambda} & 0 \\ 0 & 0 & 0 & e^{-\lambda} \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \quad (7.95)$$

We can then compute  $\mathcal{M} = \mathcal{R}'\mathcal{R}^{-1}$ , and the first three lines of  $\mathcal{M}$  correspond to the quantum process matrix

$$\chi_F = \begin{bmatrix} e^{-\lambda} & 0 & 0 & 0 \\ 0 & e^{-\lambda} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (7.96)$$

From the form of  $\chi_F$  we can see that the Bloch ball is mapped into an ellipsoid with  $z$  as symmetry axis, consistently with Eq. (7.74).

**Exercise 7.13** Construct the state process matrix  $\chi_F$  for the amplitude-damping channel (7.65).

For two qubits, the coordinates  $\{r_{\alpha_1 \alpha_2}\}$  in the Bloch-Fano representation (7.84) are the expectation values of the polarization measurements  $\{\sigma_{\alpha_1} \otimes \sigma_{\alpha_2}\}$ . The coefficients in the quantum process matrix  $\chi_F$  representing a quantum operation  $\mathbb{S}$  can therefore be interpreted in terms of modification of these expectation values. For instance, in exercise 7.15 we show that the quantum process matrix for fully correlated phase damping has a pattern that allows to clearly distinguish it from the quantum process matrix for the uncorrelated phase-damping channel.

**Exercise 7.14** Compute matrices  $\mathcal{R}$  and  $\mathcal{R}^{-1}$  for two qubits, corresponding to the 16 tensor-product states of single-qubit states (7.93).

**Exercise 7.15** Compute the quantum process matrix  $\chi_F$  for uncorrelated phase-damping acting on both channels with the same noise strength and for the case of a fully correlated phase-damping channel. In the latter case, we model the interaction of the two qubits with the environment as a phase-kick rotating both qubits through the same angle  $\theta$  about the  $z$  axis of the Bloch ball. This rotation is described in the  $\{|0\rangle, |1\rangle\}$  basis by the unitary matrix

$$R_z(\theta) \otimes R_z(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix} \otimes \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}, \quad (7.97)$$

with the rotation angle drawn from random distribution (7.72).

## 7.4 The master equation

The master equation describes the continuous temporal evolution of open quantum systems. In this section, we shall discuss two derivations of this equation. The first one (Sec. 7.4.1) is based on a microscopic model and gives a clear physical picture of the approximations made to arrive at the master equation. The second one (Sec. 7.4.2) clarifies the link between the master-equation approach and the quantum-operation formalism.

Before going into technical details, let us state the main approximations involved in deriving the master equation:

- (i) *Born approximation* – The environment is large and practically unaffected by interaction with system.
- (ii) *Markov approximation* – The system density matrix  $\rho(t)$  evolves under a first-order differential equation in time. Therefore, the knowledge of the density matrix  $\rho(t_0)$  at a given time  $t_0$  is sufficient to determine  $\rho(t)$  at any time  $t > t_0$ . We stress that this is a non-trivial requirement since the system interacts with

the environment, and so, in general, the environmental state at time  $t_0$  depends on the system density matrix  $\rho(t')$  at earlier times  $t' < t_0$ . In other words, the environment acquires information on the system, but this information can flow back, at least in part, to the system. Therefore, the knowledge of the system density matrix  $\rho(t_0)$  at time  $t_0$  is in general not sufficient to determine  $\rho(t)$  at later times. Indeed, we have

$$\begin{aligned}\rho(t_0 + dt) &= \text{Tr}_{\text{env}} [\rho_{\text{tot}}(t_0 + dt)] \\ &= \text{Tr}_{\text{env}} [U(t_0 + dt, t_0) \rho_{\text{tot}}(t_0) U^\dagger(t_0 + dt, t_0)],\end{aligned}\quad (7.98)$$

where  $\rho_{\text{tot}}$  is the density matrix of the system plus environment, whose evolution from time  $t_0$  to time  $t_0 + dt$  is driven by the unitary operator  $U(t_0 + dt, t_0)$ . As stated above,  $\rho_{\text{tot}}(t_0)$  depends on  $\rho(t)$ , for all times  $t \leq t_0$ . This means that we cannot fully determine  $\rho(t_0 + dt)$  from  $\rho(t_0)$  alone. In the Markovian approximation, we assume that the environment is memoryless; that is its state at time  $t_0$  is essentially unaffected by the history of the system. This means that the information flow is essentially one-way, namely from the system to the environment. The Markovian approximation provides a good description of quantum noise if the memory of any effect that the system has on the environment is limited to a time scale much shorter than the time scales of interest for the dynamics of the system.

#### 7.4.1 \* Derivation of the master equation

Let us consider a system in interaction with the environment (also known as a *bath* or *reservoir*). The most general Hamiltonian describing this situation reads as follows:

$$H = H_S \otimes I_R + I_S \otimes H_R + H_{SR} \equiv H_0 + H_{SR}, \quad (7.99)$$

where  $H_S$ ,  $H_R$  and  $H_{SR}$  describe the system, the reservoir and the interaction, respectively.

We call  $\chi$  the density matrix describing the system plus reservoir, and

$$\rho = \text{Tr}_R(\chi) \quad (7.100)$$

the reduced density matrix describing the system. As we saw in Sec. 2.6, the evolution of  $\chi$  is governed by the von Neumann equation:

$$i\hbar\dot{\chi} = [H, \chi]. \quad (7.101)$$

We assume that the interaction is weak so that we shall be able to separate the fast motion, due to  $H_0 = H_S + H_R$ , from the slow motion, due to the interaction  $H_{SR}$ . For this purpose, we exploit the so-called *interaction picture*, defining

$$i\hbar\dot{U}_S = H_S U_S, \quad U_S(0) = I_S, \quad (7.102)$$

$$i\hbar\dot{U}_R = H_R U_R, \quad U_R(0) = I_R, \quad (7.102)$$

$$U = U_S \otimes U_R, \quad (7.103)$$

$$\tilde{\chi} = U^\dagger \chi U, \quad \tilde{H}_{SR} = U^\dagger H_{SR} U. \quad (7.104)$$

After substitution of (7.102)–(7.104) into (7.101), we obtain

$$i\hbar\dot{\tilde{\chi}} = [\tilde{H}_{SR}, \tilde{\chi}], \quad (7.105)$$

which is equivalent to the integro-differential equation

$$\dot{\tilde{\chi}}(t) = \tilde{\chi}(0) + \frac{1}{i\hbar} \int_0^t d\tau [\tilde{H}_{SR}(\tau), \tilde{\chi}(\tau)], \quad (7.106)$$

where  $\tilde{\chi}(0) = \chi(0)$ . We now insert this expression into the right-hand side of Eq. (7.105) and obtain

$$\dot{\tilde{\chi}}(t) = \frac{1}{i\hbar} [\tilde{H}_{SR}(t), \chi(0)] - \frac{1}{\hbar^2} \int_0^t d\tau [\tilde{H}_{SR}(t), [\tilde{H}_{SR}(\tau), \tilde{\chi}(\tau)]] . \quad (7.107)$$

Let us compute  $\text{Tr}_R(\tilde{\chi})$ . We have

$$\text{Tr}_R(\tilde{\chi}) = \text{Tr}_R(U^\dagger \chi U) = U_S^\dagger \text{Tr}_R(U_R^\dagger \chi U_R) U_S = U_S^\dagger \rho U_S \equiv \tilde{\rho}, \quad (7.108)$$

where we have used the definitions (7.100) and (7.103) and the cyclic property of the trace, leading to  $\text{Tr}_R(U_R^\dagger \chi U_R) = \text{Tr}_R(\chi U_R U_R^\dagger) = \text{Tr}_R(\chi)$ . We can now trace Eq. (7.107) over the environmental degrees of freedom and obtain

$$\dot{\tilde{\rho}}(t) = \frac{1}{i\hbar} \text{Tr}_R \left\{ [\tilde{H}_{SR}(t), \chi(0)] \right\} - \frac{1}{\hbar^2} \int_0^t d\tau \text{Tr}_R \left\{ [\tilde{H}_{SR}(t), [\tilde{H}_{SR}(\tau), \tilde{\chi}(\tau)]] \right\} . \quad (7.109)$$

This equation is exact.<sup>3</sup> In order to proceed with the derivation, we need a few assumptions that are detailed below.

#### Factorization of the initial state

First of all, we assume that at time  $t = 0$  the system and the environment are not entangled:

$$\chi(0) = \rho(0) \otimes \rho_R(0), \quad (7.110)$$

where  $\rho$  and  $\rho_R$  denote the system and environment density matrices, respectively. The condition (7.110) enables us to simplify the second term in the right-hand side of Eq. (7.109). Let us now comment this assumption in more details.

A very useful criterion that can be used in order to quantify the interaction strength between the system and the reservoir, is to check whether the presence of interactions, i.e.  $H_{SR}$ , could modify the global system-plus-bath ground state. In order to illustrate this idea, we consider a simple example where the global system is composed of a single qubit and an environment that can be modeled as a single quantum harmonic oscillator. The simplest non-trivial scheme of qubit-oscillator interaction is given by the so-called *Rabi Hamiltonian* (see, e.g., Meystre and Sargent III, 2007), describing in the dipole approximation the interaction between a two-level atom and a single mode of the quantized electromagnetic field:

$$H_{\text{Rabi}} = -\frac{1}{2}\hbar\omega_0\sigma_z + \hbar\omega(a^\dagger a + \frac{1}{2}) + \hbar(\lambda\sigma_+ + \lambda^*\sigma_-)(a^\dagger + a), \quad (7.111)$$

---

<sup>3</sup>Sometimes one also includes an averaging operation over the fluctuations of the bath.

where the Pauli matrices are written in  $\{|g\rangle, |e\rangle\}$  basis spanning the Hilbert space associated with the two-level atom,  $\hbar\omega_0 = E_e - E_g$  is the difference between the energies of the atomic levels  $|g\rangle$  and  $|e\rangle$ ,  $\hbar\omega$  the single-photon energy. The raising and lowering operators  $\sigma_+$  and  $\sigma_-$ ,

$$\sigma_+ = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \sigma_- = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad (7.112)$$

are such that  $\sigma_+|g\rangle = |e\rangle$ ,  $\sigma_+|e\rangle = 0$ ,  $\sigma_-|g\rangle = 0$ ,  $\sigma_-|e\rangle = |g\rangle$ .<sup>4</sup> The single-mode field is described by the Hamiltonian  $\hbar\omega(a^\dagger a + \frac{1}{2})$ , where  $a^\dagger$  and  $a$  the photon creation and annihilation operators. Here we assume the coupling  $\lambda$  to be real.<sup>5</sup> The first two terms in Eq. (7.111) denote the Hamiltonian for the system  $H_S$  and the reservoir  $H_R$ , respectively, while the latest terms describe the interaction.<sup>6</sup>

We stress that, in the absence of interaction, the ground state (the so-called “vacuum state”) of the global system is simply given by the tensor product of the qubit’s ground state and the oscillator’s ground state, and has energy  $E_{\text{vac}} = -\frac{1}{2}\hbar\omega_0 + \frac{1}{2}\hbar\omega$ . In the interacting case, the ground-state energy can be lower than that of the non-interacting vacuum state  $E_{\text{vac}}$ , as it is visible from Fig. 7.16. Therefore, if the interaction is switched on adiabatically, the system may evolve from the vacuum state to the interacting ground state. On the other hand, if the interaction sets in non-adiabatically, the vacuum state may evolve in a state containing photons. This mechanism is discussed in the literature as one of the various effects that are framed under the name of “dynamical Casimir effect”. For what we are interested in here, the absence or presence of such phenomena can be considered as a marker of the system-bath interaction strength.

#### Assumption on the initial condition

In our derivation, we also assume the following condition:

$$\text{Tr}_R \left\{ [\tilde{H}_{SR}, \chi(0)] \right\} = 0. \quad (7.113)$$

If this is not the case,  $\text{Tr}_R \{[\tilde{H}_{SR}, \chi(0)]\}$  is an operator acting on the system alone. Therefore, it is possible to show that we can always redefine  $H_S$  and  $H_{SR}$  (while keeping the global Hamiltonian constant) in order to fulfill Eq. (7.113).

---

<sup>4</sup>Note that, in contrast with the usual notations, for the Rabi model and the Jaynes–Cummings model of Sec. 10.1.2, we found convenient to set  $\sigma_+ = \frac{1}{2}(\sigma_x - i\sigma_y)$  and  $\sigma_- = \frac{1}{2}(\sigma_x + i\sigma_y)$ .

<sup>5</sup>In the absence of the so-called counter-rotating terms proportional to  $a\sigma_-$  and  $a^\dagger\sigma_+$ , Eq. (7.111) would describe the so-called Jaynes–Cummings model, which conserves the total number of excitations  $N_T = \sigma_+\sigma_- + a^\dagger a$ , but does not present the features that we wish to discuss here. We shall discuss the Rabi and Jaynes–Cummings models in more details in Secs. 10.1.2 and Sec. 10.3.3

<sup>6</sup>Despite the apparent simplicity of the Rabi model, and although the numerical calculation of the eigenvectors is easy to be performed (in a truncated Hilbert space, where a cutoff in the maximum admissible number of excitations is introduced), an analytic formal solution has been put forward only recently by Braak (2011). This solution, however, still presents some difficulties, since it is expressed in terms of the zeros of a trascendental function given as a power series of the coupling strength  $\lambda$ .

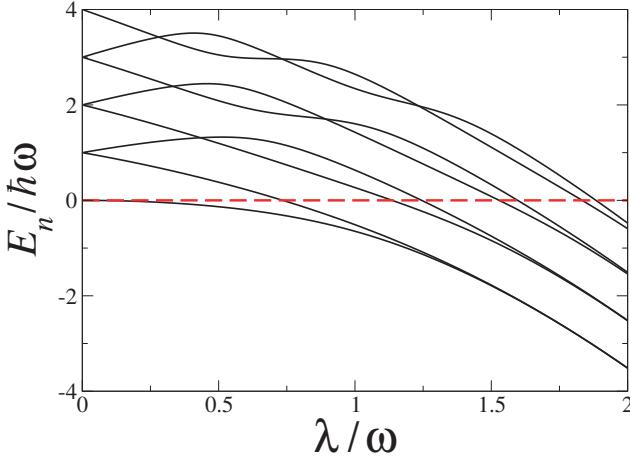


Fig. 7.16 First energy levels of the Rabi model (7.111) as a function of the interaction strength  $\lambda/\omega$ , for  $\omega_0 = \omega$ . The dashed line indicates the energy of the non-interacting vacuum state.

To verify this possibility, we start writing  $H_{SR}$  over a basis of Hermitian operators  $\{\sigma_i\}$  which act on the system:

$$H_{SR} = \sum_{i=0}^{M-1} \sigma_i B_i, \quad (7.114)$$

where the operators  $B_i$  act on the environment. Notice that, if the Hilbert space of the system has dimension  $N$ , we have  $M = N^2$  since the  $N \times N$  matrices constitute a linear vector space of dimension  $N^2$ . For instance, if  $N = 2$ , we can take  $\sigma_0 = I$ ,  $\sigma_1 = \sigma_x$ ,  $\sigma_2 = \sigma_y$  and  $\sigma_3 = \sigma_z$ . Using the interaction picture (7.104), we have

$$\tilde{H}_{SR}(t) = U^\dagger(t) H_{SR} U(t) = \sum_i \tilde{\sigma}_i(t) \tilde{B}_i(t), \quad (7.115)$$

where

$$\tilde{\sigma}_i(t) \equiv U_S^\dagger(t) \sigma_i U_S(t), \quad \tilde{B}_i(t) \equiv U_R^\dagger(t) B_i U_R(t). \quad (7.116)$$

Now, using the expansion of  $\tilde{\mathcal{H}}_{SR}(t)$  in Eq. (7.115), we have

$$\text{Tr}_R \left\{ [\tilde{\mathcal{H}}_{SR}(t), \chi(0)] \right\} = \sum_i \left[ \tilde{\sigma}_i(t), \tilde{\rho}(0) \text{Tr}_R \{ \tilde{B}_i(t) \tilde{\rho}_R(0) \} \right], \quad (7.117)$$

where we also employed the hypothesis of a factorized initial condition (7.110). From this expression we can see that, in order to evaluate the first term in the right-hand side of Eq. (7.109) under the factorization condition, we need to know the initial expectation values of the operators  $\tilde{B}_i$ . If we now suppose that

$$\text{Tr}_R \{ \tilde{B}_i(t) \tilde{\rho}_R(0) \} = 0, \quad (7.118)$$

then such term would be trivially vanishing. On the other hand, if (7.118) is not verified, we can substitute

$$\tilde{\mathcal{H}}_{SR}(t) \rightarrow \sum_i \tilde{\sigma}_i(t) [\tilde{B}_i(t) - \text{Tr}_R \{ \tilde{B}_i(t) \tilde{\rho}_R(0) \}] \quad (7.119)$$

and simultaneously sum  $\sum_i \tilde{\sigma}_i(t) \text{Tr}_R \{ \tilde{B}_i(t) \tilde{\rho}_R(0) \}$  to  $\tilde{\mathcal{H}}_{SR}(t)$ .

### Born approximation

We assume that the coupling is so weak and the reservoir so large that its state is essentially unaffected by the interaction,  $\rho_R(0) = \rho_R(\tau)$ . Therefore,

$$\tilde{\chi}(\tau) \approx \tilde{\rho}(\tau) \otimes \tilde{\rho}_R, \quad (7.120)$$

so that Eq. (7.109) becomes

$$\dot{\tilde{\rho}}(t) = -\frac{1}{\hbar^2} \int_0^t d\tau \text{Tr}_R \left\{ \left[ \tilde{H}_{SR}(t), [\tilde{H}_{SR}(\tau), \tilde{\rho}(\tau)\tilde{\rho}_R] \right] \right\}. \quad (7.121)$$

In order to interpret this assumption, we need to consider a situation where the reservoir equilibrates on a time scales that is much faster than the typical time scale of the system's evolution. Moreover the reservoir needs to be so large that it stays always in the initial state, and it always satisfies the factorization condition. Let us stress that Eq. (7.121) is still an integro-differential equation, and needs to be further simplified, before being considered suitable for practical purposes.

### Markov approximation

We assume to perform the following replacement in Eq. (7.121):

$$\tilde{\rho}(\tau) \rightarrow \tilde{\rho}(t), \quad (7.122)$$

thus obtaining an equation

$$\dot{\tilde{\rho}}(t) = -\frac{1}{\hbar^2} \int_0^t d\tau \text{Tr}_R \left\{ \left[ \tilde{H}_{SR}(t), [\tilde{H}_{SR}(\tau), \tilde{\rho}(t)\tilde{\rho}_R] \right] \right\}, \quad (7.123)$$

which is no longer integro-differential but simply differential, since now  $\tilde{\rho}(t)$  on the right-hand side of Eq (7.123) has not to be integrated in time. It is important to remark that the substitution (7.122) is possible under the hypothesis that the typical time scale of the system's evolution is much longer than that of the reservoir. This assumption holds only in the cases where the system-bath interaction is sufficiently small and the bath is sufficiently large. It is instructive to test such hypothesis numerically in simple models. This will be done in Sec. 7.5, where we shall consider an exactly solvable system composed of a single harmonic oscillator interacting with a bath consisting of a *finite* number of harmonic oscillators.

We finally insert (7.115) and (7.116) into (7.123) and obtain

$$\dot{\tilde{\rho}}(t) = -\frac{1}{\hbar^2} \sum_{i,j} \int_0^t d\tau \text{Tr}_R \left\{ \left[ \tilde{\sigma}_i(t)\tilde{B}_i(t), [\tilde{\sigma}_j(\tau)\tilde{B}_j(\tau), \tilde{\rho}(t)\tilde{\rho}_R] \right] \right\}. \quad (7.124)$$

In order to proceed beyond this differential equation, we shall specify more in details our models of interest.

### 7.4.1.1 Damping of an harmonic oscillator

Let us consider an harmonic oscillator ( $S$ ) that is coupled to a reservoir ( $R$ ) which can be modeled by a set of harmonic oscillators, such that

$$H_S = \hbar\omega_0(a^\dagger a + \frac{1}{2}), \quad H_R = \sum_{j=1}^{\infty} \hbar\omega_j(b_j^\dagger b_j + \frac{1}{2}), \quad H_{SR} = \hbar(a\Gamma^\dagger + a^\dagger\Gamma), \quad (7.125)$$

where in the latter equation we defined

$$\Gamma \equiv \sum_{j=1}^{\infty} k_j b_j. \quad (7.126)$$

We also suppose that the bath of harmonic oscillators is initially in thermal equilibrium, at temperature  $T$ . Since the different modes  $b_j$  are independent, we have

$$\rho_R^{(0)} = \prod_j \left[ 1 - \exp\left(-\frac{\hbar\omega_j}{k_B T}\right) \right] \exp\left(-\frac{\hbar\omega_j b_j^\dagger b_j}{k_B T}\right), \quad (7.127)$$

with  $k_B$  being the Boltzmann constant.

From the definition of  $H_{SR}$  in (7.125), and using Eqs. (7.102)–(7.104), we get

$$\tilde{H}_{SR}(t) = e^{\frac{i}{\hbar}(H_S+H_R)t} H_{SR} e^{-\frac{i}{\hbar}(H_S+H_R)t} = \hbar\{\tilde{a}(t)\tilde{\Gamma}^\dagger(t) + \tilde{a}^\dagger(t)\tilde{\Gamma}(t)\}, \quad (7.128)$$

where

$$\tilde{a}(t) = e^{\frac{i}{\hbar}H_S t} a e^{-\frac{i}{\hbar}H_S t} = e^{i\omega_0 a^\dagger a t} a e^{-i\omega_0 a^\dagger a t}, \quad (7.129)$$

and similarly for  $\tilde{\Gamma}$ . This expression can be explicitly evaluated by expanding the exponentials in series:

$$\begin{aligned} \tilde{a}(t) &= \left\{ 1 + i\omega_0 a^\dagger a t + \frac{(i\omega_0 a^\dagger a t)^2}{2!} + \dots \right\} a \left\{ 1 - i\omega_0 a^\dagger a t + \frac{(i\omega_0 a^\dagger a t)^2}{2!} + \dots \right\} \\ &= a + i\omega_0 a^\dagger a t - i\omega_0 a a^\dagger a t + \dots = \dots = a e^{-i\omega_0 t}, \end{aligned} \quad (7.130)$$

where we also used the bosonic commutation relation  $[a, a^\dagger] = 1$ . Proceeding in a similar way, we also obtain

$$\tilde{\Gamma}(t) = e^{\frac{i}{\hbar}H_R t} \Gamma e^{-\frac{i}{\hbar}H_R t} = \sum_j k_j b_j e^{-i\omega_j t}. \quad (7.131)$$

Now, inserting (7.128) into (7.124), we obtain

$$\begin{aligned} \dot{\tilde{\rho}}(t) &= -\frac{1}{\hbar^2} \int_0^t d\tau \text{Tr}_R \left\{ \left[ \hbar\{\tilde{a}(t)\tilde{\Gamma}^\dagger(t) + \tilde{a}^\dagger(t)\tilde{\Gamma}(t)\}, \right. \right. \\ &\quad \left. \left. [\hbar\{\tilde{a}(\tau)\tilde{\Gamma}^\dagger(\tau) + \tilde{a}^\dagger(\tau)\tilde{\Gamma}(\tau)\}, \tilde{\rho}(t)\rho_R^{(0)}] \right] \right\} \end{aligned} \quad (7.132)$$

$$\begin{aligned} &= - \int_0^t d\tau \text{Tr}_R \left\{ \left[ [\tilde{a}(t)\tilde{\Gamma}^\dagger(t), [\tilde{a}(\tau)\tilde{\Gamma}^\dagger(\tau), \tilde{\rho}(t)\rho_R^{(0)}]] + [\tilde{a}^\dagger(t)\tilde{\Gamma}(t), [\tilde{a}(\tau)\tilde{\Gamma}^\dagger(\tau), \tilde{\rho}(t)\rho_R^{(0)}]] \right] \right. \\ &\quad \left. + [\tilde{a}(t)\tilde{\Gamma}^\dagger(t), [\tilde{a}^\dagger(\tau)\tilde{\Gamma}(\tau), \tilde{\rho}(t)\rho_R^{(0)}]] + [\tilde{a}^\dagger(t)\tilde{\Gamma}(t), [\tilde{a}^\dagger(\tau)\tilde{\Gamma}(\tau), \tilde{\rho}(t)\rho_R^{(0)}]] \right\}. \end{aligned} \quad (7.133)$$

By expanding the various commutators, we obtain sixteen terms in the integrand, which can be further simplified using the cyclicity property of the trace. These are given by:

$$+ \tilde{a}(t) \tilde{a}(\tau) \tilde{\rho}(t) \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}^\dagger(\tau) \rangle, \quad (7.134)$$

$$- \tilde{a}(\tau) \tilde{\rho}(t) \tilde{a}(t) \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}^\dagger(\tau) \rangle, \quad (7.135)$$

$$+ \tilde{a}^\dagger(t) \tilde{a}(\tau) \tilde{\rho}(t) \langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle, \quad (7.136)$$

$$- \tilde{a}(\tau) \tilde{\rho}(t) \tilde{a}^\dagger(t) \langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle, \quad (7.137)$$

$$+ \tilde{a}(t) \tilde{a}^\dagger(\tau) \tilde{\rho}(t) \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle, \quad (7.138)$$

$$- \tilde{a}^\dagger(\tau) \tilde{\rho}(t) \tilde{a}(t) \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle, \quad (7.139)$$

$$+ \tilde{a}^\dagger(t) \tilde{a}^\dagger(\tau) \tilde{\rho}(t) \langle \tilde{\Gamma}(t) \tilde{\Gamma}(\tau) \rangle, \quad (7.140)$$

$$- \tilde{a}^\dagger(\tau) \tilde{\rho}(t) \tilde{a}^\dagger(t) \langle \tilde{\Gamma}(t) \tilde{\Gamma}(\tau) \rangle, \quad (7.141)$$

where average expressions like  $\langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle = \text{Tr}_R[\tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau)]$  denote two-point correlations. It is then easy to show that

$$\langle \tilde{\Gamma}(\tau) \tilde{\Gamma}(t) \rangle = \langle \tilde{\Gamma}^\dagger(\tau) \tilde{\Gamma}^\dagger(t) \rangle = 0, \quad (7.142)$$

since the expectation value of the product of two annihilation or creation operators is zero. Let us now detail the calculation of

$$\begin{aligned} \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle &= \text{Tr}_R \left[ \sum_i k_i^* b_i^\dagger e^{i\omega_i t} \sum_l k_l b_l e^{-i\omega_l \tau} \rho_R^{(0)} \right] \\ &= \text{Tr}_R \left[ \sum_{i,l} k_i^* k_l b_i^\dagger b_l e^{i(\omega_i t - \omega_l \tau)} \prod_j \left( 1 - e^{-\frac{\hbar\omega_j}{k_B T}} \right) e^{-\frac{\hbar\omega_j b_j^\dagger b_j}{k_B T}} \right]. \end{aligned} \quad (7.143)$$

It is clear that in the latter expression, when tracing over  $R$ , the terms with  $i \neq l$  are zero, therefore by retaining only the terms with  $i = l$ , we finally obtain

$$\langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle = \sum_i |k_i|^2 e^{i\omega_i(t-\tau)} \bar{n}(\omega_i, T), \quad (7.144)$$

where  $n_i = b_i^\dagger b_i$ , and  $\bar{n}(\omega_i, T) = [\exp(\hbar\omega_i/k_B T) - 1]^{-1}$  denotes the average occupation number of the  $i$ -th mode at temperature  $T$ . Analogously, using the commutation relation  $[b_i, b_j^\dagger] = \delta_{ij}$  which implies  $b_i b_i^\dagger = b_i^\dagger b_i + 1$ , we obtain

$$\langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle = \sum_i |k_i|^2 e^{-i\omega_i(t-\tau)} [\bar{n}(\omega_i, T) + 1]. \quad (7.145)$$

If we now suppose that the number of modes is sufficiently dense, we can substitute the sum with an integral over the density of states  $g(\omega)$ :

$$\langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle = \int_0^\infty d\omega e^{i\omega(t-\tau)} g(\omega) |k(\omega)|^2 \bar{n}(\omega, T), \quad (7.146)$$

$$\langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle = \int_0^\infty d\omega e^{-i\omega(t-\tau)} g(\omega) |k(\omega)|^2 [\bar{n}(\omega, T) + 1]. \quad (7.147)$$

Eventually, the expressions (7.142), (7.146), (7.147) have to be substituted in the reduced master equation (7.133). From the sixteen integrand terms written above, only the ones in (7.136)–(7.139) will survive.

Collecting everything we obtain

$$\begin{aligned} \dot{\tilde{\rho}}(t) = & - \int_0^t d\tau \left\{ \left[ (\tilde{a}(t) \tilde{a}^\dagger(\tau) \tilde{\rho}(t) - \tilde{a}^\dagger(\tau) \tilde{\rho}(t) \tilde{a}(t)) \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle + \text{H.c.} \right] \right. \\ & \left. + \left[ (\tilde{a}^\dagger(t) \tilde{a}(\tau) \tilde{\rho}(t) - \tilde{a}(\tau) \tilde{\rho}(t) \tilde{a}^\dagger(t)) \langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle + \text{H.c.} \right] \right\}, \end{aligned} \quad (7.148)$$

which can be further simplified by using (7.130), that implies  $\tilde{a}(t)\tilde{a}^\dagger(\tau) = aa^+e^{-i\omega_0(t-\tau)}$ :

$$\begin{aligned} \dot{\tilde{\rho}}(t) = & - \int_0^t d\tau \left\{ \left[ (aa^\dagger \tilde{\rho}(t) - a^\dagger \tilde{\rho}(t) a) e^{-i\omega_0(t-\tau)} \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle + \text{H.c.} \right] \right. \\ & \left. + \left[ (a^\dagger a \tilde{\rho}(t) - a \tilde{\rho}(t) a^\dagger) e^{i\omega_0(t-\tau)} \langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle + \text{H.c.} \right] \right\}. \end{aligned} \quad (7.149)$$

This expression is easy to be handled, since we only have to integrate the terms in  $\tau$ , while, due to Born and Markov approximations, the density matrix  $\tilde{\rho}(t)$  does not depend on  $\tau$ . From (7.146) and (7.147) we obtain

$$\int_0^t d\tau e^{-i\omega_0(t-\tau)} \langle \tilde{\Gamma}^\dagger(t) \tilde{\Gamma}(\tau) \rangle = \beta^*, \quad \int_0^t d\tau e^{i\omega_0(t-\tau)} \langle \tilde{\Gamma}(t) \tilde{\Gamma}^\dagger(\tau) \rangle = \beta + \alpha, \quad (7.150)$$

where we defined

$$\alpha \equiv \int_0^t d\tau \int_0^\infty d\omega e^{i(\omega_0 - \omega)(t-\tau)} g(\omega) |k(\omega)|^2, \quad (7.151)$$

$$\beta \equiv \int_0^t d\tau \int_0^\infty d\omega e^{i(\omega_0 - \omega)(t-\tau)} g(\omega) |k(\omega)|^2 \bar{n}(\omega, T). \quad (7.152)$$

Substituting these expressions into (7.149), we obtain

$$\dot{\tilde{\rho}} = \{\alpha(a\tilde{\rho}a^\dagger - a^\dagger a\tilde{\rho}) + \text{H.c.}\} + \{\beta(a\tilde{\rho}a^\dagger + a^\dagger \tilde{\rho}a - a^\dagger a\tilde{\rho} - \tilde{\rho}aa^\dagger) + \text{H.c.}\}. \quad (7.153)$$

The expressions for  $\alpha$  and  $\beta$  can be simplified by noticing that the time  $t$  in Eqs. (7.151) and (7.152) needs to be of the order of the typical time scale of the system  $\tilde{\rho}(t)$ , which is much larger than the times characterizing the reservoir's dynamics. One can thus safely extend the integration time to infinity. Using now

$$\lim_{t \rightarrow +\infty} \int_0^t d\tau e^{-i(\omega - \omega_0)\tau} = \pi \delta(\omega - \omega_0) + i \text{P} \left\{ \frac{1}{\omega_0 - \omega} \right\}, \quad (7.154)$$

where  $\text{P}\{\cdot\}$  indicates the Cauchy principal value, it is possible to write

$$\alpha = k + i\Delta, \quad \beta = k \cdot \bar{n}(\omega_0, T) + i\Delta', \quad (7.155)$$

after defining  $k \equiv \pi g(\omega_0) |k(\omega_0)|^2$ ,  $\bar{n} \equiv \bar{n}(\omega_0, T)$ , and

$$\Delta \equiv \text{P} \left\{ \int_0^\infty d\omega \frac{g(\omega) |k(\omega)|^2}{\omega_0 - \omega} \right\}, \quad \Delta' \equiv \text{P} \left\{ \int_0^\infty d\omega \frac{g(\omega) |k(\omega)|^2}{\omega_0 - \omega} \bar{n}(\omega, T) \right\}. \quad (7.156)$$

Substituting (7.155) into (7.153), we obtain the final form of the master equation:

$$\dot{\tilde{\rho}} = -i\Delta[a^\dagger a, \tilde{\rho}] + k(2a\tilde{\rho}a^\dagger - a^\dagger a\tilde{\rho} - \tilde{\rho}a^\dagger a) + 2k\bar{n}(\omega_0, T)(a\tilde{\rho}a^\dagger + a^\dagger \tilde{\rho}a - a^\dagger a\tilde{\rho} - \tilde{\rho}aa^\dagger). \quad (7.157)$$

### 7.4.1.2 Damping of a qubit

We can repeat a similar calculation for a qubit coupled to a set of harmonic oscillators. The analogous of Hamiltonians (7.125) is now provided by

$$H_S = -\frac{1}{2}\hbar\omega_0\sigma_z, \quad H_R = \sum_{j=1}^{\infty} \hbar\omega_j(b_j^\dagger b_j + \frac{1}{2}), \quad H_{SR} = \hbar(\sigma_- \Gamma^\dagger + \sigma_+ \Gamma). \quad (7.158)$$

Using a series expansion analogous to (7.130), we obtain

$$\tilde{\sigma}_-(t) = e^{\frac{i}{\hbar}H_S t} \sigma_- e^{-\frac{i}{\hbar}H_S t} = \sigma_- e^{-i\omega_0 t}, \quad (7.159)$$

and similarly for  $\tilde{\sigma}_+(t)$ . From (7.132) to (7.141) it is thus sufficient to substitute  $a \rightarrow \sigma_-$  and  $a^\dagger \rightarrow \sigma_+$ , since the order of operators has not been changed. Eventually we need to use commutation relations for spins, and thus we obtain

$$\begin{aligned} \dot{\tilde{\rho}} = & k(\bar{n} + 1)(2\sigma_- \tilde{\rho} \sigma_+ - \sigma_+ \sigma_- \tilde{\rho} - \tilde{\rho} \sigma_+ \sigma_-) + k\bar{n}(2\sigma_+ \tilde{\rho} \sigma_- - \sigma_- \sigma_+ \tilde{\rho} - \tilde{\rho} \sigma_- \sigma_+) \\ & + i(\Delta + \Delta')(-\sigma_+ \sigma_- \tilde{\rho} + \tilde{\rho} \sigma_+ \sigma_-) + i\Delta'(\sigma_- \sigma_+ \tilde{\rho} - \tilde{\rho} \sigma_- \sigma_+). \end{aligned} \quad (7.160)$$

### 7.4.1.3 General case

In the general case, the expression (7.124) can be further manipulated to obtain the following final form for the master equation describing the evolution of an  $N$ -level system coupled to an environment (see Gorini *et al.*, 1976):

$$\dot{\rho} = -\frac{i}{\hbar}[H_S, \rho] + \frac{1}{\hbar^2} \sum_{i,j} \frac{\gamma_{ji}}{2} \left\{ [\sigma_i, \rho \sigma_j] + [\sigma_i \rho, \sigma_j] \right\}. \quad (7.161)$$

The complex constants  $\gamma_{ji}$  depend on the details of the Markovian bath, and have to be calculated in order to determine the system dynamics.

Here we prefer to skip any tedious technical detail, and observe that Eq. (7.161) describes the most general non-unitary evolution of a density matrix  $\rho$ , under the form of a Markovian time-homogeneous equation. Specifically, the map  $\rho(0) \rightarrow \rho(t)$  is a superoperator, since it can be proven to be trace-preserving and completely positive for any initial condition, provided the coefficient matrix  $[\gamma_{ji}]$  is positive.

## 7.4.2 The master equation and quantum operations

It is instructive to derive the master equation in an alternative way, within the framework of the quantum-operation formalism. In the Kraus representation, the density matrices  $\rho(t)$  and  $\rho(t+dt)$ , which describe the system at times  $t$  and  $t+dt$ , are related as follows:

$$\rho(t+dt) = \mathbb{S}(t+dt, t)\rho(t) = \sum_{k=0}^{M-1} E_k \rho(t) E_k^\dagger, \quad (7.162)$$

where  $\mathbb{S}(t+dt, t)$  is the superoperator mapping  $\rho(t)$  into  $\rho(t+dt)$  and the operators  $E_k$  are the Kraus operators, whose number  $M$  is  $\leq N^2$  ( $N$  is the dimension of the

Hilbert space). Note that the operators  $E_k$  in (7.162), in contrast with Sec. 7.1, refer to infinitesimal transformations. In order to assure that  $\mathbb{S}(t, t)$  is equal to the identity, we may write the Kraus operators as follows:

$$\begin{aligned} E_0 &= I + \frac{1}{\hbar}(-iH + K)dt, \\ E_k &= L_k\sqrt{dt}, \quad (k = 1, \dots, M-1), \end{aligned} \quad (7.163)$$

where  $H$  and  $K$  are Hermitian operators, and the operators  $L_k$  are known as the *Lindblad operators*. The normalization condition  $\sum_k E_k^\dagger E_k = I$  gives

$$\left[ I + \frac{1}{\hbar}(iH + K)dt \right] \left[ I + \frac{1}{\hbar}(-iH + K)dt \right] + \sum_{k=1}^{M-1} L_k^\dagger L_k dt = I, \quad (7.164)$$

that is,

$$\frac{2}{\hbar} Kdt + \sum_{k=1}^{M-1} L_k^\dagger L_k dt + O((dt)^2) = 0. \quad (7.165)$$

Therefore,

$$K = -\frac{\hbar}{2} \sum_{k=1}^{M-1} L_k^\dagger L_k. \quad (7.166)$$

We now insert (7.163) and (7.166) into (7.162) and obtain

$$\rho(t+dt) = \rho(t) - \frac{i}{\hbar} [H, \rho(t)] dt + \sum_{k=1}^{M-1} \left( L_k \rho(t) L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho(t) - \frac{1}{2} \rho(t) L_k^\dagger L_k \right) dt + O(dt^2). \quad (7.167)$$

If we assume that

$$\rho(t+dt) = \rho(t) + \dot{\rho}(t)dt + O((dt)^2), \quad (7.168)$$

we obtain the GKLS (Gorini, Kossakowski, Lindblad and Sudarshan) master equation (see Gorini *et al.*, 1976 and Lindblad, 1976):

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho] + \sum_{k=1}^{M-1} \left( L_k \rho L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho - \frac{1}{2} \rho L_k^\dagger L_k \right). \quad (7.169)$$

We should stress that the expansion (7.168) is possible under the Markovian approximation previously discussed. We should also point out that the quantum-operation formalism is more general than the master-equation approach. Indeed, a quantum process described in terms of an operator-sum representation is, in general, non-Markovian and therefore cannot be described by means of a Markovian master equation.

**Exercise 7.16** As an example application of the GKLS master equation, we consider a two-level atom in a thermal radiation field. In such a situation, the master equation reads (see, *e.g.*, Gardiner and Zoller, 2000)

$$\begin{aligned} \dot{\rho} &= -\frac{i}{\hbar} [H, \rho] + \gamma (\bar{n} + 1) \left( \sigma_- \rho \sigma_+ - \frac{1}{2} \sigma_+ \sigma_- \rho - \frac{1}{2} \rho \sigma_+ \sigma_- \right) \\ &\quad + \gamma \bar{n} \left( \sigma_+ \rho \sigma_- - \frac{1}{2} \sigma_- \sigma_+ \rho - \frac{1}{2} \rho \sigma_- \sigma_+ \right), \end{aligned} \quad (7.170)$$

where the Pauli matrices are written in  $\{|g\rangle, |e\rangle\}$  basis ( $|g\rangle$  and  $|e\rangle$  stand for the ground and excited state, respectively),

$$H = -\frac{1}{2} \hbar \omega_0 \sigma_z, \quad (7.171)$$

so that  $\omega_0$  is the frequency of the radiation that the atom will emit or absorb,  $\sigma_+ = \frac{1}{2}(\sigma_x - i\sigma_y)$  and  $\sigma_- = \frac{1}{2}(\sigma_x + i\sigma_y) = \sigma_+^\dagger$  are the raising and lowering operators and  $\bar{n}$  represents the mean occupation number at temperature  $T$  ( $\bar{n} = 1/[\exp(\hbar\omega_0/k_B T) - 1]$ ). Note that Eq. (7.160) reduces to Eq. (7.170), provided  $\Delta = \Delta' = 0$  and  $\gamma = \frac{k}{2}$ . In (7.170) the Lindblad operators are  $L_1 = \sqrt{\gamma(\bar{n} + 1)} \sigma_-$  and  $L_2 = \sqrt{\gamma\bar{n}} \sigma_+$ . While  $L_1$  drives the transition  $|e\rangle \rightarrow |g\rangle$ ,  $L_2$  induces the jump  $|g\rangle \rightarrow |e\rangle$ . Solve the master equation (7.170). In particular, discuss the approach to equilibrium.

It is possible to show that the master equations (7.161) and (7.169) are equivalent. First of all, for the expansion (7.114) we choose a basis  $\{\sigma_0, \dots, \sigma_{N^2-1}\}$  such that  $\sigma_0 = I/\sqrt{N}$ ,  $\text{Tr}(\sigma_i) = 0$ ,  $\text{Tr}(\sigma_i^\dagger \sigma_j) = \delta_{ij}$  ( $i, j = 1, \dots, N^2 - 1$ ). In this basis, the master equation can be expressed in the form given by Gorini *et al.* (1976):

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho] + \frac{1}{2} \sum_{i,j=1}^{N^2-1} A_{ij} \left\{ [\sigma_i, \rho \sigma_j^\dagger] + [\sigma_i \rho, \sigma_j^\dagger] \right\}, \quad (7.172)$$

where  $H = H^\dagger$  and  $A$  is a positive complex matrix. The term  $-(i/\hbar)[H, \rho]$  describes the Hamiltonian part in the evolution of the density matrix, while the other terms in the right-hand side of (7.172) describe quantum-noise processes. Note that  $H$  is not necessarily the same as the system Hamiltonian since it may include a non-dissipative contribution coming from the interaction with the environment.

Let us introduce the matrix-valued vectors

$$\boldsymbol{\sigma} \equiv \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_{N^2-1} \end{bmatrix}, \quad \mathbf{w} \equiv \begin{bmatrix} w_1 \\ \vdots \\ w_{N^2-1} \end{bmatrix}, \quad (7.173)$$

where

$$\mathbf{w}^\dagger \equiv S \boldsymbol{\sigma}^\dagger \quad (7.174)$$

and the matrix  $S$  is such that

$$\tilde{A} \equiv SAS^\dagger \quad (7.175)$$

is diagonal. Let  $\{\lambda_i\}$  denote the eigenvalues of  $A$ . Since  $A$  is a positive matrix,  $\lambda_i \geq 0$  for all  $i$ . We order these eigenvalues in such a manner that  $\lambda_i > 0$  for  $i = 1, \dots, M$  ( $M \leq N^2 - 1$ ) and  $\lambda_i = 0$  for  $i = M + 1, \dots, N^2 - 1$ . After defining the vector

$$\mathbf{L} \equiv \begin{bmatrix} L_1 \\ \vdots \\ L_{N^2-1} \end{bmatrix} \equiv \begin{bmatrix} \sqrt{\lambda_1} w_1^\dagger \\ \vdots \\ \sqrt{\lambda_{N^2-1}} w_{N^2-1}^\dagger \end{bmatrix}, \quad (7.176)$$

Eq. (7.172) reduces to the GKLS master equation (7.169),  $L_1, \dots, L_M$  being the Lindblad operators.

**Exercise 7.17** Using Eqs. (7.173)–(7.176), show the equivalence of the master equations (7.172) and (7.169).

### 7.4.3 The master equation for a single qubit

In this section, we study, in the Markovian approximation, the most general evolution of the density matrix for a single qubit. The master equation (7.172) is given by

$$\dot{\rho}(t) = -\frac{i}{\hbar} [H, \rho(t)] + \frac{1}{2} \sum_{i,j=1}^3 A_{ij} \{ [\sigma_i, \rho(t)\sigma_j] + [\sigma_i\rho(t), \sigma_j] \}, \quad (7.177)$$

where the  $\{\sigma_i\}$  are the Pauli matrices ( $\sigma_1 = \sigma_x$ ,  $\sigma_2 = \sigma_y$ ,  $\sigma_3 = \sigma_z$ ;  $\sigma_i^\dagger = \sigma_i$ ) and  $A$  is a Hermitian matrix. We assume that both the Hamiltonian  $H$  and the matrix  $A$  are time-independent. In other words, we assume that the environment is stationary and not modified by the interaction with the system. As shown in Sec. 7.4.1, the parameters  $A_{ij}$  can be identified with the bath correlation functions.

The Hamiltonian  $H$  describes the reversible part of the qubit dynamics. It can be written as follows:

$$H = \frac{\hbar}{2} (\omega_0 \sigma_3 + \Delta \sigma_1 + \Delta' \sigma_2). \quad (7.178)$$

Therefore, the Hamiltonian part of the evolution depends on the three real parameters  $\omega_0$ ,  $\Delta$  and  $\Delta'$ . Since the dissipative part of the evolution is governed by the  $3 \times 3$  Hermitian matrix  $A$ , it depends on 9 real parameters. Therefore, the evolution (7.177) of the single-qubit density matrix is determined by 12 independent real parameters. These parameters correspond to the 12 parameters appearing in the Kraus representation of the most general quantum-noise process acting on a single qubit.

It is useful to gain an intuitive understanding of the effect of these parameters on the evolution of the single-qubit density matrix. For this purpose, we employ the Bloch-sphere representation, in which

$$\rho(t) = \frac{1}{2} \begin{bmatrix} 1+z(t) & x(t)-iy(t) \\ x(t)+iy(t) & 1-z(t) \end{bmatrix} = \frac{1}{2} [I + \mathbf{r}(t) \cdot \boldsymbol{\sigma}], \quad (7.179)$$

where  $\mathbf{r}(t) = (x(t), y(t), z(t))$ . We can derive a first-order differential equation for the Bloch vector:

$$\dot{\mathbf{r}}(t) = M\mathbf{r}(t) + \mathbf{c}. \quad (7.180)$$

Indeed, if we insert (7.179) into (7.177), for the Hamiltonian part of the evolution we obtain

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}_H = \begin{bmatrix} 0 & -\omega_0 & \Delta' \\ \omega_0 & 0 & -\Delta \\ -\Delta' & \Delta & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad (7.181)$$

and, for the dissipative part,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}_D = \begin{bmatrix} -2(A_{22}+A_{33}) & (A_{12}+A_{21}) & (A_{13}+A_{31}) \\ (A_{12}+A_{21}) & -2(A_{33}+A_{11}) & (A_{23}+A_{32}) \\ (A_{13}+A_{31}) & (A_{23}+A_{32}) & -2(A_{11}+A_{22}) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 2i(A_{23}-A_{32}) \\ 2i(A_{31}-A_{13}) \\ 2i(A_{12}-A_{21}) \end{bmatrix}. \quad (7.182)$$

After introducing the new parameters

$$\begin{aligned} \gamma_1 &= 2(A_{22} + A_{33}), & \gamma_2 &= 2(A_{33} + A_{11}), & \gamma_3 &= 2(A_{11} + A_{22}), \\ \alpha &= (A_{12} + A_{21}), & \beta &= (A_{13} + A_{31}), & \gamma &= (A_{23} + A_{32}), \\ c_1 &= 2i(A_{23} - A_{32}), & c_2 &= 2i(A_{31} - A_{13}), & c_3 &= 2i(A_{12} - A_{21}), \end{aligned} \quad (7.183)$$

we obtain

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\gamma_1 & \alpha - \omega_0 & \beta + \Delta' \\ \alpha + \omega_0 & -\gamma_2 & \gamma - \Delta \\ \beta - \Delta' & \gamma + \Delta & -\gamma_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}, \quad (7.184)$$

which is of the form (7.180), with  $\mathbf{c} = (c_1, c_2, c_3)$  and

$$M = M_H + M_D, \\ M_H = \begin{bmatrix} 0 & -\omega_0 & \Delta' \\ \omega_0 & 0 & -\Delta \\ -\Delta' & \Delta & 0 \end{bmatrix}, \quad M_D = \begin{bmatrix} -\gamma_1 & \alpha & \beta \\ \alpha & -\gamma_2 & \gamma \\ \beta & \gamma & -\gamma_3 \end{bmatrix}. \quad (7.185)$$

The matrix  $M_H$ , corresponding to the Hamiltonian part of the master equation (7.177), generates unitary evolution. The matrix  $M_D$ , corresponding to the dissipative part of this equation, is symmetric and can therefore be diagonalized. Its eigenvalues give the contraction rates of the Bloch sphere along the directions identified by the corresponding eigenvectors. Hence,  $M_D$  deforms the Bloch sphere into an ellipsoid. The term  $\mathbf{c}$  in (7.180) induces a rigid shift of the Bloch sphere. It is clear from (7.180) that the evolution of the Bloch vector over an infinitesimal time  $dt$  is given by

$$\mathbf{r}(t + dt) = (I + M dt) \mathbf{r}(t) + \mathbf{c} dt + O(dt^2). \quad (7.186)$$

This is an affine map. The evolution of the Bloch vector in a time  $t$  can be obtained by applying the infinitesimal evolution (7.186)  $\frac{t}{dt}$  times, in the limit  $dt \rightarrow 0$ . Since the composition of two affine maps is again an affine map, also the generic evolution of the Bloch vector in a finite time interval is an affine map. This conclusion allows us to obtain a precise correspondence between the 12 parameters appearing in the single-qubit master equation and the 12 parameters needed to characterize a generic quantum operation acting on a two-level system.

We now discuss a few special cases:

- (i)  $M_D = 0$ ,  $\mathbf{c} = \mathbf{0}$  (Hamiltonian case).

Equation (7.184) reduces to

$$\dot{x} = -\omega_0 y + \Delta' z, \quad \dot{y} = \omega_0 x - \Delta z, \quad \dot{z} = -\Delta' x + \Delta y. \quad (7.187)$$

The solution of these equations corresponds to a rotation of the Bloch sphere about the axis

$$\mathbf{n} = \left( \frac{\Delta}{\sqrt{\omega_0^2 + \Delta^2 + \Delta'^2}}, \frac{\Delta'}{\sqrt{\omega_0^2 + \Delta^2 + \Delta'^2}}, \frac{\omega_0}{\sqrt{\omega_0^2 + \Delta^2 + \Delta'^2}} \right), \quad (7.188)$$

with frequency  $\Omega = \sqrt{\Delta^2 + \Delta'^2 + \omega_0^2}$ .

**Exercise 7.18** Solve Eq. (7.187).

- (ii)  $M_D$  diagonal,  $M_H = 0$ ,  $\mathbf{c} = \mathbf{0}$ .

Equation (7.184) becomes

$$\dot{x} = -\gamma_1 x, \quad \dot{y} = -\gamma_2 y, \quad \dot{z} = -\gamma_3 z. \quad (7.189)$$

These equations are readily solved, and we obtain

$$x(t) = x(0)e^{-\gamma_1 t}, \quad y(t) = y(0)e^{-\gamma_2 t}, \quad z(t) = z(0)e^{-\gamma_3 t}. \quad (7.190)$$

Therefore, the Bloch sphere collapses exponentially fast onto its centre.

- (iii)  $M_D$  diagonal,  $M_H = 0$ ,  $\mathbf{c} \neq \mathbf{0}$ .

The differential equations that govern the evolution of the Bloch vector are given by

$$\dot{x} = -\gamma_1 x + c_1, \quad \dot{y} = -\gamma_2 y + c_2, \quad \dot{z} = -\gamma_3 z + c_3. \quad (7.191)$$

The solution is

$$\begin{aligned} x(t) &= x(0)e^{-\gamma_1 t} + \frac{c_1}{\gamma_1} (1 - e^{-\gamma_1 t}), \\ y(t) &= y(0)e^{-\gamma_2 t} + \frac{c_2}{\gamma_2} (1 - e^{-\gamma_2 t}), \\ z(t) &= z(0)e^{-\gamma_3 t} + \frac{c_3}{\gamma_3} (1 - e^{-\gamma_3 t}). \end{aligned} \quad (7.192)$$

As in the previous case, the Bloch sphere shrinks exponentially fast onto a single point. However, this point is no longer the centre of the Bloch sphere but has coordinates  $(c_1/\gamma_1, c_2/\gamma_2, c_3/\gamma_3)$ .

## 7.5 \* Non-Markovian quantum dynamics

Non-Markovian quantum dynamics is intuitively associated with a back-flow of information from the environment into the system. Memory effects are therefore associated to non-Markovianity: information which has been transferred to the environment, can be later retrieved by the system. In spite of such intuition, the two following main questions, namely (i) when is a quantum process Markovian or non-Markovian and (ii) how to quantify the degree of non-Markovianity of a process, are still under debate, see Rivas *et al.* (2014) and Breuer *et al.* (2016) for reviews.

---

*Trace distance, non-Markovianity, and information flow*


---

As proposed by Breuer *et al.* (2009), non-Markovian dynamics can be witnessed by means of the trace distance. The trace distance between two quantum states  $\rho_1$  and  $\rho_2$  is defined as

$$D(\rho_1, \rho_2) = \frac{1}{2} \text{Tr} |\rho_1 - \rho_2|, \quad (7.193)$$

where the modulus of an operator  $A$  is defined as  $|A| \equiv \sqrt{A^\dagger A}$ . The trace distance has a series of interesting properties (see, *e.g.*, Nielsen and Chuang, 2000). It is a measure of how close two quantum states are, since it is non-negative, symmetric and satisfies the triangle inequality. We have the bounds  $0 \leq D(\rho_1, \rho_2) \leq 1$ , where  $D(\rho_1, \rho_2) = 0$  if and only if  $\rho_1 = \rho_2$  and  $D(\rho_1, \rho_2) = 1$  if and only if  $\rho_1$  and  $\rho_2$  are orthogonal. Moreover, the trace distance is contractive under quantum operations, namely

$$D(\mathbb{S}(\rho_1), \mathbb{S}(\rho_2)) \leq D(\rho_1, \rho_2) \quad (7.194)$$

for any quantum operation  $\mathbb{S}$ . Hence a quantum operation can never increase the distinguishability of any two quantum states. For a Markovian (continuous time) dynamics, the evolution from  $\rho(t)$  to  $\rho(t+dt)$  is described by a quantum operation and therefore  $D(\rho_1(t), \rho_2(t))$  cannot increase with time,

$$D(\rho_1(t_2), \rho_2(t_2)) \leq D(\rho_1(t_1), \rho_2(t_1)), \quad t_2 \geq t_1, \quad (7.195)$$

for all states  $\rho_1$  and  $\rho_2$ . Any decrease of the trace distance with time can be interpreted as a loss of information from the system into the environment. Conversely, if there exists a couple of states  $\rho_1$  and  $\rho_2$  such that the trace distance is non-monotonic, we can conclude that the dynamics is non-Markovian, and when the trace distance increases with time, we say that information flows back from the environment into the system.

**Exercise 7.19** Compute the trace distance between two generic states of a qubit.

---

*Non-divisible quantum maps*


---

Rivas *et al.* (2010) proposed a definition of quantum Markovianity based on the notion of completely positive (CP)-divisible quantum maps. A quantum system subject for times  $t \geq t_0$  to an evolution governed by a family of trace-preserving linear maps  $\{\mathbb{S}(t_2, t_1), t_2 \geq t_1 \geq t_0\}$  is Markovian (or CP-divisible) if, for every  $t_2$  and  $t_1$ ,  $\mathbb{S}(t_2, t_1)$  is completely positive and  $\mathbb{S}$  fulfills the composition law

$$\mathbb{S}(t_3, t_1) = \mathbb{S}(t_3, t_2) \mathbb{S}(t_2, t_1), \quad t_3 \geq t_2 \geq t_1 \geq t_0. \quad (7.196)$$

The most general CP-divisible, time-homogeneous dynamics is provided by the GKLS master equation (7.169), see Rivas *et al.* (2014) for the extension to time-inhomogeneous dynamics (*i.e.*, time-dependent Hamiltonian and/or Lindblad operators).

A *collision model* of system-environment interaction, introduced by Scarani *et al.* (2002), can help visualize the memoryless property of CP-divisible dynamics. In

this model, the system interacts with the environment at discrete times  $t_1, t_2, \dots$ . Each collision produces a change in the system's state as follows:

$$\rho_S(t_{n+1}) = \text{Tr}_E[U(t_{n+1}, t_n)(\rho_S(t_n) \otimes \rho_E)U^\dagger(t_{n+1}, t_n)] = \mathbb{S}(t_{n+1}, t_n)[\rho_S(t_n)], \quad (7.197)$$

where  $U(t_{n+1}, t_n)$  is a unitary operator describing the system-environment interaction, and the environment state  $\rho_E$  is assumed to be the same for every collision. By tracing over the environment after the collision, we obtain the completely positive trace preserving map  $\mathbb{S}(t_{n+1}, t_n)$ . The concatenation of these collisions leads to a Markovian (CP-divisible) process. Indeed,

$$\rho_S(t_{n+2}) = \mathbb{S}(t_{n+2}, t_{n+1})[\rho_S(t_{n+1})] = \mathbb{S}(t_{n+2}, t_{n+1})\mathbb{S}(t_{n+1}, t_n)[\rho_S(t_n)], \quad (7.198)$$

and therefore

$$\mathbb{S}(t_{n+2}, t_n) = \mathbb{S}(t_{n+2}, t_{n+1})\mathbb{S}(t_{n+1}, t_n), \quad (7.199)$$

where the maps  $\mathbb{S}$  are completely positive trace preserving (CPT) maps. Note that every Markovian dynamics can be seen as a collision model (for continuous dynamics, the limit  $t_{n+1} - t_n \rightarrow 0$  has to be taken for every value of  $n$ ). Indeed, as we have seen in Sec. 7.1, it is possible to see any quantum operation for the evolution of a system S as the reduced dynamics of a unitary evolution for an enlarged system S+E.

The key point for a Markovian evolution is that, while  $U$  may depend on time, at each collision the environment is in the same state  $\rho_E$ , without any memory of the previous collision events.

Hereafter, we shall describe a few examples useful to grasp significant features of non-Markovian quantum dynamics.

### Quantum Brownian motion

Let us consider a model of the quantum Brownian motion where a heavy (Brownian) particle, described by an harmonic oscillator of frequency  $\omega_0$ , interacts with a *finite-size* “reservoir” made of a set of  $\mathcal{N}$  harmonic oscillators. The Hamiltonian of the model is the same as in Eqs. (7.125) and (7.126), but with a finite number of oscillators. The bilinear coupling (7.126) preserves the total number of quanta,  $N_T = N_0 + \sum_{j=1}^{\mathcal{N}} N_j$ , where  $N_0 = a^\dagger a$  and  $N_j = b_j^\dagger b_j$ . Following Gaioli *et al.* (1997), in Fig. 7.17 we plot the mean number of quanta for the Brownian particle,  $\langle N_0 \rangle$ , as a function of time. The reservoir is initially (at time  $t = 0$ ) prepared in a thermal state, at a given temperature  $T$ . It can be seen that  $\langle N_0 \rangle$ , after an initial decay, revives after a finite time, the *Poincaré recurrence time*, which diverges in the limit  $\mathcal{N} \rightarrow \infty$ . Therefore, in order to actually describe dissipation, one has to consider an infinite number of bath oscillators, corresponding to a continuous distribution of bath frequencies.

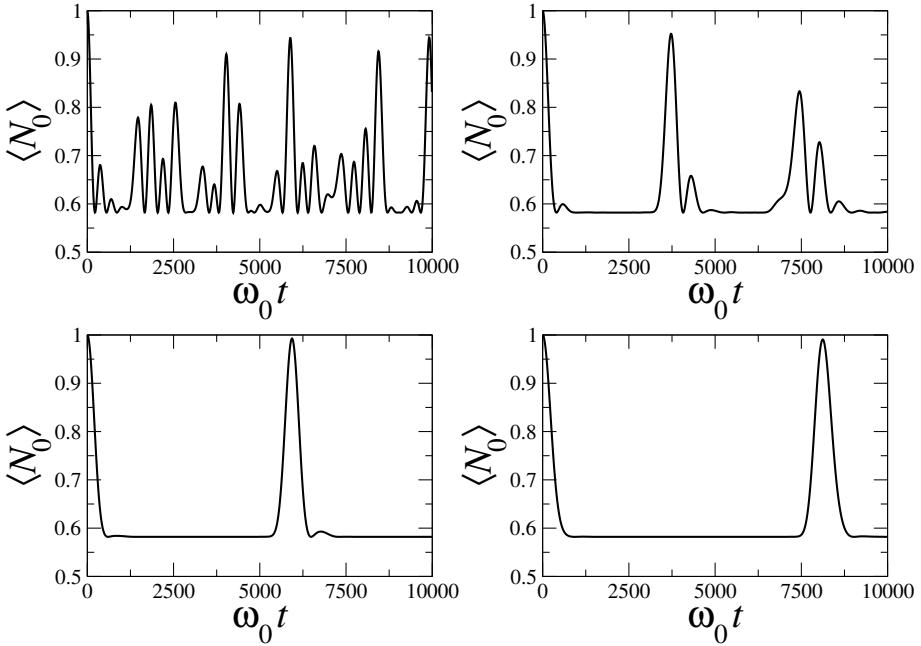


Fig. 7.17 Time evolution of the mean number of quanta,  $\langle N_0 \rangle$ , for a Brownian particle coupled to a bath of  $\mathcal{N}$  oscillators, for  $\mathcal{N} + 1 = 6$  (top left), 12 (top right), 18 (bottom left), and 24 (bottom right). The bath oscillators are initially in their equilibrium state at temperature  $T$  such that  $k_B T = \hbar\omega_0$ , where  $k_B$  is the Boltzmann constant and  $\omega_0$  the frequency of the Brownian particle. For the Brownian particle the initial number of quanta is  $\langle N_0 \rangle = 1$ . The bath frequencies  $\omega_j$  ( $j = 1, \dots, \mathcal{N}$ ) are equidistant around the frequency  $\omega_0$ , with a band width  $\omega_{\mathcal{N}} - \omega_1 = 0.018\omega_0$  for all  $\mathcal{N}$ . We note that the Poincaré recurrence time is very well approximated by  $2\pi/\delta\omega$ , where  $\delta\omega = \omega_{j+1} - \omega_j$  ( $j = 1, \dots, \mathcal{N} - 1$ ) is the spacing between nearby bath frequencies. For further details, see Gaioli *et al.* (1997).

### Entanglement revivals

Let us consider a bipartite two-qubit system AB, initially prepared in a maximally entangled state, say  $|\phi_{AB}^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . We assume that qubit A is exposed to a local bath, while qubit B is ideally isolated from decoherence sources. Since entanglement cannot increase under *local* CPT maps (a statement known as *monotonicity axiom*, see Bennett *et al.*, 1996b), it follows that for Markovian (CP-divisible) dynamics the entanglement will be monotonically decreasing. On the other hand, if the evolution is non-Markovian the request of nonmonotonicity does not hold and entanglement revivals are possible.

As a first example we assume that qubit A resonantly interacts with a quantum harmonic oscillator O via a Jaynes-Cummings Hamiltonian, which is obtained from the Rabi Hamiltonian (7.111) after neglecting the counter-rotating terms, an assumption frequently performed in cavity quantum electrodynamics, as we shall discuss in Sec. 10.1.2. Initially, the two qubits are prepared in the Bell state  $|\phi_{AB}^+\rangle$  and the oscillator in its ground state  $|0_O\rangle$ , see Fig. 7.18 (left). At time  $t = T/2$  ( $T = \pi/\lambda$ )

the states of  $A$  and  $O$  are swapped with respect the initial state, and the global state becomes  $|0_A\rangle \otimes |\phi_{OB}^+\rangle$ , see Fig. 7.18 (middle). We have  $\rho_{AB}(t = T/2) = |0_A\rangle \otimes \frac{1}{2}I_B$ , and therefore the AB entanglement is zero, being completely transferred to BO. The subsequent interaction between A and O can gradually restore the AB entanglement. The non-monotonic behaviour of the entanglement of formation  $E_F(\rho_{AB})$  is shown in Fig. 7.18 (right). At time  $T$ , when a new AO swapping is completed, maximal entanglement is retrieved. Therefore, the entanglement revival is here due to the perfect *entanglement back-transfer*.

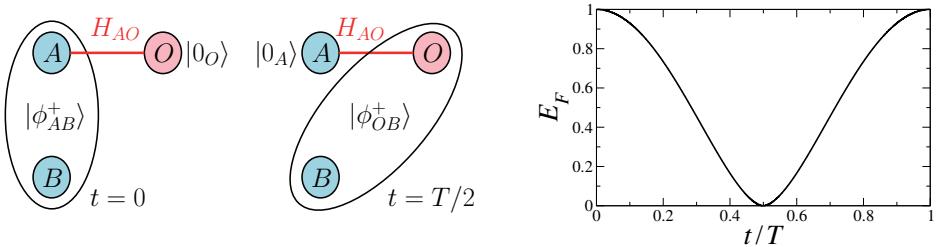


Fig. 7.18 Entanglement revival for a two-qubit system, with qubit A coupled to an harmonic oscillator O via a Jaynes-Cumming Hamiltonian.

A second example shows that entanglement revivals are possible also when the system is subject to *classical* noise sources (see D'Arrigo *et al.*, 2014), represented by a stochastic process. We consider a *random local field*, inducing local random unitaries  $U_\alpha(t) \otimes V_\beta(t)$  on a bipartite system, with the operators  $U_\alpha$  and  $V_\beta$  acting respectively on the first and on the second subsystem, and depending on the random variables  $\alpha, \beta$ . If the system is initially prepared in the pure state  $|\varphi(0)\rangle$ , then at any later time it is in the mixed state  $\rho(t) = \sum_{\alpha, \beta} p_{\alpha\beta} |\varphi_{\alpha\beta}(t)\rangle\langle\varphi_{\alpha\beta}(t)|$ , where  $|\varphi_{\alpha\beta}(t)\rangle = (U_\alpha(t) \otimes V_\beta(t))|\varphi(0)\rangle$  and the probabilities  $\{p_{\alpha\beta}\}$  satisfy the condition of unit total probability,  $\sum_{\alpha, \beta} p_{\alpha\beta} = 1$ .

Let us consider a two-qubit system  $AB$  initially prepared in the maximally entangled Bell state  $|\phi^+\rangle$ . The time evolution consists of local unitaries, but we have no complete information about which local unitary is acting (classical noise source). In particular, we suppose that the qubit A undergoes, with equal probability, a rotation about the  $x$ -axis of its Bloch sphere,  $U_x(t) = e^{-i\sigma_x \omega t/2}$ , or a rotation around the  $z$ -axis,  $U_z(t) = e^{-i\sigma_z \omega t/2}$ , while the qubit B remains unchanged.

On the other hand, the entanglement of the state  $\rho(t)$  changes in time, with period  $T = \frac{2\pi}{\omega}$ . At time  $t = T/2$ ,  $\rho(T/2) = \frac{1}{2}|\phi^-\rangle\langle\phi^-| + \frac{1}{2}|\psi^+\rangle\langle\psi^+|$  is separable, whereas at time  $T$ ,  $U_x(T) = U_z(T) = I_A$  and the initial maximally entangled state is recovered. In the interval  $[T/2, T]$  the entanglement revives from zero to one without the action of any nonlocal quantum operation, thus apparently violating the monotonicity axiom. On the other hand, there is no violation, since this axiom is fulfilled by all entanglement measures provided we consider local operations which are CPT maps. In the local random field example, entanglement recovery from

time  $T/2$  to time  $T$  is induced by purely local operations, which however cannot be described by a CPT map. To prove this point, it is enough to observe that the density matrix is such that  $\rho(T) = \rho(0)$ . Therefore driving the system from the state  $\rho(T/2)$  to  $\rho(T)$  is equivalent to driving it to  $\rho(0)$ . This operation cannot be described by a CPT map since the (CPT) evolution from time 0 to time  $T$  is not invertible.

In this latter example, it is easy to check that also a non-monotonous behaviour of the trace distance witnesses non-Markovian dynamics. For instance, if we consider the two initial states  $\rho_1(0) = |\phi^+\rangle\langle\phi^+|$  and  $\rho_2(0) = |\psi^+\rangle\langle\psi^+|$ , since the two states are orthogonal we have  $D(\rho_1(0), \rho_2(0)) = 1$ . At  $t = T/2$ ,  $\rho_1(T/2) = \rho_2(T/2) = \frac{1}{2}|\phi^-\rangle\langle\phi^-| + \frac{1}{2}|\psi^+\rangle\langle\psi^+|$  and therefore  $D(\rho_1(T/2), \rho_2(T/2)) = 0$ . On the other hand, the trace distance grows at later times, since  $\rho_1(T) = \rho_1(0)$  and  $\rho_2(T) = \rho_2(0)$ , implying  $D(\rho_1(T), \rho_2(T)) = 1$ .

## 7.6 Quantum to classical transition

### 7.6.1 Schrödinger's cat

The problem of the emergence of classical behaviour in a world governed by the laws of quantum mechanics has fascinated scientists since the dawn of quantum theory. The heart of the problem is the superposition principle, which entails consequences that appear unacceptable according to classical intuition. This point is clearly elucidated by Schrödinger's cat paradox. Inside a box we have a radioactive source, a detector, a hammer, a vial of poison and a cat. The source is a two-level atom, initially in its excited state  $|1\rangle$ . The atom can decay to the ground state  $|0\rangle$  by emission of a photon, which triggers the detector. The click of the detector induces the hammer to break the vial of poison and kill the cat. We assume that initially the state of the composite atom–cat system is

$$|\psi_0\rangle = |1\rangle|\text{live}\rangle. \quad (7.200)$$

Since the poison kills the cat if the atom decays to the state  $|0\rangle$ , we obtain, after a time corresponding to the half-life of the atom, the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|1\rangle|\text{live}\rangle + |0\rangle|\text{dead}\rangle), \quad (7.201)$$

which is a superposition of the live- and dead-cat states. We emphasize that the cat and the atom are now entangled. Let us consider the density matrix of the state (7.201):

$$\begin{aligned} \rho = |\psi\rangle\langle\psi| &= \frac{1}{2} \left( |1\rangle|\text{live}\rangle\langle 1|\langle\text{live}| + |0\rangle|\text{dead}\rangle\langle 0|\langle\text{dead}| \right. \\ &\quad \left. + |1\rangle|\text{live}\rangle\langle 0|\langle\text{dead}| + |0\rangle|\text{dead}\rangle\langle 1|\langle\text{live}| \right). \end{aligned} \quad (7.202)$$

In the basis  $\{|0\rangle|\text{live}\rangle, |0\rangle|\text{dead}\rangle, |1\rangle|\text{live}\rangle, |1\rangle|\text{dead}\rangle\}$  we have

$$\rho = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (7.203)$$

This density matrix contains non-zero matrix elements not only along the diagonal but also off-diagonal. These latter elements, known as *coherences*, have no classical analogue.

Decoherence plays a key role in understanding the transition from the quantum to classical world. The atom–cat system is never perfectly isolated from the environment, so that, instead of the state (7.201), we must consider the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left( |1\rangle |\text{live}\rangle |E_1\rangle + |0\rangle |\text{dead}\rangle |E_0\rangle \right), \quad (7.204)$$

where  $|E_0\rangle$  and  $|E_1\rangle$  are states of the environment. If  $|E_0\rangle$  and  $|E_1\rangle$  are orthogonal, then, after tracing over the environment, we obtain a diagonal density matrix:

$$\rho_{\text{dec}} = \frac{1}{2} \left( |1\rangle |\text{live}\rangle \langle 1| \langle \text{live}| + |0\rangle |\text{dead}\rangle \langle 0| \langle \text{dead}| \right). \quad (7.205)$$

This diagonal density matrix corresponds to a mixed state and is compatible with a classical description of the system in terms of probabilities. The cat is dead with probability  $p = 1/2$  and alive with the same probability, and we discover its state upon observation. Note that this situation is different from that described by (7.202). In that case, the atom–cat system is in a non-classical superposition state and only collapses onto a “classical” state (corresponding to the live or dead cat) after a measurement.

### 7.6.2 Decoherence and destruction of cat states

In this subsection, by means of a simple model, we shall show that a very weak interaction with the environment can lead to very fast coherence decay. These studies are of interest not only to understand the quantum to classical correspondence but also from the viewpoint of quantum computation. Since a quantum computer is never perfectly isolated from the external world, it is important to estimate the degree of isolation required to reliably implement a given quantum algorithm.

Let us recall that a free particle localized at  $x_0$  and moving along a line is described by the Gaussian wave function of Eq. (5.67), where the mean values of position and momentum are  $\langle x \rangle = x_0$  and  $\langle p \rangle = 0$ , and the variances are  $\langle (\Delta x)^2 \rangle = \delta^2/2$  and  $\langle (\Delta p)^2 \rangle = \hbar^2/(2\delta^2)$ . The corresponding density matrix is given by Eq. (5.71). We now consider the superposition of two Gaussian wave packets centred at  $+x_0$  and  $-x_0$ , respectively. We assume that the distance  $2x_0$  between these two packets is much larger than their width  $\delta$ . These states are known as *cat states*, for a reason that will soon become clear. If  $\psi_+(x)$  and  $\psi_-(x)$  denote the two Gaussian packets, we have

$$\begin{aligned} \psi_{\text{cat}}(x) &\equiv \frac{1}{\sqrt{2}} [\psi_+(x) + \psi_-(x)] \\ &= \frac{1}{\sqrt{2\sqrt{\pi}\delta}} \left\{ \exp\left[-\frac{(x-x_0)^2}{2\delta^2}\right] + \exp\left[-\frac{(x+x_0)^2}{2\delta^2}\right] \right\}. \end{aligned} \quad (7.206)$$

An example of probability distribution for a cat state is shown in Fig. 7.19. The corresponding density matrix, drawn in Fig. 7.20, has four components:

$$\begin{aligned}\langle x | \rho_{\text{cat}} | x' \rangle &= \langle x | \psi \rangle_{\text{cat}} \langle \psi | x' \rangle = \psi_{\text{cat}}(x) \psi_{\text{cat}}^*(x') = \psi_{\text{cat}}(x) \psi_{\text{cat}}(x') \\ &= \frac{1}{2} [\psi_+(x)\psi_+(x') + \psi_-(x)\psi_-(x') + \psi_+(x)\psi_-(x') + \psi_-(x)\psi_+(x')].\end{aligned}\quad (7.207)$$

The peaks along the diagonal ( $x = x'$ ) correspond to the two possible locations of the particle,  $x = x_0$  or  $x = -x_0$ . The off-diagonal peaks ( $x = -x'$ ) are purely quantum and demonstrate that the particle is neither localized in  $x_0$  nor in  $-x_0$ . We have a coherent superposition of the states  $\psi_+(x)$  and  $\psi_-(x)$ ; as in Schrödinger's cat paradox, we have a superposition of the states  $|1\rangle|\text{live}\rangle$  and  $|0\rangle|\text{dead}\rangle$ .

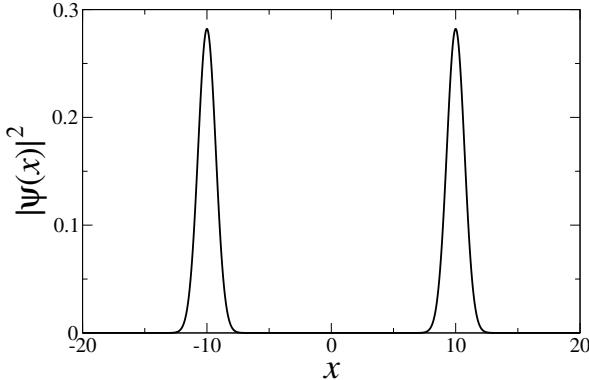


Fig. 7.19 The probability distribution for the cat state (7.206), with  $x_0 = 10\delta$ . Here and in the other figures of this subsection we set  $\delta = 1$ .

In order to investigate the physical origin of decoherence, it is useful to compare the Fourier transforms of a Gaussian and of a cat state. For the Gaussian wave packet (5.67) we have

$$|\langle p | \psi \rangle|^2 = \frac{\delta}{\sqrt{\pi\hbar}} \exp\left(-\frac{\delta^2 p^2}{\hbar^2}\right), \quad (7.208)$$

while for a cat state we obtain

$$|\langle p | \psi \rangle_{\text{cat}}|^2 = \frac{2\delta}{\sqrt{\pi\hbar}} \exp\left(-\frac{\delta^2 p^2}{\hbar^2}\right) \cos^2\left(\frac{px_0}{\hbar}\right). \quad (7.209)$$

Both (7.208) and (7.209) are shown in Fig. 7.21.

We emphasize the presence of interference fringes in the case of the cat state. These fringes are of pure quantum origin and are due to the coherent superposition of  $|\psi_+\rangle$  and  $|\psi_-\rangle$  in  $|\psi_{\text{cat}}\rangle$ . It is clear from (7.209) that these fringes have a period given by

$$\tilde{p} = \frac{\hbar}{x_0}. \quad (7.210)$$

The important point is that this period drops with increasing the separation  $2x_0$  between the two Gaussian packets. The interaction with the environment quickly

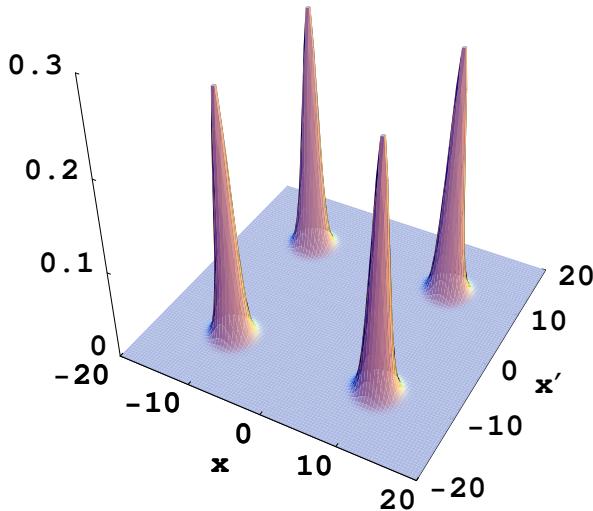


Fig. 7.20 The density matrix corresponding to the cat state (7.206), with  $x_0 = 10\delta$ .

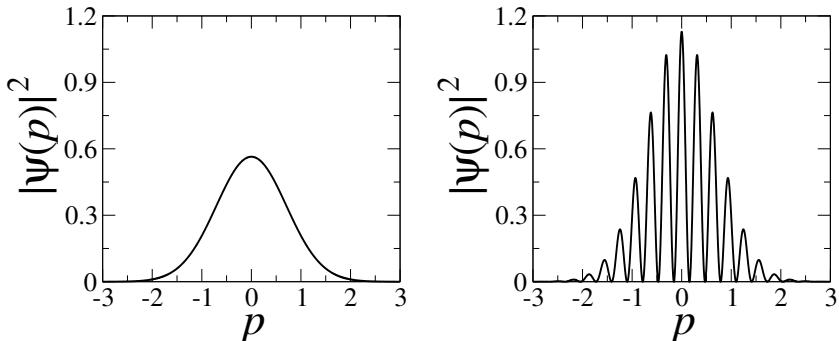


Fig. 7.21 The probability distributions  $|\psi(p)|^2$  of a Gaussian state (left) and a cat state (right), with  $x_0 = 10\delta$ . In this figure we set  $\delta = \hbar = 1$ .

weakens the visibility of the interference fringes. It is intuitive that this process is faster when the frequency of the fringes is higher; that is, when the separation of the packets is larger. Therefore, “non-local” quantum superpositions are very fragile. Table 7.1 gives the order of magnitude of  $\tilde{p}$  for a few relevant cases. Since a photon with wavelength 600 nm has a momentum of approximately  $10^{-27}$  Kg m/s, it is sufficient the collision of a such a photon to destroy quantum coherences in the cases shown in Table 7.1.

It is instructive to present a simple microscopic model illustrating the destruction of cat states (here we follow Cohen-Tannoudji, unpublished lecture notes). To simplify the discussion, we assume that the particle subjected to the decoherence process is heavy enough to neglect the variation of its kinetic energy over the decoherence time scale. We

Table 7.1 Relevant orders of magnitude for different cat states.

System	Mass	$x_0$	$\tilde{p}(\text{Kg m/s})$
atom	30 a.m.u.	600 nm	$10^{-27}$
dust particle	$10^{-9}$ g	$10^{-2}$ cm	$7 \times 10^{-30}$
cat	1 Kg	10 cm	$7 \times 10^{-33}$

model a system–environment interaction event as the scattering of a heavy particle (the system) with a light particle. We write the composite initial state of these two particles as follows:

$$\Phi_i(\mathbf{X}, \mathbf{x}) = \psi(\mathbf{X}) \otimes \phi_{\mathbf{p}_i}(\mathbf{x}), \quad (7.211)$$

where  $\mathbf{X}$  and  $\mathbf{x}$  denote the positions of the heavy and the light particles, moving in three-dimensional space, and  $\mathbf{p}_i$  is the initial momentum of the light particle. Let us describe the scattering of these two particles. For this purpose, it is convenient to write the Fourier expansion of the wave function  $\psi(X)$ :

$$\psi(\mathbf{X}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d\mathbf{P} \exp\left(\frac{i\mathbf{P} \cdot \mathbf{X}}{\hbar}\right) \tilde{\psi}(\mathbf{P}). \quad (7.212)$$

We insert (7.212) into (7.211) and obtain

$$\Phi_i(\mathbf{X}, \mathbf{x}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d\mathbf{P} \exp\left(\frac{i\mathbf{P} \cdot \mathbf{X}}{\hbar}\right) \tilde{\psi}(\mathbf{P}) \otimes \phi_{\mathbf{p}_i}(\mathbf{x}). \quad (7.213)$$

The scattering of the two particles changes momenta:  $\mathbf{P}_i = \mathbf{P} \rightarrow \mathbf{P}_f$  and  $\mathbf{p}_i \rightarrow \mathbf{p}_f$ . Momentum is conserved, so that

$$\mathbf{P}_i + \mathbf{p}_i = \mathbf{P}_f + \mathbf{p}_f. \quad (7.214)$$

Therefore, the effect of scattering on the wave function (7.213) is described by the following transformation:

$$\exp\left(\frac{i\mathbf{P} \cdot \mathbf{X}}{\hbar}\right) \otimes \phi_{\mathbf{p}_i}(\mathbf{x}) \rightarrow \exp\left(\frac{i(\mathbf{P} + \mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{X}}{\hbar}\right) \otimes \phi_{\mathbf{p}_f}(\mathbf{x}). \quad (7.215)$$

In order to obtain the final state of the composite system  $\Phi_f(\mathbf{X}, \mathbf{x})$ , we must integrate over all possible states after scattering. We have

$$\begin{aligned} \Phi_f(\mathbf{X}, \mathbf{x}) &= \frac{1}{(2\pi\hbar)^{3/2}} \int d\mathbf{P} \int d\mathbf{p}_f A(\mathbf{p}_f, \mathbf{p}_i) \exp\left(\frac{i\mathbf{P} \cdot \mathbf{X}}{\hbar}\right) \tilde{\psi}(\mathbf{P}) \exp\left(\frac{i(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{X}}{\hbar}\right) \otimes \phi_{\mathbf{p}_f}(\mathbf{x}) \\ &= \psi(\mathbf{X}) \int d\mathbf{p}_f A(\mathbf{p}_f, \mathbf{p}_i) \exp\left(\frac{i(\mathbf{p}_i - \mathbf{p}_f) \cdot \mathbf{X}}{\hbar}\right) \otimes \phi_{\mathbf{p}_f}(\mathbf{x}), \end{aligned} \quad (7.216)$$

where we have assumed that the transition amplitude  $A(\mathbf{p}_f, \mathbf{p}_i)$  is independent of the state of the heavy particle. It is important to stress that the collision has transformed the separable state (7.211) into an entangled state. In other words, the system is now entangled with the environment. The key point is that, while the global (system plus environment) final state  $\Phi_f(\mathbf{X}, \mathbf{x})$  is a pure state, this is not the case for the state of the

system alone. Therefore, it must be described by means of a density matrix, obtained after tracing over the environmental degrees of freedom:

$$\begin{aligned} \langle \mathbf{X}' | \rho_f | \mathbf{X}'' \rangle &= \int d\mathbf{x} \Phi_f(\mathbf{X}', \mathbf{x}) \Phi_f^*(\mathbf{X}'', \mathbf{x}) \\ &= \int d\mathbf{x} \psi(\mathbf{X}') \psi^*(\mathbf{X}'') \int d\mathbf{p}'_f A(\mathbf{p}'_f, \mathbf{p}_i) \exp\left(\frac{i(\mathbf{p}_i - \mathbf{p}'_f) \cdot \mathbf{X}'}{\hbar}\right) \otimes \phi_{\mathbf{p}'_f}(\mathbf{x}) \\ &\quad \times \int d\mathbf{p}''_f A^*(\mathbf{p}''_f, \mathbf{p}_i) \exp\left(-\frac{i(\mathbf{p}_i - \mathbf{p}''_f) \cdot \mathbf{X}''}{\hbar}\right) \otimes \phi_{\mathbf{p}''_f}^*(\mathbf{x}) \\ &= \langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle \int d\mathbf{p}_f |A(\mathbf{p}_f, \mathbf{p}_i)|^2 \exp\left(\frac{i(\mathbf{p}_i - \mathbf{p}_f) \cdot (\mathbf{X}' - \mathbf{X}'')}{\hbar}\right), \end{aligned} \quad (7.217)$$

where we have used the orthogonality relation

$$\int d\mathbf{x} \phi_{\mathbf{p}'_f}(\mathbf{x}) \phi_{\mathbf{p}''_f}^*(\mathbf{x}) = \delta(\mathbf{p}'_f - \mathbf{p}''_f) \quad (7.218)$$

and factored out the initial density matrix of the system:

$$\langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle = \psi(\mathbf{X}') \psi^*(\mathbf{X}''). \quad (7.219)$$

We must evaluate the integral appearing in the last line of (7.217). First of all, we assume that the collision is elastic, so that energy is conserved; that is,

$$\frac{P_i^2}{2M} + \frac{p_i^2}{2m} = \frac{P_f^2}{2M} + \frac{p_f^2}{2m}, \quad (7.220)$$

where  $M$  and  $m$  are the masses of the heavy and the light particle, respectively. It is reasonable to assume that the kinetic energy of the heavy particle is essentially unchanged ( $\frac{P_i^2}{2M} \approx \frac{P_f^2}{2M}$ ). Thus,  $\frac{p_i^2}{2m} \approx \frac{p_f^2}{2m}$  and so  $|\mathbf{p}_i| \approx |\mathbf{p}_f|$ . This implies that the integral in the last line of (7.217) averages to zero when  $|\mathbf{X}' - \mathbf{X}''| \gg \hbar/|\mathbf{p}_i|$ , as in this case its argument oscillates rapidly. Therefore, the matrix elements of  $\rho_f$  for  $|\mathbf{X}' - \mathbf{X}''| \gg \hbar/|\mathbf{p}_i|$  are eliminated after a single scattering event. This means that we may limit ourselves to the case  $|\mathbf{X}' - \mathbf{X}''| \ll \hbar/|\mathbf{p}_i|$ . In this limit we may expand the exponent appearing in (7.217) as follows:

$$\exp\left(\frac{i(\mathbf{p}_i - \mathbf{p}_f) \cdot (\mathbf{X}' - \mathbf{X}'')}{\hbar}\right) \approx 1 + \frac{i(\mathbf{p}_i - \mathbf{p}_f) \cdot (\mathbf{X}' - \mathbf{X}'')}{\hbar} - \frac{1}{2} \left[ \frac{(\mathbf{p}_i - \mathbf{p}_f) \cdot (\mathbf{X}' - \mathbf{X}'')}{\hbar} \right]^2. \quad (7.221)$$

We now insert this expression into (7.217). The first term gives

$$\langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle \int d\mathbf{p}_f |A(\mathbf{p}_f, \mathbf{p}_i)|^2 = \langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle. \quad (7.222)$$

We can transfer  $\langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle$  to the left-hand side of (7.217) to define the variation of the density matrix due to the collision as follows:

$$\delta \langle \mathbf{X}' | \rho | \mathbf{X}'' \rangle \equiv \langle \mathbf{X}' | \rho_f | \mathbf{X}'' \rangle - \langle \mathbf{X}' | \rho_i | \mathbf{X}'' \rangle. \quad (7.223)$$

Assuming invariance of  $A$  with respect to reflections, namely  $A(\mathbf{p}_f, \mathbf{p}_i) = A(-\mathbf{p}_f, -\mathbf{p}_i)$ , we see that the linear term in (7.221) does not contribute to the integral in (7.217). Finally,

we insert the second-order term of (7.221) into (7.217) and obtain

$$\begin{aligned}\delta\langle \mathbf{X}'|\rho|\mathbf{X}''\rangle &\approx -\frac{1}{2}\int d\mathbf{p}_f |A(\mathbf{p}_i, \mathbf{p}_f)|^2 \left[ \frac{(\mathbf{p}_i - \mathbf{p}_f)}{\hbar} \cdot (\mathbf{X}' - \mathbf{X}'') \right]^2 \langle \mathbf{X}'|\rho_i|\mathbf{X}''\rangle \\ &\propto |\mathbf{X}' - \mathbf{X}''|^2 \langle \mathbf{X}'|\rho_i|\mathbf{X}''\rangle.\end{aligned}\quad (7.224)$$

Therefore, given a cat state, the interaction with the environment drops coherences and leaves the diagonal terms of the density matrix practically unchanged, corresponding to “classical” probability distributions.

## 7.7 Decoherence and quantum measurements

The role of measurement is to convert quantum states into classical outcomes. In this section, we shall discuss the role of decoherence in the quantum measurement process. This issue is important in the problem of the transition from quantum physics to the classical world. Moreover, it is of interest in quantum information processing since any quantum protocol must end with a measurement.

We seek a purely unitary model of measurement that does not require the collapse of the wave packet. In this model, the first stage, known as *preambleasurement*, is to establish *correlations* between the system and the measurement apparatus. In order to clarify this point, we consider a simple example, sketched in Fig. 7.22. A one-qubit system, prepared in a generic pure state  $|\psi\rangle_S = \alpha|0\rangle_S + \beta|1\rangle_S$ , interacts with a measurement apparatus, initially in the state  $|0\rangle_A$ . We assume that this interaction induces a CNOT gate. Therefore, the preambleasurement process maps the initial state

$$|\Phi_0\rangle = |\psi\rangle_S |0\rangle_A = (\alpha|0\rangle_S + \beta|1\rangle_S) |0\rangle_A \quad (7.225)$$

into the state

$$|\Phi\rangle = \alpha|00\rangle_{SA} + \beta|11\rangle_{SA}, \quad (7.226)$$

corresponding to the density matrix

$$\rho_{SA} = \begin{bmatrix} |\alpha|^2 & 0 & 0 & \alpha\beta^* \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \alpha^*\beta & 0 & 0 & |\beta|^2 \end{bmatrix}. \quad (7.227)$$

We stress that this final state is entangled. This means that purely quantum correlations between the system and the measurement apparatus have been established. As a consequence, in the density matrix there are non-zero off-diagonal terms.

The presence of entanglement in the state (7.226) engenders ambiguity in the measurement process. The problem arises if we wish to associate the possible states of the apparatus with those appearing in (7.226),  $|0\rangle_A$  and  $|1\rangle_A$ . At first glance, it would be tempting to say that the measured quantity is  $\sigma_z$  and that the possible

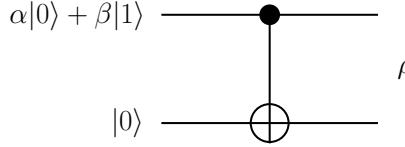


Fig. 7.22 A single-qubit premeasurement. The top line represents the system, the bottom line the measurement apparatus.

outcomes are  $\pm 1$ , corresponding to the states  $|0\rangle_A$  and  $|1\rangle_A$ . However, this interpretation is not correct. To illustrate this point, let us take  $\alpha = \beta = \frac{1}{\sqrt{2}}$  in (7.226). In this case we may write

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle), \quad (7.228)$$

where  $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$  are the eigenstates of  $\sigma_x$ . Therefore, the above interpretation would lead to an ambiguity in the definition of the measured quantity. Moreover, the direction of the information flow is not uniquely determined either. The action of the CNOT gate is

$$\text{CNOT}(|x\rangle_S |y\rangle_A) = |x\rangle_S |y \oplus x\rangle_A, \quad (7.229)$$

where  $x, y = 0, 1$ . Since the first qubit is unchanged, the direction of the information transfer is from the first qubit to the second qubit. Thus, it appears reasonable to identify the first (control) qubit with the system and the second (target) qubit with the measurement apparatus. However, this identification is not always correct since we have

$$\begin{aligned} \text{CNOT}(|\pm\rangle_S |+\rangle_A) &= |\pm\rangle_S |+\rangle_A, \\ \text{CNOT}(|\pm\rangle_S |-\rangle_A) &= |\mp\rangle_S |-\rangle_A. \end{aligned} \quad (7.230)$$

Thus, for the states of the  $\{|+\rangle, |-\rangle\}$  basis the second qubit is unchanged while the first qubit is flipped when the state of the second is  $|-\rangle$  (this is the backward sign propagation discussed in Sec. 3.5). This means that, while the information on the observable  $\sigma_z$  flows from the first to the second qubit, the information on  $\sigma_x$  travels from the second to the first qubit. If we require the information to always go from the system to the apparatus, then we must identify the first qubit with the system and the second qubit with the measurement apparatus in (7.229) and *vice versa* in (7.230).

The above ambiguities are resolved if we take into account the interaction of the measurement apparatus with the environment. For instance, we can represent the environment as a third qubit, interacting with the apparatus by means of a CNOT gate (see Fig. 7.23). The system-apparatus density matrix is obtained after tracing over the environmental degrees of freedom. In this example we obtain

$$\rho_{SA} = \begin{bmatrix} |\alpha|^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & |\beta|^2 \end{bmatrix}. \quad (7.231)$$

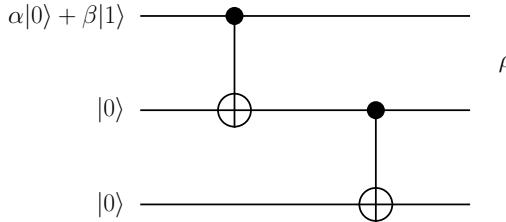


Fig. 7.23 A single-qubit premeasurement followed by a system–environment interaction process. The lines represent the system (top), the measurement apparatus (middle) and the environment (bottom).

In this density matrix, quantum correlations have disappeared, and we may interpret the density matrix (7.231) as follows. There exist *classical correlations* between the states  $|0\rangle_S$  and  $|0\rangle_A$  as well as between the states  $|1\rangle_S$  and  $|1\rangle_A$ : if the apparatus is in the state  $|0\rangle_A$ , we know that the system is in the state  $|0\rangle_S$ ; on the contrary, if the apparatus is in the state  $|1\rangle_A$ , the system is in the state  $|1\rangle_S$ . The state  $|0\rangle_A$  is obtained with probability  $p_0 = |\alpha|^2$ , the state  $|1\rangle_A$  with probability  $p_1 = |\beta|^2$ . Due to the perfect system-apparatus classical correlations, the average polarization for the system is  $\langle \sigma_z \rangle = p_0 - p_1 = |\alpha|^2 - |\beta|^2$ .

We emphasize that the density matrix (7.231) is diagonal in a *preferential basis* whose states (known as *pointer states*) are determined by the form of the apparatus–environment interaction. It is the CNOT interaction of the apparatus with the environment that determines the preferential basis in which the density matrix (7.231) is diagonal. A different apparatus–environment interaction would determine a different preferential basis. Therefore, according to the pointer-state theory, it is the interaction with the environment that determines which observable is measured by the apparatus (see Zurek, 2003).

### 7.7.1 \* Weak measurements

An interesting kind of generalized measurement is the weak measurement; that is, a measurement that disturbs the state of the system very little. In this section, we provide concrete examples of weak measurement, obtained from a projective measurement performed on an ancilla weakly coupled to the system. We assume that both the system and the ancilla can be described as qubits, and their interaction is a CNOT, with the system acting as the control and the ancilla as the target qubit. The system qubit is initially in a generic state  $\alpha|0\rangle_S + \beta|1\rangle_S$ , while the ancillary qubit is in the state  $\cos \frac{\theta}{2}|0\rangle_A + \sin \frac{\theta}{2}|1\rangle_A$ . After the CNOT interaction, the overall state reads as follows:

$$|\Phi\rangle = \alpha \cos \frac{\theta}{2} |00\rangle_{SA} + \alpha \sin \frac{\theta}{2} |01\rangle_{SA} + \beta \sin \frac{\theta}{2} |10\rangle_{SA} + \beta \cos \frac{\theta}{2} |11\rangle_{SA}. \quad (7.232)$$

The reduced density matrices for the system and the ancilla are readily computed:

$$\rho_S = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \sin\theta \\ \alpha^*\beta \sin\theta & |\beta|^2 \end{bmatrix}, \quad (7.233)$$

$$\rho_A = \begin{bmatrix} |\alpha|^2 \cos^2 \frac{\theta}{2} + |\beta|^2 \sin^2 \frac{\theta}{2} & \frac{1}{2} \sin\theta \\ \frac{1}{2} \sin\theta & |\alpha|^2 \sin^2 \frac{\theta}{2} + |\beta|^2 \cos^2 \frac{\theta}{2} \end{bmatrix}. \quad (7.234)$$

For  $\theta = 0$  we recover the result of Sec. 7.7, and the mean polarization on the ancilla coincides with that of the system:  $\langle \sigma_z \rangle_A = \langle \sigma_z \rangle_S = |\alpha|^2 - |\beta|^2$ . On the other hand, for  $\theta = \frac{\pi}{2}$ , the final state of the system coincides with the initial one,  $|\psi\rangle_S = \alpha|0\rangle_S + \beta|1\rangle_S$ , while the ancillary qubit is left in the state  $\rho_A = \frac{1}{2}I$ . Hence, the system is untouched but we do not obtain any information on its state. In general,

$$\langle \sigma_z \rangle_A = (|\alpha|^2 - |\beta|^2) \cos\theta \quad (7.235)$$

and therefore we can recover the system polarization as

$$\langle \sigma_z \rangle_S = \frac{\langle \sigma_z \rangle_A}{\cos\theta}. \quad (7.236)$$

Weak measurements correspond to  $\epsilon = \frac{\pi}{2} - \theta \ll 1$ . The great advantage in this case is that the system is weakly perturbed, and nevertheless we can recover the exact value of the polarization. Obviously this comes at a price: since  $\langle \sigma_z \rangle_A \ll \langle \sigma_z \rangle_S$  when  $\epsilon \ll 1$ , it is necessary to repeat the experiment many times (always with the system prepared in the same initial state) to obtain an accurate estimate of  $\langle \sigma_z \rangle_S$ .

It is instructive to consider also a different weak measurement model, following Brun (2002). We assume that the evolution of the two qubits is described by the unitary transformation  $\{[R_z(\theta)]_S \otimes I_A\}(\cos\theta I - i\sin\theta \text{CNOT})$ , with  $\theta \ll 1$ . After this, the two qubits are (up to an overall phase factor  $\exp(i\frac{\theta}{2})$ ) in the state

$$|\Psi\rangle = (\alpha|0\rangle_S + \beta \cos\theta|1\rangle_S)|0\rangle_A - i\beta \sin\theta|1\rangle_S|1\rangle_A. \quad (7.237)$$

If we measure the ancillary qubit in the  $z$ -basis, we obtain outcomes 0 or 1 with probabilities  $p_0 = |\alpha|^2 + |\beta|^2 \cos^2\theta \approx 1 - |\beta|^2\theta^2$  and  $p_1 = |\beta|^2 \sin^2\theta \approx |\beta|^2\theta^2 \ll 1$ . In both cases, the system and the ancilla are no longer entangled after the measurement. If the outcome 0 occurs, then the system is left in the state

$$|\psi_0\rangle_S = \frac{\alpha|0\rangle_S + \beta \cos\theta|1\rangle_S}{\sqrt{|\alpha|^2 + |\beta|^2 \cos^2\theta}} \approx \alpha\left(1 + \frac{1}{2}|\beta|^2\theta^2\right)|0\rangle_S + \beta\left(1 - \frac{1}{2}|\alpha|^2\theta^2\right)|1\rangle_S. \quad (7.238)$$

In this case, the system is weakly perturbed and the (weak) information obtained from the measurement of the ancilla is that it is now more probable that the system is found in the state  $|0\rangle_S$ . If instead the outcome 1 is obtained from this measurement, then the system is left in the state

$$|\psi_1\rangle_S = |1\rangle_S. \quad (7.239)$$

Therefore, in this example the measurement is weak in the sense that most of the time (with probability  $p_0 \approx 1 - |\beta|^2\theta^2$ ) the system is weakly perturbed. However, in rare occasions (with probability  $p_1 \approx |\beta|^2\theta^2$ ) the system changes abruptly (in the language of quantum trajectories, to be discussed in Sec. 7.7.2, we say that a *jump* occurs). We stress that the system–ancilla interaction plus the projective measurement acting on the ancilla can be conveniently described as a generalized measurement acting on the system, with the measurement operators

$$M_0 = |0\rangle_S \langle 0| + \cos \theta |1\rangle_S \langle 1|, \quad M_1 = \sin \theta |1\rangle_S \langle 1|, \quad (7.240)$$

satisfying the completeness relation  $M_0^\dagger M_0 + M_1^\dagger M_1 = I_S$ .

**Exercise 7.20** Show that, if the weak measurement (7.240) is repeated a very large number of times, then the effect is the same as a strong measurement: given a state  $|\psi\rangle_S = \alpha|0\rangle_S + \beta|1\rangle_S$ , the system is at the end left in the state  $|0\rangle_S$  with probability  $p_0 \approx |\alpha|^2$  or in the state  $|1\rangle_S$  with probability  $p_1 \approx |\beta|^2$ .

It is important to point out that the state of the system after the measurement of the ancillary qubit depends on the selected measurement basis. Let us consider, for instance, what happens if we measure the ancilla in the  $x$ -basis. For this purpose, it is useful to rewrite the state (7.237) as follows:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left( \alpha|0\rangle_S + \beta e^{-i\theta}|1\rangle_S \right) |+\rangle_A + \frac{1}{\sqrt{2}} \left( \alpha|0\rangle_S + \beta e^{i\theta}|1\rangle_S \right) |-\rangle_A, \quad (7.241)$$

where  $|\pm\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle_A \pm |1\rangle_A)$  are the eigenstates of  $(\sigma_x)_A$  corresponding to the eigenvalues  $\pm 1$ . The two measurement outcomes  $(\sigma_x)_A = \pm 1$  leave the system in the new states

$$|\psi_+\rangle_S = \alpha|0\rangle_S + \beta e^{-i\theta}|1\rangle_S, \quad |\psi_-\rangle_S = \alpha|0\rangle_S + \beta e^{i\theta}|1\rangle_S. \quad (7.242)$$

It can be clearly seen that in both cases the state of the system is weakly perturbed: a small relative phase  $\pm\theta$  is added. The sign of this phase is chosen randomly due to the inherent randomness of the quantum measurement process. If we repeat the entire procedure (system–ancilla interaction plus ancilla measurement) several times we do not have jumps but a slow *diffusion* in the relative phase  $\theta$  between the coefficients in front of the states  $|0\rangle_S$  and  $|1\rangle_S$ . Note that, also in this case in which the measurement is performed in the  $x$ -basis, we can give a convenient description in terms of generalized measurement, with the measurement operators

$$M_0 = \frac{1}{\sqrt{2}} \left( |0\rangle \langle 0| + e^{-i\theta} |1\rangle \langle 1| \right), \quad M_1 = \frac{1}{\sqrt{2}} \left( |0\rangle \langle 0| + e^{i\theta} |1\rangle \langle 1| \right). \quad (7.243)$$

Of course, this measurement weakly disturbs the state but also gives a small amount of information on it: we only know that, as a result of the weak measurement, a relative phase  $\theta$  has been added or subtracted.

### 7.7.2 \* Decoherence and quantum trajectories

In this section we shall describe the quantum-trajectory approach, a theory developed mainly in the field of quantum optics to investigate physical phenomena such as spontaneous emission, resonance fluorescence and Doppler cooling, to name but a few. Here we discuss quantum trajectories as a powerful technique for numerical simulation of quantum information processing in a noisy environment.

As a consequence of the undesired environmental coupling, a quantum processor becomes, in general, entangled with its environment. Therefore, under the assumption that the environment is Markovian, the state is described by a density matrix whose evolution is governed by a master equation. Solving this equation for a state of several qubits is a prohibitive task in terms of memory cost. Indeed, for a system whose Hilbert space has dimension  $N$ , one has to store and evolve a density matrix of size  $N \times N$ . Quantum trajectories allow us instead of doing so, to store only a stochastically evolving state vector of size  $N$ . By averaging over many runs we obtain the same probabilities (within statistical errors) as those obtained directly through the density matrix. Therefore, quantum trajectories are the natural approach for simulating equations otherwise very hard to solve.

The GKLS master equation (7.169) can be written as

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] - \frac{1}{2} \sum_{\mu} \{L_{\mu}^{\dagger} L_{\mu}, \rho\} + \sum_{\mu} L_{\mu} \rho L_{\mu}^{\dagger}, \quad (7.244)$$

where  $L_{\mu}$  are the Lindblad operators ( $\mu \in [1, \dots, \mathcal{M}]$ , the number  $\mathcal{M}$  depending on the noise model),  $H$  is the system Hamiltonian and  $\{, \}$  denotes the anticommutator. The first two terms of this equation can be regarded as the evolution generated by an effective non-Hermitian Hamiltonian,  $H_{\text{eff}} = H + iK$ , with  $K = -\frac{\hbar}{2} \sum_{\mu} L_{\mu}^{\dagger} L_{\mu}$ . In fact, we see that

$$-\frac{i}{\hbar}[H, \rho] - \frac{1}{2} \sum_{\mu} \{L_{\mu}^{\dagger} L_{\mu}, \rho\} = -\frac{i}{\hbar}(H_{\text{eff}}\rho - \rho H_{\text{eff}}^{\dagger}), \quad (7.245)$$

which reduces to the usual evolution equation for the density matrix in the case when  $H_{\text{eff}}$  is Hermitian. The last term in (7.244) is responsible for the so-called *quantum jumps*. In this context the Lindblad operators  $L_{\mu}$  are also called *quantum-jump operators*. If the initial density matrix is in a pure state  $\rho(t_0) = |\phi(t_0)\rangle\langle\phi(t_0)|$ , after an infinitesimal time  $dt$  it evolves to the following statistical mixture:

$$\begin{aligned} \rho(t_0 + dt) &= \rho(t_0) - \frac{i}{\hbar} \left[ H_{\text{eff}}\rho(t_0) - \rho(t_0)H_{\text{eff}}^{\dagger} \right] dt + \sum_{\mu} L_{\mu}\rho(t_0)L_{\mu}^{\dagger}dt \\ &\approx \left( I - \frac{i}{\hbar}H_{\text{eff}}dt \right) \rho(t_0) \left( I + \frac{i}{\hbar}H_{\text{eff}}^{\dagger}dt \right) + \sum_{\mu} L_{\mu}\rho(t_0)L_{\mu}^{\dagger}dt \\ &= \left( 1 - \sum_{\mu} dp_{\mu} \right) |\phi_0\rangle\langle\phi_0| + \sum_{\mu} dp_{\mu} |\phi_{\mu}\rangle\langle\phi_{\mu}|, \end{aligned} \quad (7.246)$$

with the probabilities  $dp_\mu$  defined by

$$dp_\mu = \langle \phi(t_0) | L_\mu^\dagger L_\mu | \phi(t_0) \rangle dt, \quad (7.247)$$

and the new states by

$$|\phi_0\rangle = \frac{(I - \frac{i}{\hbar} H_{\text{eff}} dt) |\phi(t_0)\rangle}{\sqrt{1 - \sum_\mu dp_\mu}} \quad (7.248)$$

and

$$|\phi_\mu\rangle = \frac{L_\mu |\phi(t_0)\rangle \sqrt{dt}}{\sqrt{dp_\mu}} = \frac{L_\mu |\phi(t_0)\rangle}{\|L_\mu |\phi(t_0)\rangle\|}. \quad (7.249)$$

The *quantum-jump picture* turns out then to be clear: a jump occurs and the system is prepared in the state  $|\phi_\mu\rangle$  with probability  $dp_\mu$ . With probability  $1 - \sum_\mu dp_\mu$  there are no jumps and the system evolves according to the effective Hamiltonian  $H_{\text{eff}}$  (normalization is also included in this case because the evolution is given by a non-unitary operator).

In order to simulate the master equation one may employ a numerical method usually known as the *Monte Carlo wave function* approach. We start from a pure state  $|\phi(t_0)\rangle$  and at intervals  $dt$ , much smaller than the time scales relevant for the evolution of the density matrix, we perform the following evaluation. We choose a random number  $\epsilon$  from a uniform distribution in the unit interval  $[0, 1]$ . If  $\epsilon \leq dp$ , where  $dp = \sum_\mu dp_\mu$ , the system jumps to one of the states  $|\phi_\mu\rangle$  (to  $|\phi_1\rangle$  if  $0 \leq \epsilon \leq dp_1$ , to  $|\phi_2\rangle$  if  $dp_1 < \epsilon \leq dp_1 + dp_2$ , and so on). On the other hand, if  $\epsilon > dp$ , evolution with the non-Hermitian Hamiltonian  $H_{\text{eff}}$  takes place, ending up in the state  $|\phi_0\rangle$ . In both circumstances we renormalize the state. We repeat this process as many times as  $n_{\text{steps}} = \bar{t}/dt$  where  $\bar{t}$  is the entire time elapsed during the evolution. Each realization provides a different *quantum trajectory* and a particular set of them (given a choice of the Lindblad operators) is an “unravelling” of the master equation.<sup>7</sup> It is easy to see that if we average over different runs, we recover the probabilities obtained with the density operator. In fact, given an operator  $A$ , we can write the mean value  $\langle A \rangle_t = \text{Tr}[A \rho(t)]$  as the average over  $\mathcal{N}$  trajectories:

$$\langle A \rangle_t = \lim_{\mathcal{N} \rightarrow \infty} \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \langle \phi_i(t) | A | \phi_i(t) \rangle. \quad (7.250)$$

The advantage of using the quantum-trajectory method is clear since we need to store a vector of length  $N$  ( $N$  being the dimension of the Hilbert space) rather than an  $N \times N$  density matrix. Moreover, there is also an advantage in computation time with respect to direct density-matrix calculations. It is indeed generally found that a reasonably small number of trajectories ( $\mathcal{N} \approx 100 - 500$ ) is needed in order

---

<sup>7</sup>Different unravellings are possible since there is always freedom in the choice of the Lindblad operators that induce a given temporal evolution of the density matrix  $\rho(t)$  (see, e.g., Brun, 2002). This corresponds to the freedom in the operator-sum representation discussed in Sec. 7.1.

to obtain a satisfactory statistical convergence, so that there is an advantage in computer time provided  $N > \mathcal{N}$ .<sup>8</sup>

We can say that a quantum trajectory represents a single member of an ensemble whose density operator satisfies the corresponding master equation (7.244). This picture can be formalized by means of the stochastic Schrödinger equation

$$|d\phi\rangle = -iH|\phi\rangle dt - \frac{1}{2} \sum_{\mu} \left( L_{\mu}^{\dagger} L_{\mu} - \langle \phi | L_{\mu}^{\dagger} L_{\mu} | \phi \rangle \right) |\phi\rangle dt + \sum_{\mu} \left( \frac{L_{\mu}}{\sqrt{\langle \phi | L_{\mu}^{\dagger} L_{\mu} | \phi \rangle}} - I \right) |\phi\rangle dN_{\mu}, \quad (7.251)$$

where the stochastic differential variables  $dN_{\mu}$  are statistically independent and represent measurement outcomes (for instance, in indirect measurement models the environment is measured, see the example from quantum optics below). Their ensemble average is given by  $M[dN_{\mu}] = \langle \phi | L_{\mu}^{\dagger} L_{\mu} | \phi \rangle dt$ . The probability that the variable  $dN_{\mu}$  is equal to 1 during a given time step  $dt$  is  $\langle \phi | L_{\mu}^{\dagger} L_{\mu} | \phi \rangle dt$ . Therefore, most of the time the variables  $dN_{\mu}$  are 0 and as a consequence the system evolves continuously by means of the non-Hermitian effective Hamiltonian  $H_{\text{eff}}$ . However, when a variable  $dN_{\mu}$  is equal to 1, the corresponding term in Eq. (7.251) is the most significant. In these cases a quantum jump occurs. Therefore, Eq. (7.251) is a stochastic non-linear differential equation, where the stochasticity is due to the measurement and non-linearity appears as a consequence of the renormalization of the state vector after each measurement process. We point out that, in contrast to the master equation (7.244) for the density operator, Eq. (7.251) represents the evolution of an individual quantum system, as exemplified by a single run of a laboratory experiment.

There is a close connection between the quantum-jump picture and the Kraus-operator formalism. To see this, we write the solution to the master equation (7.244) as a completely positive map:

$$\rho(t_0 + dt) = \mathbb{S}(t_0 + dt, t_0)\rho(t_0) = \sum_{\mu=0}^{\mathcal{M}} E_{\mu}(dt)\rho(t)E_{\mu}^{\dagger}(dt), \quad (7.252)$$

where, for  $\mu = 0$ , we have  $E_0 = I - iH_{\text{eff}}dt/\hbar$  and, for  $\mu > 0$ ,  $E_{\mu} = L_{\mu}\sqrt{dt}$  (see Sec. 7.4.2), satisfying  $\sum_{\mu=0}^{\mathcal{M}} E_{\mu}^{\dagger} E_{\mu} = I$  to first order in  $dt$ . The action of the superoperator  $\mathbb{S}$  in (7.252) can be interpreted as  $\rho$  being randomly replaced by  $E_{\mu}\rho E_{\mu}^{\dagger}/\text{Tr}(E_{\mu}\rho E_{\mu}^{\dagger})$ , with probability  $\text{Tr}(E_{\mu}\rho E_{\mu}^{\dagger})$ . Equivalently, the set  $\{E_{\mu}\}$  defines a Positive Operator-Valued Measurement (POVM), with POVM elements  $F_{\mu} = E_{\mu}^{\dagger} E_{\mu}$  satisfying  $\sum_{\mu=0}^{\mathcal{M}} F_{\mu} = I$ . The process outlined is equivalent to performing a continuous (weak) measurement on the system, which can be seen as an indirect measurement if the environment is actually measured.

<sup>8</sup>The updating of a density matrix and of a wave vector, performed after each time step  $dt$ , require  $O(N^3)$  and  $O(N^2)$  operations, respectively. In the first case, we must multiply  $N \times N$  matrices, in the latter  $N \times N$  matrices by a vector of size  $N$ . Hence, the cost in computer time for the quantum-trajectory approach is  $\propto \mathcal{N}N^2$ , to be compared with the cost  $\propto N^3$  for the density-matrix calculations.

### Example

A simple example will help us clarify the general quantum-trajectory theory sketched above. We consider the simplest, zero-temperature instance of the quantum optical master equation (7.170):

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] - \frac{\gamma}{2}(\sigma_+ \sigma_- \rho + \rho \sigma_+ \sigma_-) + \gamma \sigma_- \rho \sigma_+, \quad (7.253)$$

where the Hamiltonian  $H = -\frac{1}{2}\hbar\omega_0\sigma_z$  describes the free evolution of a two-level atom<sup>9</sup> and  $\gamma$  is the atom-field coupling constant. In this case there is a single Lindblad operator  $L_1 = \sqrt{\gamma}\sigma_-$  and a jump is a transition from the excited state  $|e\rangle$  to the ground state  $|g\rangle$  of the atom. Starting from a pure initial state  $|\phi(t_0)\rangle = \alpha|g\rangle + \beta|e\rangle$  and evolving it for an infinitesimal time  $dt$ , the probability of a jump in a time  $dt$  is given by

$$dp = \langle\phi(t_0)|L_1^\dagger L_1|\phi(t_0)\rangle dt = \gamma\langle\phi(t_0)|\sigma_+ \sigma_-|\phi(t_0)\rangle dt = \gamma p_e(t_0)dt, \quad (7.254)$$

where  $p_e(t_0) = |\beta|^2$  is the population of the excited state  $|e\rangle$  at time  $t_0$ . If a jump occurs, the new state of the atom is

$$|\phi_1\rangle = \frac{L_1|\phi(t_0)\rangle}{\|L_1|\phi(t_0)\rangle\|} = \frac{\sqrt{\gamma}\sigma_-(\alpha|g\rangle + \beta|e\rangle)\sqrt{dt}}{\sqrt{dp}} = |g\rangle. \quad (7.255)$$

In this case, the transition  $|e\rangle \rightarrow |g\rangle$  takes place and the emitted photon is detected. As a consequence, the atomic state vector collapses onto the ground state  $|g\rangle$ . If instead there are no jumps, the system evolution is governed by the non-Hermitian effective Hamiltonian  $H_{\text{eff}} = H - i\frac{\hbar}{2}L_1^\dagger L_1 = H - i\frac{\hbar}{2}\gamma\sigma_+ \sigma_-$ , so that the state of the atom at time  $t_0 + dt$  is

$$|\phi_0\rangle = \frac{(I - i\frac{\hbar}{\hbar}H_{\text{eff}}dt)|\phi(t_0)\rangle}{\sqrt{1-dp}} = \frac{(1 - i\frac{\omega_0}{2}dt)\alpha|g\rangle + (1 + i\frac{\omega_0}{2}dt - \frac{\gamma}{2}dt)\beta|e\rangle}{\sqrt{1-\gamma|\beta|^2dt}}. \quad (7.256)$$

Note that the normalization factor  $1/\sqrt{1-dp}$  is due to the fact that, if no counts are registered by the photodetector, then we consider it more probable that the system is unexcited. To illustrate the fact that the normalization factor leads to the correct physical result, let us consider the evolution without jumps in a finite time interval, from  $t_0$  to  $t_0 + t$ , and then let  $t \rightarrow \infty$ . If we first write the unnormalized state vector as

$$|\phi_0^{(u)}(t)\rangle = \alpha^{(u)}(t)|g\rangle + \beta^{(u)}(t)|e\rangle, \quad (7.257)$$

we see that the coefficients  $\alpha^{(u)}$  and  $\beta^{(u)}$  obey the simple equations of motion

$$\dot{\alpha}^{(u)}(t) = -i\frac{\omega_0}{2}\alpha^{(u)}(t), \quad \dot{\beta}^{(u)}(t) = \left[i\frac{\omega_0}{2} - \frac{\gamma}{2}\right]\beta^{(u)}(t), \quad (7.258)$$

which imply

$$\begin{aligned} \alpha^{(u)}(t_0 + t) &= \exp\left[-i\frac{\omega_0}{2}(t - t_0)\right]\alpha^{(u)}(t_0), \\ \beta^{(u)}(t_0 + t) &= \exp\left[\left(i\frac{\omega_0}{2} - \frac{\gamma}{2}\right)(t - t_0)\right]\beta^{(u)}(t_0). \end{aligned} \quad (7.259)$$

<sup>9</sup>The exact expression of the Hamiltonian  $H$  is not important here and one could equally well consider the same example for a generic Hamiltonian  $H$ , provided the interaction picture is considered (see, e.g., Scully and Zubairy, 1997).

Therefore, after normalization, the evolution of the state vector conditional on there being no photons detected is

$$|\phi_0(t_0 + t)\rangle = \frac{\alpha \exp[-i\frac{\omega_0}{2}(t - t_0)] |g\rangle + \beta \exp[(i\frac{\omega_0}{2} - \frac{\gamma}{2})(t - t_0)] |e\rangle}{\sqrt{|\alpha|^2 + |\beta|^2 \exp[-\gamma(t - t_0)]}}. \quad (7.260)$$

We stress that as  $t \rightarrow +\infty$  the state  $|\phi_0(t)\rangle \rightarrow |g\rangle$  (up to an overall phase factor). That is, if after some long time we have never seen a count, then we conclude that we have been in the ground state  $|g\rangle$  from the beginning.

**Exercise 7.21** Solve numerically the example of this section, using  $\mathcal{N}$  trajectories. In particular compute the time evolution of the Bloch-sphere coordinates for the two-level atom and compare, for different values of  $\mathcal{N}$ , with the exact solution.

## 7.8 A guide to the bibliography

A Review on decoherence is Zurek (2003), while a simple introduction can be found in Zurek (1991).

The Bloch-Fano representation is discussed in Fano (1957, 1983); Hioe and Eberly (1981); Schlienz and Mahler (1995) and, for quantum process tomography, in Benenti and Strini (2009b).

A discussion of the master equation from the perspective of quantum optics can be found in Gardiner and Zoller (2000) and also in Breuer and Petruccione (2002). References on dissipative quantum systems are Caldeira and Leggett (1983), Weiss (2012), Dittrich *et al.* (1998) and Prokof'ev and Stamp (2000). Reviews on characterization, quantification and detection of Non-Markovian quantum dynamics are Rivas *et al.* (2014) and Breuer *et al.* (2016). For a review of the dynamical Casimir effect and other quantum vacuum amplification phenomena see Nation *et al.* (2012).

The quantum-trajectory approach to quantum noise is discussed in Carmichael (1993), Gardiner and Zoller (2000), Scully and Zubairy (1997) and Plenio and Knight (1998). This approach can be generalized to treat non-Markovian effects, see for instance Breuer *et al.* (1999). An introduction to quantum trajectories closer to quantum information can be found in Brun (2002). The use of quantum trajectories in the simulation of quantum-information protocols is investigated, for instance, in Carlo *et al.* (2004).

## Chapter 8

# Quantum information theory

Classical information theory deals with the transmission of messages (say, binary strings) over communication channels. Its fundamental questions are: How much can a message be *compressed* and still be *transmitted reliably*? Can we protect this message against errors that will appear in noisy communication channels? In this chapter, we discuss the first questions in the light of quantum mechanics, which opens up new possibilities for information theory, while the second question will be postponed to Chap. 9, devoted to error correction.

We review the main results of classical information theory. It turns out that it is possible to compress a message into a shorter string of letters, the compression factor being the Shannon entropy. This is the content of Shannon's celebrated noiseless coding theorem. We discuss the natural extension of this result to quantum mechanics. To this end one may consider a message whose letters are quantum states, transmitted through a *quantum communication channel*. Such quantum states may be treated as though they were (quantum) information and one might thus ask to what extent this quantum message can be compressed. Schumacher's quantum noiseless coding theorem states that the optimal compression factor is given by the von Neumann entropy. Therefore, the von Neumann entropy is the appropriate measure of quantum information, just as the Shannon entropy is for classical information. On the other hand, if Alice codes a classical message by means of quantum states, it is natural to ask how much information Bob can gain on the message by performing (generalized) measurements on the quantum states received. This is not an easy question since the transmitted quantum states are not necessarily orthogonal and they cannot therefore be perfectly distinguished. The Holevo bound establishes an upper limit on the information accessible to Bob.

The performance of a noisy classical channel can be characterized by a single number, i.e., its *capacity*, defined as the maximum rate at which information can be reliably transmitted down the channel. On the other hand, noisy quantum communication channels can use quantum systems as carriers of both classical or quantum information, by encoding classical bits by means of quantum states or by transferring (unknown) quantum states between, say, subunits of a quantum computer. Therefore, different capacities must be defined. The *classical capacity*

$C$  and the *quantum capacity*  $Q$  of a noisy quantum channel are defined as the maximum number of, respectively, bits and qubits that can be reliably transmitted per channel use.

Noise effects can be conveniently described in the quantum operation formalism described in Sec. 7.1: any input state  $\rho$  is mapped onto the output state  $\rho' = \mathbb{S}(\rho)$  by a linear, completely positive, trace preserving map (superoperator)  $\mathbb{S}$ . In the simplest setting each channel use is independent of the previous ones. It means that, if a *quantum channel* use is described by the map  $\mathbb{S}$ ,  $n$  uses of the channel are described by the map  $\mathbb{S}_n = \mathbb{S}^{\otimes n}$ . This assumption is not always justified. Indeed, noise can have significant low frequency components, which traduce themselves in *memory* effects, leading to relevant correlations in the errors affecting successive transmissions. Important examples in this context are photons traveling across fibers with birefringence fluctuating with characteristic time scales longer than the separation between consecutive light pulses or low-frequency impurity noise in solid state implementations of quantum hardware. Memory effects become unavoidably relevant when trying to increase the *transmission rate*, that is, to reduce the time interval that separates two consecutive channel uses. With increasing the transmission rate, the environment may retain memory of the previous channel uses. In this case noise introduces memory (or *correlation*) effects among consecutive channel uses, and  $\mathbb{S}_n \neq \mathbb{S}^{\otimes n}$  (*memory channels*).

The present chapter requires more formal development than those preceding. This is quite natural since we are concerned with the most general results on the properties of quantum information. Nonetheless, in order to illustrate these general concepts and aiming to provide a gentle introduction to the field, we shall skip proofs of theorems and rather describe significant concrete examples in detail.

## 8.1 Classical data compression

### 8.1.1 Shannon's noiseless coding theorem

We show that the Shannon entropy is a good *measure* of information. Let us consider the following fundamental problem: how much can a message be *compressed* while still obtaining essentially the same information? In other words, what are the minimal physical resources required in order to store a message without loosing its information content?

As an example, we consider a message written using an alphabet with four letters,  $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ . We assume that these letters occur with probabilities  $p_1 = \frac{1}{2}$ ,  $p_2 = \frac{1}{4}$ ,  $p_3 = p_4 = \frac{1}{8}$ . To specify a letter out of four we need 2 bits of information. It is instead more convenient to encode the letters as follows:

$$a_1 \rightarrow c_1 \equiv 0, \quad a_2 \rightarrow c_2 \equiv 10, \quad a_3 \rightarrow c_3 \equiv 110, \quad a_4 \rightarrow c_4 \equiv 111. \quad (8.1)$$

To send one coded letter we need, on average,  $\sum_{i=1}^4 p_i l_i$  bits, where  $l_i$  is the length, in bits, of the coded letter  $c_i$  (we have  $l_1 = 1$ ,  $l_2 = 2$ ,  $l_3 = l_4 = 3$ ). Since  $\sum_i p_i l_i = \frac{7}{4} < 2$ , we have compressed the information. Note that the good strategy,

here as in any other useful compression code, is to encode the most probable strings in the shortest sequences and the less probable strings in the longest sequences.

Shannon proved that the optimal compression rate is given by the Shannon entropy. If Alice sends Bob a string of  $n$  letters taken from the alphabet  $\mathcal{A} = \{a_1, \dots, a_k\}$  and each letter  $a_i$  occurs with the *a priori* probability  $p_i$ , then, for large  $n$ , Alice can reliably communicate her message by sending only  $nH(p_1, \dots, p_k)$  bits of information. This is the content of the Shannon's noiseless coding theorem.

**Theorem 8.1** Shannon's noiseless coding theorem: *Given a message in which the letters have been chosen independently from the ensemble  $\mathcal{A} = \{a_1, \dots, a_k\}$  with a priori probabilities  $\{p_1, \dots, p_k\}$ , there exists, asymptotically in the length of the message, an optimal and reliable code compressing the message to  $H(p_1, \dots, p_k)$  bits per letter.*

Note that in the example considered above the optimal compression rate is  $H = -\sum_{i=1}^4 p_i \log p_i = \frac{7}{4}$ . Since  $\sum_i p_i l_i = \frac{7}{4} = H$ , the optimal compression established by the Shannon's theorem has been attained.

A proof of Shannon's theorem can be found in Cover and Thomas (1991). Here, we shall limit ourselves to explaining the basic argument. First of all, it is useful to introduce the concept of *typical sequence*. A particular  $n$ -letter message,  $x_1, x_2, \dots, x_n$ , where  $x_i \in \mathcal{A}$ , occurs with *a priori* probability

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2)\cdots p(x_n), \quad (8.2)$$

where we have assumed that the different letters of the message are independent and identically distributed according to the probability distribution  $\{p_1, p_2, \dots, p_k\}$ . A typical sequence contains approximately  $np_1$  times the letter  $a_1$ ,  $np_2$  times the letter  $a_2$ , ... and  $np_k$  times the letter  $a_k$ . The number of such strings is given by  $n!/\prod_{i=1}^k (np_i)!$ , which represents the number of distinct strings having  $np_1$  times  $a_1$ ,  $np_2$  times  $a_2$  and so on. It is easy to show (see exercise 8.1) that

$$\frac{n!}{\prod_{i=1}^k (np_i)!} \approx 2^{nH(p_1, \dots, p_k)}. \quad (8.3)$$

**Exercise 8.1** Using Stirling's formula,  $\log n! = n \log n - n/\ln 2 + O(\log n)$ , where  $\ln$  denotes the natural logarithm (having base  $e$ ), prove Eq. (8.3).

The probability of obtaining any given typical sequence  $x_1, x_2, \dots, x_n$  is

$$p(x_1, x_2, \dots, x_n) \approx 2^{-nH(p_1, \dots, p_k)}. \quad (8.4)$$

Indeed, from Eq. (8.2) we obtain

$$-\frac{1}{n} \log p(x_1, \dots, x_n) = -\frac{1}{n} \sum_{i=1}^n \log p(x_i) \approx H(p_1, \dots, p_k), \quad (8.5)$$

where the last (approximate) equality is guaranteed by the law of large numbers and is obtained as follows: for all  $j = 1, 2, \dots, k$ , the frequency  $n_j/n$  of the letter  $a_j$  in the message is substituted by the *a priori* probability  $p_j$  ( $n_j$  is the number of

times that  $a_j$  appears in the message). The law of large numbers also tells us that, if we fix  $\epsilon > 0$  and we say that a sequence is  $\epsilon$ -typical when

$$\left| -\frac{1}{n} \log p(x_1, \dots, x_n) - H(p_1, \dots, p_k) \right| < \epsilon, \quad (8.6)$$

then, for any  $\delta > 0$ , the probability that a given sequence is  $\epsilon$ -typical is larger than  $1 - \delta$ , for sufficiently large  $n$ . Therefore, most of the sequences are  $\epsilon$ -typical in the limit of large  $n$ .

Since there are  $2^{nH}$  typical sequences (asymptotically in  $n$ ), each occurring with probability  $2^{-nH}$ , we can identify which one of these sequences actually occurred using  $nH$  bits. Moreover, it can be shown that this asymptotic compression to  $H$  bits per letter is optimal. Note that it is sufficient to code only the typical sequences since the probability that a message is atypical becomes negligible for large  $n$ .

### 8.1.2 Examples of data compression

It is clear that an “asymptotic” data compression strategy, that is, a strategy based on the compression of long typical sequences, is not practical: to compress a long  $n$ -letter message, we must accumulate all  $n$  letters before identifying the typical sequence and compressing it. Fortunately, there exist quite efficient methods to encode smaller strings of letters.

A first example was shown in Sec. 8.1.1. Here we consider further examples. First of all, we apply the encoding (8.1) to a four-letter alphabet, with  $p_1 = 0.9$ ,  $p_2 = 0.05$ ,  $p_3 = p_4 = 0.025$ . The optimal compression is determined by  $H(p_1, p_2, p_3, p_4) \approx 0.62$ , while the code gives  $\sum p_i l_i = 1.15$  and therefore data compression in this case, even though useful, is not optimal.

Let us apply the same code to the case in which the four letters are equiprobable,  $p_i = \frac{1}{4}$  for  $i = 1, \dots, 4$ . In this case, no compression is possible, because  $H = 2$  and we send exactly two bits to specify a letter. Furthermore, if we try to apply the previous code, we obtain  $\sum p_i l_i = 2.25 > 2$  and therefore the code is in this case detrimental to the efficiency of data transmission.

Finally, let us consider the Huffman code, shown in Table 8.1. We consider a binary alphabet  $\{0, 1\}$  and the encoding procedure is applied to strings four bits long. There are  $2^4 = 16$  such strings ( $0 \equiv 0000$ ,  $1 \equiv 0001$ ,  $\dots$ ,  $15 \equiv 1111$ ). Let  $P_i$  denote the probability that the string  $i$  occurs, with  $i = 0, \dots, 15$ . We have  $P_0 = p_0^4$ ,  $P_1 = p_0^3 p_1$ ,  $\dots$ ,  $P_{15} = p_1^4$ . If we consider, for instance, the case with  $p_0 = \frac{3}{4}$  and  $p_1 = \frac{1}{4}$ , we find that the best possible compression for a four-letter message is given by  $4H(p_0, p_1) \approx 3.25$ , while the Huffman code gives on average  $\sum_{i=0}^{15} P_i l_i \approx 3.27$  bits, which is very close to the optimal value. This shows the power of data compression codes.

The enormous practical importance of data compression in fields such as telecommunication is self-evident. Data compression allows us to increase the *transmission rate* or the *storage capacity* of a computer. To achieve such results, we simply exploit the *redundancies* that any message contains: for instance, the letters of an

(English) text are not equiprobable but appear with different frequencies. Shannon's theorem tells us that, as far as the letters of a message are not equiprobable, data compression is possible.<sup>1</sup>

Table 8.1 Data encoding by means of the Huffman code, with  $p_0 = \frac{3}{4}$ ,  $p_1 = \frac{1}{4}$ .

Message	Huffman's encoding
0000	10
0001	000
0010	001
0011	11000
0100	010
0101	11001
0110	11010
0111	1111000
1000	011
1001	11011
1010	11100
1011	111111
1100	11101
1101	111110
1110	111101
1111	1111001

### 8.1.3 Capacity of classical channels

In a communication system a message is encoded by a sender (Alice) in a string of bits, which are sent down a *classical* communication channel to a receiver (Bob) which should be able to unambiguously decode the message. We call the channel classical in that classical systems are sent across it. In general the channel is *noisy*, that is, some error is introduced in the communication process. Nevertheless we expect that it is still possible to reliably communicate through a noisy channel, provided noise is not too strong. For example, Alice could send many times each bit of information and Bob could apply the majority voting, namely accept as correct the value received most of the times. This would be a reasonable strategy if the possibility that the channel corrupts a bit of information is small.

A key question in information theory is what is the channel *capacity*, namely the highest rate at which information can be reliably transmitted through the channel, where reliably means that the error rate can be made arbitrarily small in the limit  $n \rightarrow \infty$ , where  $n$  is the number of channel uses. The input message is characterized by the random variable  $X$ , which takes the value  $x \in \mathcal{A} = \{a_1, \dots, a_k\}$ , with probability  $p(x) \in \{p_1, \dots, p_k\}$ , and similarly the output message is characterized by the random variable  $Y$ , which takes the value  $y \in \mathcal{A}' = \{a'_1, \dots, a'_l\}$ , with probability

---

<sup>1</sup>The notion of differing probabilities should not be confused with correlations, which are present in a real language but are not being considered here.

$p(y) \in \{p'_1, \dots, p'_l\}$ . The channel is characterized by the transfer probability  $p(y|x)$ , the joint probability of the input and output letters being  $p(x,y) = p(x)p(y|x)$ . We assume here that the channel is *stationary*, namely  $p(x)$  and  $p(y)$  are the same for all channel uses, and that the channel is *memoryless*, that is, the action of the channel on each bit is independent of the past history. The capacity of a classical memoryless channel is given by Shannon's noisy channel coding theorem.

**Theorem 8.2** Shannon's noisy channel coding theorem: The capacity  $C$  of a classical communication channel, namely the highest number of bits that can be reliably transmitted per channel use, is given by

$$C = \max_{\{p(x)\}} \mathcal{I}(X:Y), \quad (8.7)$$

where reliable means that the error probability can be made arbitrarily small in the limit of channel uses  $n \rightarrow \infty$ .

The mutual information  $\mathcal{I}(X:Y)$  indicates the number of bits of the input message that can be faithfully decoded per output letter. In the limiting case of a noiseless channel,  $X$  and  $Y$  are perfectly correlated,  $p(y|x) = \delta_{xy}$  and therefore  $H(X) = H(Y) = H(X,Y)$ , implying  $\mathcal{I}(X:Y) = H(X)$ , so that the input information can be perfectly recovered from the output. In the opposite limiting case, the channel noise is so strong to make  $X$  and  $Y$  independent, so that  $\mathcal{I}(X:Y) = 0$  and the capacity of the channel to transmit information vanishes,  $C = 0$ .

A proof of Shannon's noisy channel coding theorem can be found in Cover and Thomas (1991). Here, we limit ourselves to illustrative examples and qualitative arguments. Let us first consider the *binary symmetric channel*. It flips each input bit with probability  $p$ , that is  $p(1|0) = p(0|1) = p$ , while the bit is transmitted without error with probability  $1-p$ :  $p(0|0) = p(1|1) = 1-p$ . If we encode  $k$  bits in blocks of size  $n$  channel uses, then the channel rate is  $R = \frac{k}{n}$  (we have  $2^n$  possible strings of  $n$  bits used to encode  $2^k = 2^{nR}$  codewords), and we want to maximize  $R$ , with negligible error rate in the limit  $n \rightarrow \infty$ . It is natural to require that the input typical sequences are sufficiently far apart to consider highly improbable that the channel maps two different inputs into the same output, otherwise we would have decoding ambiguities. The *Hamming distance* between two strings  $i$  and  $j$  of  $n$  bits is the number of binary digits in which  $i$  and  $j$  differ. For any  $n$ -bit input string, due to channel noise about  $np$  digits flip, so that about  $2^{nH_{\text{bin}}(p)}$  typical output strings are possible ( $H_{\text{bin}}(p) = -p \log p - (1-p) \log(1-p)$  is the binary entropy), occupying a “Hamming sphere” of “Hamming radius”  $np$ . To avoid decoding ambiguities, the output Hamming spheres should have negligible overlap. Since there are  $2^n$  possible output messages, this happens when

$$2^{nH_{\text{bin}}(p)} 2^{nR} \leq 2^n, \quad (8.8)$$

namely

$$R \leq 1 - H_{\text{bin}}(p). \quad (8.9)$$

As shown in exercise 8.2, this upper bound to the rate  $R$  is the capacity of the binary symmetric channel.

The above arguments can be extended to generic alphabets and channels. For each typical  $n$ -letters input sequence, there are approximately  $2^{nH(Y|X)}$  typical output sequences. In order to avoid decoding errors, we must avoid that two typical input sequences produce the same output sequence. The total number of typical output sequences is of the order of  $2^{nH(Y)}$ . By dividing this set into subsets (Hamming spheres) of size  $2^{nH(Y|X)}$ , corresponding to the different input sequences, we find that the total number of disjoint subsets cannot be larger than  $2^{n(H(Y)-H(Y|X))} = 2^{n\mathcal{I}(X:Y)}$ . We can therefore send at most  $2^{n\mathcal{I}(X:Y)}$  distinguishable sequences of length  $n$  and therefore the channel capacity is upper bounded by the mutual information  $\mathcal{I}(X:Y)$ . It turns out that this upper bound is actually attainable, with arbitrarily low error rates asymptotically in  $n$ , for a proof see Cover and Thomas (1991).

**Exercise 8.2** Using the Shannon's noisy channel coding theorem, show that the capacity of the binary symmetric channel is given by  $C = 1 - H_{\text{bin}}(p)$ .

**Exercise 8.3** In the *binary erasure channel* some bits are lost (rather than corrupted as in the binary symmetric channel), with probability  $p$ . This is equivalent to stating that there is a third output state, which we call  $e$ , with  $p(e|0) = p(e|1) = p$ , while  $p(0|0) = p(1|1) = 1 - p$  and  $p(1|0) = p(0|1) = 0$ . Compute the capacity of this channel and compare the result with the one obtained in exercise 8.2 for the binary symmetric channel.

## 8.2 Quantum data compression

Quantum information is the information related to a *quantum source*, that is, a source  $\Sigma$  of identical quantum systems  $\mathcal{Q}$ , which are prepared in an unknown quantum state chosen from the ensemble  $\{\rho_i\}$ , according to a given stationary probability distribution  $\{p_i\}$ .

### 8.2.1 Schumacher's quantum noiseless coding theorem

The Schumacher's quantum noiseless coding theorem is an extension to the quantum case of the Shannon's noiseless coding theorem discussed in Sec. 8.1. Alice sends Bob a message of  $n$  letters, each letter being chosen at random from the alphabet (ensemble of pure states)  $\mathcal{A} = \{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_k\rangle\}$ .<sup>2</sup> The state  $|\psi_i\rangle$  is extracted with *a priori* probability  $p_i$  and  $\sum_i p_i = 1$ . Therefore, each letter in the message is described by the density matrix

$$\rho = \sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|, \quad (8.10)$$

---

<sup>2</sup>The problem of quantum data compression for ensembles of mixed states is more complicated, see Wilde (2013).

and the density matrix for the entire message is

$$\rho^n = \rho^{\otimes n}, \quad (8.11)$$

where  $\rho^{\otimes n}$  denotes the tensor product  $\rho \otimes \rho \otimes \cdots \otimes \rho$ . It is clear that we have assumed that all the letters in the message are statistically independent and described by the same density matrix  $\rho$ . Schumacher's theorem tells us that it is possible to compress the message, namely to encode it in a shorter message, the optimal compression rate being the von Neumann entropy.

**Theorem 8.3** Schumacher's quantum noiseless coding theorem: *Given a message whose letters are pure quantum states drawn independently from the ensemble  $\mathcal{A} = \{|\psi_1\rangle, \dots, |\psi_k\rangle\}$  with a priori probabilities  $\{p_1, \dots, p_k\}$ , there exists, asymptotically in the length of the message, an optimal and reliable code compressing the message to  $S(\rho)$  qubits per letter, where  $\rho = \sum_{i=1}^k p_i |\psi_i\rangle\langle\psi_i|$ .*

The proof of this theorem can be found in Schumacher (1995) and closely follows the techniques used in the proof of the Shannon's noiseless coding theorem. Here, we simply illustrate the basic ideas of the proof. Let us first write the spectral decomposition of the density operator  $\rho$ :

$$\rho = \sum_{i=1}^k \lambda_i |a_i\rangle\langle a_i|. \quad (8.12)$$

Clearly, we have  $H(\lambda_1, \dots, \lambda_k) = S(\rho)$ . The ensemble  $\mathcal{A}' = \{|a_1\rangle, \dots, |a_k\rangle\}$  constitutes an alphabet of orthogonal pure quantum states. Following the definition of  $\epsilon$ -typical sequence given Sec. 8.1.1, we say that a state  $|x_1\rangle \otimes \cdots \otimes |x_n\rangle$ , with  $|x_i\rangle \in \mathcal{A}'$ , is  $\epsilon$ -typical when

$$\left| -\frac{1}{n} \log [\lambda(x_1) \cdots \lambda(x_n)] - S(\rho) \right| < \epsilon, \quad (8.13)$$

where  $\lambda(x_i) = \lambda_j$  if  $|x_i\rangle$  is the letter  $|a_j\rangle$ . We define the  $\epsilon$ -typical subspace as the subspace spanned by the  $\epsilon$ -typical states. It can be shown that the dimension of this subspace is  $\approx 2^{nS(\rho)}$ . If  $P_{\text{typ}}$  denotes the projector on this subspace, then, for any  $\delta > 0$ , we have  $\text{Tr}(P_{\text{typ}}\rho^n) > 1 - \delta$ , provided  $n$  is large enough. Therefore, for  $n \rightarrow \infty$  the density matrix  $\rho^n$  has its support on a typical subspace of dimension  $2^{nS(\rho)}$ . A typical  $n$ -state message can then be encoded using  $nS(\rho)$  qubits.

### 8.2.2 Compression of an $n$ -qubit message

In this section, we follow the presentation of Schumacher (1998). Let us consider the binary alphabet  $\mathcal{A} = \{|\psi_0\rangle, |\psi_1\rangle\}$ , where  $|\psi_0\rangle \equiv |\tilde{0}\rangle$  and  $|\psi_1\rangle \equiv |\tilde{1}\rangle$  are the qubit states defined by Eq. (6.50). Assume that Alice wishes to send the following  $n$ -qubit message to Bob:

$$|\Psi_K\rangle = |\psi_{k_1}\rangle \otimes |\psi_{k_2}\rangle \otimes \cdots \otimes |\psi_{k_n}\rangle, \quad (8.14)$$

where  $K = \{k_1, k_2, \dots, k_n\}$  singles out the message ( $k_i = 0, 1$ ). The states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  are drawn from the alphabet  $\mathcal{A}$  with probabilities  $p$  and  $1 - p$ , respectively. Any  $n$ -letter message  $|\Psi_K\rangle$  belongs to the Hilbert space

$$\mathcal{H}^n = \mathcal{H}^{\otimes n}, \quad (8.15)$$

where  $\mathcal{H}$  is the Hilbert space for a single qubit. Thus,  $\mathcal{H}^n$  has dimension  $2^n$ . It is possible to diagonalize the density matrix

$$\rho = p|\tilde{0}\rangle\langle\tilde{0}| + (1-p)|\tilde{1}\rangle\langle\tilde{1}| \quad (8.16)$$

and then construct the typical subspace as explained in the previous subsection. A generic message  $|\psi_K\rangle$  can then be decomposed into a component belonging to the typical subspace (we call it  $\mathcal{H}_{\text{typ}}$ ) and another belonging to its orthogonal complement, known as the atypical subspace ( $\mathcal{H}_{\text{atyp}}$ ). We can write

$$|\Psi_K\rangle = \alpha_K|\tau_K\rangle + \beta_K|\tau_K^\perp\rangle, \quad (8.17)$$

where  $|\tau_K\rangle \in \mathcal{H}_{\text{typ}}$  and  $|\tau_K^\perp\rangle \in \mathcal{H}_{\text{atyp}}$ .

Alice performs a measurement to determine if  $|\Psi_K\rangle$  belongs to the typical subspace or not. If this is the case, the message is encoded and sent to Bob. Since the typical subspace has dimension  $\approx 2^{nS(\rho)}$ , we need only  $nS(\rho)$  qubits for the encoding (we shall see an example of encoding for  $n = 3$  qubits in Sec. 8.2.4). If instead  $|\Psi_K\rangle$  belongs to the atypical subspace, we substitute it with some reference state  $|R\rangle$  living in the typical subspace. Finally, Bob decodes the  $nS(\rho)$  qubits received from Alice and obtains a state described by the density matrix

$$\tilde{\rho}_K = |\alpha_K|^2|\tau_K\rangle\langle\tau_K| + |\beta_K|^2|R\rangle\langle R|. \quad (8.18)$$

How reliable is the transmission of quantum information by means of this procedure? A method to answer this question is to compute the *fidelity*  $f$ , defined as (see Sec. 5.2.3)

$$f = \langle\Psi_K|\tilde{\rho}_K|\Psi_K\rangle. \quad (8.19)$$

We have  $0 \leq f \leq 1$ , where the maximum value  $f = 1$  is obtained when the initial and final states coincide ( $\tilde{\rho}_K = |\Psi_K\rangle\langle\Psi_K|$ ), while  $f = 0$  when the initial and final states are orthogonal. The average fidelity  $\bar{f}$  is obtained after averaging over all the possible messages  $|\Psi_K\rangle$ , each weighted with the probability  $p_K$  of its occurrence:

$$\begin{aligned} \bar{f} &= \sum_K p_K \langle\Psi_K|\tilde{\rho}_K|\Psi_K\rangle \\ &= \sum_K p_K \langle\Psi_K| \left( |\alpha_K|^2|\tau_K\rangle\langle\tau_K| + |\beta_K|^2|R\rangle\langle R| \right) |\Psi_K\rangle \\ &= \sum_K p_K |\alpha_K|^4 + \sum_K |\beta_K|^2 |\langle\Psi_K|R\rangle|^2. \end{aligned} \quad (8.20)$$

It is possible to show that the average fidelity tends to 1 as  $n \rightarrow \infty$ . This means that, in this limit, messages have unit overlap with the typical subspace. Hence, we can code only the typical subspace and still achieve good fidelity.

### Comments

- (i) Alice could send Bob classical information and Bob could use this information to reconstruct Alice's  $n$ -qubit message (8.14). Indeed, she could send the sequence  $K = \{k_1, k_2, \dots, k_n\}$ , as this sequence uniquely determine the message  $|\Psi_K\rangle$ . According to Shannon's noiseless coding theorem, this sequence can be compressed by a factor given by the Shannon entropy  $H$ . However, this compression is not optimal if the alphabet is made of non-orthogonal quantum states. For instance, if the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  are taken with equal probability  $p_0 = p = \frac{1}{2}$  and  $p_1 = 1 - p = \frac{1}{2}$ , then  $H(\frac{1}{2}) = 1$ , whereas

$$S(\rho) = -\frac{1}{2}(1 + \sin 2\theta) \log\left[\frac{1}{2}(1 + \sin 2\theta)\right] - \frac{1}{2}(1 - \sin 2\theta) \log\left[\frac{1}{2}(1 - \sin 2\theta)\right], \quad (8.21)$$

see Eqs. (6.54)–(6.55), which is smaller than  $H(\frac{1}{2})$  as far as  $\theta \neq 0$ .

- (ii) The price to pay to compress the quantum information by a factor  $S < H$  is that Bob can reliably reconstruct the quantum state that Alice sent to him, but cannot know exactly what state he received. Indeed, each letter received is taken from a source of non-orthogonal quantum states and, as we know, non-orthogonal states cannot be distinguished with perfect reliability. Nevertheless, the compression of quantum information may be useful for several foreseen applications. For instance, one could compress the quantum memory of a quantum computer or transfer compressed quantum information between different quantum processors.

#### 8.2.3 Example 1: two-qubit messages

This simple example illustrates the difference between the compression of classical and quantum messages in the case in which the letters are represented by non-orthogonal quantum states. Let us consider the alphabet  $\mathcal{A} = \{|\tilde{0}\rangle, |\tilde{1}\rangle\}$  defined by Eq. (6.50). We assume that the state  $|\tilde{0}\rangle$  is drawn from the alphabet  $\mathcal{A}$  with probability  $p$  and the state  $|\tilde{1}\rangle$  with probability  $1 - p$ . Alice generates a two-qubit message but she can only afford to send Bob a single qubit. Bob receives this qubit and guesses that the second letter of the message is some reference state, say  $|\tilde{0}\rangle$ . What is the fidelity of his guess? Let us first compute the fidelities  $f_K = |\langle\psi_2|\tilde{0}\rangle|^2$  of the four possible messages,  $|\psi_2\rangle$  being the actual state of the second qubit. We have  $f_K = 1$  if  $|\psi_2\rangle = |\tilde{0}\rangle$  and  $f_K = \sin^2 2\theta$  if  $|\psi_2\rangle = |\tilde{1}\rangle$ .

$K$	Message	$p_K$	Bob's guess	$f_K$
0	$ \tilde{0}\tilde{0}\rangle$	$p^2$	$ \tilde{0}\tilde{0}\rangle$	1
1	$ \tilde{0}\tilde{1}\rangle$	$p(1 - p)$	$ \tilde{0}\tilde{0}\rangle$	$\sin^2 2\theta$
2	$ \tilde{1}\tilde{0}\rangle$	$p(1 - p)$	$ \tilde{1}\tilde{0}\rangle$	1
3	$ \tilde{1}\tilde{1}\rangle$	$(1 - p)^2$	$ \tilde{1}\tilde{0}\rangle$	$\sin^2 2\theta$

We can readily compute the average fidelity

$$\bar{f} = \sum_K p_K f_K = p \cos^2 2\theta + \sin^2 2\theta, \quad (8.22)$$

which is shown in Fig. 8.1 for various values of  $\theta$ . We note that  $\theta = 0$  (transmission of orthogonal states) corresponds to the classical case. Indeed, we can define a classical fidelity  $f_{c,K}$ , which is equal to 1 if a message is correctly transmitted (in our example, for  $K = 0$  and  $K = 2$ ) and equal to 0 otherwise (for  $K = 1$  and  $K = 3$ ). It turns out that the average classical fidelity  $\bar{f}_c = \sum_K p_K f_{c,K} = p$  is equal to the quantum fidelity for  $\theta = 0$ . For  $\theta \neq 0$ , the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  are no longer orthogonal and therefore the fidelity is higher (we have  $f_1 = f_3 = \sin^2 2\theta > 0$ , while  $f_{c,1} = f_{c,3} = 0$ ). In the limiting case  $\theta = \pi/4$ , the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  coincide and therefore  $f = 1$  for any value of  $p$ . Note that in this case no information is transmitted since the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  cannot be distinguished by any measurement.

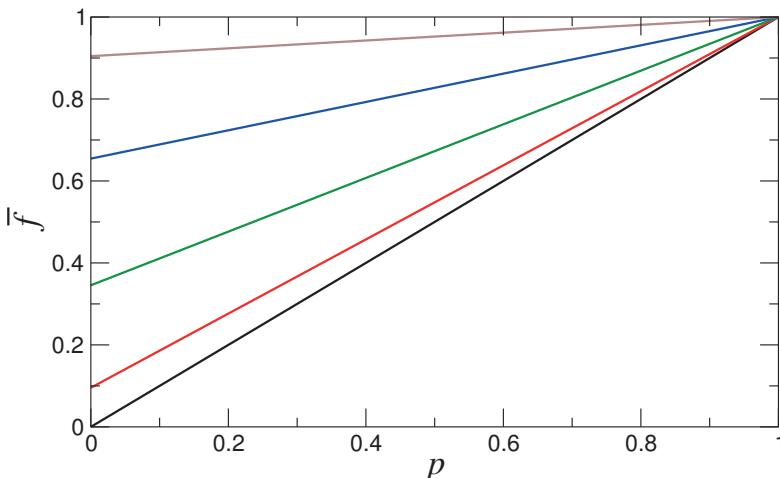


Fig. 8.1 The average fidelity  $\bar{f}$  for a two-qubit message when only the first qubit is sent (see text). The values of the angle  $\theta$  are:  $\theta = 0$  (black),  $\theta = 0.2 \times \frac{\pi}{4}$  (red),  $\theta = 0.4 \times \frac{\pi}{4}$  (green),  $\theta = 0.6 \times \frac{\pi}{4}$  (blue) and  $\theta = 0.8 \times \frac{\pi}{4}$  (brown).

### 8.2.4 Example 2: three-qubit messages

In order to clarify the principles of quantum data compression, it is useful to consider the example of a message consisting of three qubits chosen from the ensemble  $\{|\tilde{0}\rangle, |\tilde{1}\rangle\}$  with *a priori* probabilities  $\{p, 1-p\}$ , where  $|\tilde{0}\rangle$  is the letter generated with probability  $p \geq \frac{1}{2}$ . Let us assume that Alice can only afford to send Bob two qubits and that Alice and Bob wish to devise a strategy to maximize the average fidelity for the transmission of a three-qubit messages.

Each letter of the message is described by the density matrix  $\rho = p|\tilde{0}\rangle\langle\tilde{0}| + (1-p)|\tilde{1}\rangle\langle\tilde{1}|$ , whose eigenvalues  $\lambda_{\pm}$  were written down in Eq. (6.54). The corresponding

eigenstates are given by

$$|\pm\rangle = \frac{1}{\sqrt{(\lambda_{\pm} + p \cos 2\theta - C^2)^2 + C^2 S^2}} \begin{bmatrix} \lambda_{\pm} + p \cos 2\theta - C^2 \\ CS \end{bmatrix}, \quad (8.23)$$

where we have used the shorthand notation  $C = \cos \theta$  and  $S = \sin \theta$ . It is also useful to write down the inner products

$$\langle \tilde{0} | \pm \rangle = \frac{C[\lambda_{\pm} + p \cos 2\theta - C^2] + CS^2}{\sqrt{N_{\pm}}}, \quad (8.24a)$$

$$\langle \tilde{1} | \pm \rangle = \frac{S[\lambda_{\pm} + p \cos 2\theta - C^2] + C^2 S}{\sqrt{N_{\pm}}}, \quad (8.24b)$$

where we have defined

$$N_{\pm} = (\lambda_{\pm} + p \cos 2\theta - C^2)^2 + C^2 S^2. \quad (8.25)$$

We call  $|\Psi_K\rangle$  the 8 possible messages,

$$|\Psi_0\rangle = |\tilde{0}\tilde{0}\tilde{0}\rangle, \quad |\Psi_1\rangle = |\tilde{0}\tilde{0}\tilde{1}\rangle, \quad \dots, \quad |\Psi_7\rangle = |\tilde{1}\tilde{1}\tilde{1}\rangle, \quad (8.26)$$

and  $|\chi_J\rangle$  the eigenstates of  $\rho^{\otimes 3}$ :

$$|\chi_0\rangle = |+++ \rangle, \quad |\chi_1\rangle = |++-\rangle, \quad \dots, \quad |\chi_7\rangle = |---\rangle, \quad (8.27)$$

where  $|+\rangle$  and  $|-\rangle$  are the eigenstates (8.23) of  $\rho$ . The states  $\{|\chi_J\rangle\}$  constitute a basis for the three-qubit Hilbert space and we can therefore decompose the possible messages as follows:

$$|\Psi_K\rangle = \sum_J c_{KJ} |\chi_J\rangle, \quad (8.28)$$

where we have defined  $c_{KJ} = \langle \chi_J | \Psi_K \rangle$ .

Since  $\lambda_+ > \lambda_-$  for  $p > \frac{1}{2}$ , then in the spectral decomposition of the density matrix  $\rho$  the weight  $\lambda_+$  of the eigenstate  $|+\rangle$  is higher than the weight  $\lambda_-$  of the eigenstate  $|-\rangle$ . The most likely subspace is spanned by the most likely states, namely

$$\{|\chi_0\rangle = |+++ \rangle, \quad |\chi_1\rangle = |++-\rangle, \quad |\chi_2\rangle = |+-+\rangle, \quad |\chi_4\rangle = |-++\rangle\}, \quad (8.29)$$

while the unlikely subspace is spanned by

$$\{|\chi_3\rangle = |+--\rangle, \quad |\chi_5\rangle = |-+-\rangle, \quad |\chi_6\rangle = |--+\rangle, \quad |\chi_7\rangle = |---\rangle\}. \quad (8.30)$$

The states  $|\Psi_K\rangle$  of the message can be decomposed into a component  $|\tau_K\rangle$  along the likely subspace and a component  $|\tau_K^\perp\rangle$  along the unlikely subspace; that is,  $|\Psi_K\rangle = \alpha_K |\tau_K\rangle + \beta_K |\tau_K^\perp\rangle$ . The coefficients  $\alpha_K$  and  $\beta_K$  are given by

$$\begin{aligned} \alpha_K &= \sqrt{|c_{K0}|^2 + |c_{K1}|^2 + |c_{K2}|^2 + |c_{K4}|^2}, \\ \beta_K &= \sqrt{|c_{K3}|^2 + |c_{K5}|^2 + |c_{K6}|^2 + |c_{K7}|^2}, \end{aligned} \quad (8.31)$$

where the coefficients  $c_{Ki}$  can be easily computed by exploiting expressions (8.24a) and (8.24b) for the inner products  $\langle \tilde{0} | \pm \rangle$  and  $\langle \tilde{1} | \pm \rangle$ .

In order to code the message, Alice employs the following strategy. She applies a unitary transformation  $U$  that rotates the basis states spanning the likely subspace ( $|\chi_0\rangle, |\chi_1\rangle, |\chi_2\rangle$  and  $|\chi_4\rangle$ ) into the states  $|i_1\rangle|i_2\rangle|0\rangle$  (with  $i_1, i_2 = 0, 1$ ), whereas the unlikely states  $|\chi_3\rangle, |\chi_5\rangle, |\chi_6\rangle$  and  $|\chi_7\rangle$  are rotated into  $|i_1\rangle|i_2\rangle|1\rangle$ . She then performs a measurement of the third qubit: if she obtains 0, her state  $|\Psi_K\rangle$  has been projected onto the likely subspace. In this case, she sends the first two qubits to Bob. If instead she obtains outcome 1, her state has been projected onto the unlikely subspace and she sends Bob the first two qubits of  $U|R\rangle$ , where  $|R\rangle$  is some reference state belonging to the likely subspace. For instance, she takes  $|R\rangle$  equal to the most likely state  $|\chi_0\rangle$ . Bob appends to the two qubits received an ancillary qubit, prepared in the state  $|0\rangle$ . He then applies the operator  $U^{-1}$  to these three qubits and ends up with a state described by the density matrix

$$\tilde{\rho}_K = |\alpha_K|^2|\tau_K\rangle\langle\tau_K| + |\beta_K|^2|R\rangle\langle R|. \quad (8.32)$$

The average fidelity is then given by

$$\bar{f} = \sum_{K=0}^7 p_K \langle \Psi_K | \tilde{\rho}_K | \Psi_K \rangle = \sum_{K=0}^7 p_K \left( |\alpha_K|^4 + |\beta_K|^2 |\langle \Psi_K | R \rangle|^2 \right), \quad (8.33)$$

where  $p_K$  is the probability that the message  $|\Psi_K\rangle$  is generated. The graph of  $\bar{f}$  as a function of  $p$  is shown in Fig. 8.2, for various values of  $\theta$ . This figure exhibits several interesting features. First of all, for  $\theta = 0$  we recover the classical case, in which the average fidelity  $\bar{f}_c$  is obtained after summing the probabilities of all messages correctly transmitted. The calculation is similar to that performed in the previous subsection for two-qubit messages and gives

$$\bar{f}_c = p^3 + 3p^2(1-p) = 3p^2 - 2p^3. \quad (8.34)$$

We note that for  $p = \frac{1}{2}$  we have  $\bar{f}_c = \frac{1}{2}$ . Indeed, in this case we have 8 messages occurring with the same probability and only 4 are correctly transmitted. The average quantum fidelity  $\bar{f}$  is instead larger than  $\frac{1}{2}$  when  $\theta > 0$ . This is because our *a priori* ignorance for non-orthogonal states is smaller than for orthogonal states. In the limiting case  $\theta = \pi/4$  the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  superimpose. Thus,  $\bar{f}(\pi/4) = 1$  but there is no transmission of information.

### 8.3 Accessible information

We assume that Alice sends Bob a message whose letters are chosen independently from the alphabet  $\mathcal{A} = \{a_1, \dots, a_k\}$  with *a priori* probabilities  $\{p_1, \dots, p_k\}$ . The letters of the alphabet are coded by quantum states that are not necessarily orthogonal. In this section we consider the following problem: how much information can Bob gain on the message by performing measurements on the quantum states received? This problem is non-trivial since non-orthogonal quantum states cannot be perfectly distinguished. It is important to emphasize that, as we saw in Chap. 5, this property lies at the heart of quantum cryptography.

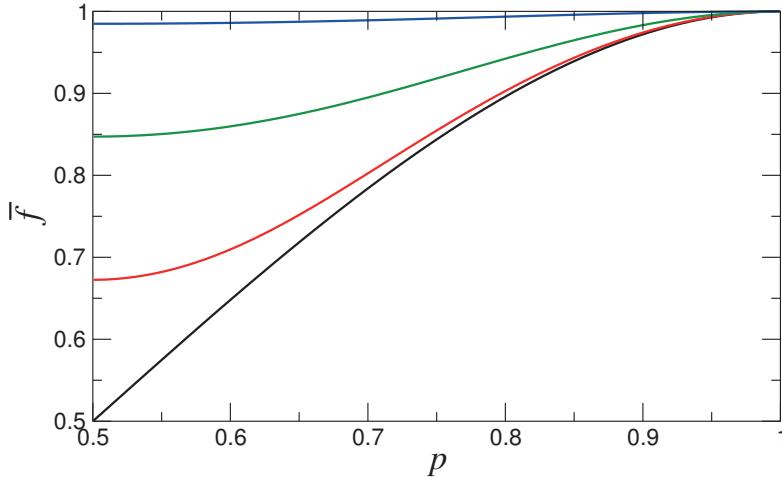


Fig. 8.2 The average fidelity  $\bar{f}$  for the transmission of a three-qubit message by means of a two-qubit code (see details in the text). From bottom to top:  $\theta = 0$  (black),  $\pi/16$  (red),  $\pi/10$  (green) and  $\pi/6$  (blue).

If  $X$  and  $Y$  denote the random variables associated with the letters generated by Alice and with Bob's measurement outcomes, respectively, then the *accessible information* is defined as the maximum of the mutual information  $\mathcal{I}(X:Y)$  over all possible measurement schemes.

### 8.3.1 The Holevo bound

The Holevo bound (proved in Holevo, 1973) establishes an upper bound on the accessible information.

**Theorem 8.4** The Holevo Bound: *If Alice prepares a (mixed) state  $\rho_X$  chosen from the ensemble  $\mathcal{A} = \{\rho_0, \dots, \rho_k\}$  with a priori probabilities  $\{p_1, \dots, p_k\}$  and Bob performs a POVM measurement on that state, with POVM elements  $\{F_1, \dots, F_l\}$  and measurement outcome described by the random variable  $Y$ , then the mutual information  $\mathcal{I}(X:Y)$  is bounded as follows:*

$$\mathcal{I}(X:Y) \leq S(\rho) - \sum_{i=1}^k p_i S(\rho_i) \equiv \chi(\mathcal{E}), \quad (8.35)$$

where  $\rho = \sum_{i=1}^k p_i \rho_i$  and  $\chi(\mathcal{E})$  is known as the Holevo information of the ensemble  $\mathcal{E} \equiv \{\rho_1, \dots, \rho_k; p_1, \dots, p_k\}$ .

A proof of this theorem can be found in Nielsen and Chuang (2000). Here, we shall limit ourselves to discuss the Holevo bound in a few concrete examples.

### 8.3.2 Example 1: two non-orthogonal pure states

If Alice sends Bob pure orthogonal quantum states drawn from the ensemble  $\{|\psi_1\rangle, \dots, |\psi_k\rangle\}$ , then Bob can unambiguously distinguish these states by means of projective measurements described by the POVM elements (in this case, simple projectors)  $\{F_1 = |\psi_1\rangle\langle\psi_1|, \dots, F_k = |\psi_k\rangle\langle\psi_k|\}$ . It is easy to check that  $\mathcal{I}(X:Y) = H(X)$  (we have  $H(X|Y) = 0$ ) and therefore this case is not different from the transmission of classical information over a noiseless channel: if we send the letter  $a_x$ , we recover the same letter; that is,  $a_y = a_x$ .

The simplest example that cannot be reduced to classical information theory is that in which Alice sends Bob states generated by a source of non-orthogonal pure quantum states. We assume that the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$ , defined by Eq. (6.50), are generated with probabilities  $p_0 = p$  and  $p_1 = 1 - p$ , respectively.

Since the single letters are represented in this case by pure states, their von Neumann entropy is equal to zero:  $S(\rho_0) = S(|\tilde{0}\rangle\langle\tilde{0}|) = 0$  and  $S(\rho_1) = S(|\tilde{1}\rangle\langle\tilde{1}|) = 0$ . Therefore, the Holevo information  $\chi(\mathcal{E})$  reduces to

$$\chi(\mathcal{E}) = S(\rho), \quad (8.36)$$

where  $\rho = p\rho_0 + (1 - p)\rho_1$ . Hence, the Holevo bound gives

$$\mathcal{I}(X:Y) \leq S(\rho). \quad (8.37)$$

A plot of  $S(\rho)$  was already shown in Fig. 6.5. It reveals that, for non-orthogonal states ( $\theta \neq 0$ ),  $S(\rho) < H(X)$  and therefore  $\mathcal{I}(X:Y) < H(X)$ . It is possible to show that this strict inequality also has general validity for mixed states  $\{\rho_i\}$ , provided they do not have orthogonal support (for orthogonal support, it readily follows from Eq. (6.42) that  $\mathcal{I}(X:Y) = H(X)$ ).

It is instructive to consider the following special case: we assume that Bob performs a projective measurement on the received qubits along the direction  $\hat{\mathbf{n}}$  (that is, he measures  $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$ ) and we show that in this case the Holevo bound is satisfied. For this purpose, we compute the mutual information. Bob's measurement along the direction  $\hat{\mathbf{n}}$  is described by the POVM elements (projectors)

$$F_0 = \frac{1}{2}(I + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}), \quad F_1 = \frac{1}{2}(I - \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}). \quad (8.38)$$

For instance, if  $\hat{\mathbf{n}} = (0, 0, 1)$ , then  $F_0 = |0\rangle\langle 0|$  and  $F_1 = |1\rangle\langle 1|$ . We compute the conditional probability

$$p(y|x) = \text{Tr}(\rho_x F_y), \quad (x, y = 0, 1), \quad (8.39)$$

which is the probability that Bob's measurement gives outcome  $y$ , provided the state  $\rho_x$  was sent by Alice. For this purpose, we write down the Bloch-sphere representation of the density matrices associated with the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  (see Sec. 3.1):

$$\rho_0 = |\tilde{0}\rangle\langle\tilde{0}| = \frac{1}{2}(I + \mathbf{r}_0 \cdot \boldsymbol{\sigma}), \quad \rho_1 = |\tilde{1}\rangle\langle\tilde{1}| = \frac{1}{2}(I + \mathbf{r}_1 \cdot \boldsymbol{\sigma}), \quad (8.40)$$

where the Cartesian components of the Bloch vectors  $\mathbf{r}_0$  and  $\mathbf{r}_1$  are given by

$$\mathbf{r}_0 = (\sin 2\theta, 0, \cos 2\theta), \quad \mathbf{r}_1 = (\sin 2\theta, 0, -\cos 2\theta). \quad (8.41)$$

Taking into account that  $\text{Tr}(\sigma_i) = 0$  and  $\text{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}$  for  $i, j = x, y, z$  (see exercise 3.2), it is now straightforward to compute the conditional probabilities:

$$\begin{aligned} p(0|0) &= \text{Tr}(\rho_0 F_0) = \frac{1}{2}(1 + \mathbf{r}_0 \cdot \hat{\mathbf{n}}), \\ p(1|0) &= \text{Tr}(\rho_0 F_1) = \frac{1}{2}(1 - \mathbf{r}_0 \cdot \hat{\mathbf{n}}), \\ p(0|1) &= \text{Tr}(\rho_1 F_0) = \frac{1}{2}(1 + \mathbf{r}_1 \cdot \hat{\mathbf{n}}), \\ p(1|1) &= \text{Tr}(\rho_1 F_1) = \frac{1}{2}(1 - \mathbf{r}_1 \cdot \hat{\mathbf{n}}). \end{aligned} \quad (8.42)$$

If, for the sake of simplicity, we assume that the measurement direction lies in the  $(x, z)$  plane of the Bloch sphere; that is,  $\hat{\mathbf{n}} = (\sin \bar{\theta}, 0, \cos \bar{\theta})$  (see Fig. 8.3), we have

$$\begin{aligned} p(0|0) &= \frac{1}{2}[1 + \cos(\bar{\theta} - 2\theta)], & p(1|0) &= \frac{1}{2}[1 - \cos(\bar{\theta} - 2\theta)], \\ p(0|1) &= \frac{1}{2}[1 - \cos(\bar{\theta} + 2\theta)], & p(1|1) &= \frac{1}{2}[1 + \cos(\bar{\theta} + 2\theta)]. \end{aligned} \quad (8.43)$$

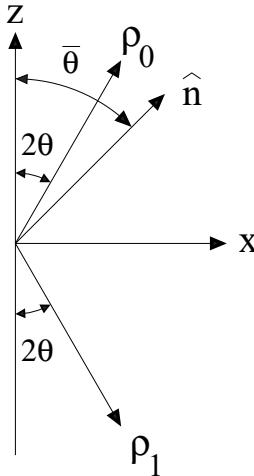


Fig. 8.3 A geometric visualization of the Bloch sphere vectors  $\rho_0$  and  $\rho_1$  and of the measurement axis  $\hat{\mathbf{n}}$ .

We now compute  $p(x, y) = p(x)p(y|x)$ , where, as stated at the beginning of this subsection, we assume that the states  $|\tilde{0}\rangle$  and  $|\tilde{1}\rangle$  are generated with probabilities  $p(X = 0) = p$  and  $p(X = 1) = 1 - p$ , respectively. We thus have

$$\begin{aligned} p(0, 0) &= \frac{1}{2}p[1 + \cos(\bar{\theta} - 2\theta)], \\ p(0, 1) &= \frac{1}{2}p[1 - \cos(\bar{\theta} - 2\theta)], \\ p(1, 0) &= \frac{1}{2}(1 - p)[1 - \cos(\bar{\theta} + 2\theta)], \\ p(1, 1) &= \frac{1}{2}(1 - p)[1 + \cos(\bar{\theta} + 2\theta)]. \end{aligned} \quad (8.44)$$

Then we compute  $p(y) = \sum_x p(x, y)$  and obtain

$$\begin{aligned} p(Y=0) &= \frac{1}{2}[1 + p \cos(\bar{\theta} - 2\theta) - (1 - p) \cos(\bar{\theta} + 2\theta)], \\ p(Y=1) &= \frac{1}{2}[1 - p \cos(\bar{\theta} - 2\theta) + (1 - p) \cos(\bar{\theta} + 2\theta)]. \end{aligned} \quad (8.45)$$

Finally, we insert the expressions derived for  $p(x)$ ,  $p(y)$  and  $p(x,y)$  into Eq. (6.33), obtaining the mutual information  $\mathcal{I}(X:Y)$ .

As an example, in Fig. 8.4 we show the mutual information  $\mathcal{I}(X:Y)$  for  $\theta = \pi/10$  and  $p = 0.8$ . Within the chosen measurement scheme, the only free parameter that may be varied in order to maximize  $\mathcal{I}$  is  $\bar{\theta}$ . The maximum value  $\mathcal{I}_{\max} \equiv \max_{\bar{\theta}} \mathcal{I}(\bar{\theta}) \approx 0.40$  is attained for  $\bar{\theta} \approx 0.14\pi$ . We stress that this value is below the Holevo bound  $\chi = S(\rho) \approx 0.526$ . Of course, this value is also smaller than the classical bound,  $\mathcal{I}(X:Y) \leq H(X) \approx 0.722$ .

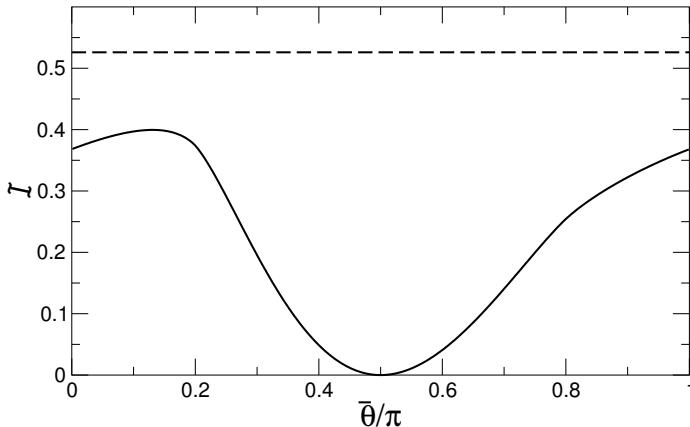


Fig. 8.4 The mutual information  $\mathcal{I}(X:Y)$  for a message coded by means of the non-orthogonal states (6.50), with  $\theta = \pi/10$  and  $p = 0.8$ . The angle  $\bar{\theta}$  determines the measurement direction  $\hat{n} = (\sin \bar{\theta}, 0, \cos \bar{\theta})$ . The dashed line shows the Holevo bound  $\chi \approx 0.526$ .

**Exercise 8.4** Alice sends Bob the state  $|0\rangle$  with probability  $p$  or the state  $|1\rangle$  with probability  $1 - p$ . For this purpose, they employ a quantum channel whose action on a state with Bloch vector  $(x, y, z)$  is given by

$$x \rightarrow x' = ax, \quad y \rightarrow y' = ay, \quad z \rightarrow z' = z, \quad (8.46)$$

where  $0 < a < 1$  (note that this quantum channel corresponds to the phase-flip channel, which was discussed in Sec. 7.2.5). Finally, Bob performs a standard projective measurement along the direction  $\hat{n}$ . Compute Alice and Bob's mutual information.

**Exercise 8.5** Repeat the previous exercise for the case in which the action of the quantum channel on a state with Bloch vector  $(x, y, z)$  is given by

$$x \rightarrow x' = x \cos \theta, \quad y \rightarrow y' = y \cos \theta, \quad z \rightarrow z' = \sin^2 \theta + z \cos^2 \theta, \quad (8.47)$$

where  $0 \leq \theta \leq \frac{\pi}{2}$  (this quantum channel corresponds to the amplitude-damping channel, which was discussed in Sec. 7.2.8).

### 8.3.3 \* Example 2: three non-orthogonal pure states

Let us now consider the case in which Alice's alphabet is  $\mathcal{A} = \{|\phi_0\rangle, |\phi_1\rangle, |\phi_2\rangle\}$ , where

$$|\phi_0\rangle = |0\rangle, \quad |\phi_1\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle, \quad |\phi_2\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle. \quad (8.48)$$

A graphical representation of these three non-orthogonal quantum states is shown in Fig. 8.5. We call  $\rho_0 = |\phi_0\rangle\langle\phi_0|$ ,  $\rho_1 = |\phi_1\rangle\langle\phi_1|$  and  $\rho_2 = |\phi_2\rangle\langle\phi_2|$  the density operators associated with these quantum states. We assume that each letter of Alice's message is one of the three states of this alphabet, chosen with *a priori* probabilities  $\{p_0, p_1, p_2\}$ . In the following we assume that  $p_0 = p_1 = p_2 = p = 1/3$  and  $\theta = 2\pi/3$ . Under these conditions, the matrix representations of the density operators  $\rho_0$ ,  $\rho_1$  and  $\rho_2$  in the  $\{|0\rangle, |1\rangle\}$  basis read:

$$\rho_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \rho_1 = \frac{1}{4} \begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix}, \quad \rho_2 = \frac{1}{4} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}. \quad (8.49)$$

The density matrix that describes the above ensemble of pure quantum states is

$$\rho = p_0\rho_0 + p_1\rho_1 + p_2\rho_2 = \frac{1}{2}I, \quad (8.50)$$

and therefore  $S(\rho) = 1$ . Since the letters of Alice's message are pure states, the Holevo bound on mutual information gives  $\mathcal{I}(X:Y) \leq S(\rho) = 1$ .

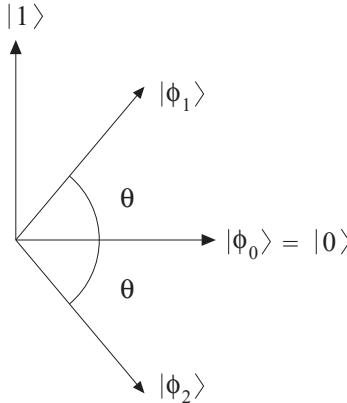


Fig. 8.5 A graphical representation of the three quantum states of the alphabet  $\mathcal{A} = \{|\phi_0\rangle, |\phi_1\rangle, |\phi_2\rangle\}$ .

Bob measures the received qubits by means of the POVM scheme described in Sec. 2.9.1, with POVM elements

$$F_0 = \frac{1}{2} \begin{bmatrix} 1 & r \\ r & r^2 \end{bmatrix}, \quad F_1 = \frac{1}{2} \begin{bmatrix} 1 & -r \\ -r & r^2 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1-r^2 \end{bmatrix}. \quad (8.51)$$

As in the previous example, we compute the conditional probabilities

$$p(y|x) = \text{Tr}(\rho_x F_y), \quad (x, y = 0, 1, 2), \quad (8.52)$$

namely the probability that Bob's POVM measurement gives outcome  $y$ , provided the state  $x$  was sent by Alice. We obtain

$$\begin{aligned} p(0|0) &= \frac{1}{2}, & p(0|1) &= \frac{1}{8}(1 - \sqrt{3}r)^2, & p(0|2) &= \frac{1}{8}(1 + \sqrt{3}r)^2, \\ p(1|0) &= \frac{1}{2}, & p(1|1) &= \frac{1}{8}(1 + \sqrt{3}r)^2, & p(1|2) &= \frac{1}{8}(1 - \sqrt{3}r)^2, \\ p(2|0) &= 0, & p(2|1) &= \frac{3}{4}(1 - r^2), & p(2|2) &= \frac{3}{4}(1 - r^2). \end{aligned} \quad (8.53)$$

We now compute  $p(x,y) = p(x)p(y|x)$ . In this case,  $p(x,y) = \frac{1}{3}p(y|x)$  since  $p(X=0) = p(X=1) = p(X=2) = \frac{1}{3}$ . We then compute  $p(y) = \sum_x p(x,y)$  and obtain

$$p(Y=0) = p(Y=1) = \frac{1}{4}(1 + r^2), \quad p(Y=2) = \frac{1}{2}(1 - r^2). \quad (8.54)$$

Finally, we insert the above expressions for  $p(x)$ ,  $p(y)$  and  $p(x,y)$  into Eq. (6.33), thus obtaining the mutual information  $\mathcal{I}(X:Y)$ . The graph of  $\mathcal{I}$  as a function of the parameter  $r$  is shown in Fig. 8.6. Its maximum value  $\mathcal{I}_{\max} \approx 0.585$  is attained for  $r \approx 0.577$ . Note that  $\mathcal{I}_{\max}$  is well below the Holevo bound  $\chi = S(\rho) = 1$ .

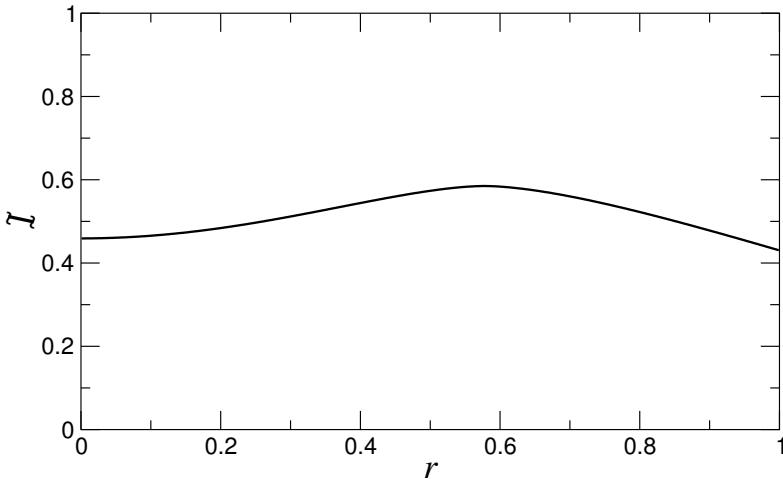


Fig. 8.6 The mutual information  $\mathcal{I}(X:Y)$  for a message coded by means of the three non-orthogonal states given by Eq. (8.48), with  $\theta = \frac{2}{3}\pi$  and  $p = \frac{1}{3}$ . Bob's measurement is described by the POVM operators (8.51). The Holevo bound gives  $\mathcal{I} \leq 1$ .

## 8.4 Capacities of quantum channels

This section deals with analogues to Shannon's noisy channel coding theorem for quantum information, when a quantum channel is used to transmit either classical or quantum information. We assume that the channel is memoryless and stationary, namely the channel acts the same way each time it is used, independently of the previous history. Proofs of the theorems stated below can be found, for instance, in Wilde (2013). Here, we will limit ourselves to illustrative examples.

### 8.4.1 Classical capacity

The classical capacity  $C$  of a quantum channel is the maximum number of bits of classical information that can be reliably transmitted per channel use. The term reliably means, as usual in information theory, that the probability of errors or information losses can be made arbitrarily small when the number  $n$  of channel uses is sufficiently large. The channel is generally noisy and we assume that its action of an input state  $\rho$  is described by the quantum operation  $\mathbb{S}$ . If the sender (Alice) encodes her messages by means of product states of the form  $\rho_{x_1} \otimes \rho_{x_2} \otimes \dots$ , then we are considering the so-called *product-state capacity*  $C_1$ . The Holevo-Schumacher-Westmoreland (HSW) theorem provides a formula for the product-state capacity.

**Theorem 8.5** Holevo-Schumacher-Westmoreland theorem: *Given a quantum channel described by a superoperator  $\mathbb{S}$ , the product-state capacity for the channel is given by the Holevo information  $\chi(\mathbb{S})$ :*

$$C_1(\mathbb{S}) = \chi(\mathbb{S}) \equiv \max_{\mathcal{E}} \left\{ S \left[ \mathbb{S} \left( \sum_i p_i \rho_i \right) \right] - \sum_i p_i S[\mathbb{S}(\rho_i)] \right\}, \quad (8.55)$$

where the Holevo information of the quantum channel is the maximum over all ensembles  $\mathcal{E} \equiv \{\rho_1, \rho_2 \dots; p_1, p_2 \dots\}$  of possible input states  $\{\rho_i\}$  to the channel, with a priori probabilities  $\{p_i\}$ .

Equation (8.55) is very convenient in that it is sufficient to optimize for single uses of the channel.<sup>3</sup> However, there exist channels (see Hastings, 2009) for which the Holevo information is *superadditive*, that is,

$$\chi(\mathbb{S}_1 \otimes \mathbb{S}_2) > \chi(\mathbb{S}_1) + \chi(\mathbb{S}_2). \quad (8.56)$$

Therefore, one should in general optimize over all possible  $n$ -letter input ensembles, possibly entangled across many uses of the channel, in the limit  $n \rightarrow \infty$ . The corresponding *regularized* formula for the classical capacity is given by

$$C(\mathbb{S}) = \lim_{n \rightarrow \infty} \frac{C_1(\mathbb{S}^{\otimes n})}{n} = \lim_{n \rightarrow \infty} \frac{\chi(\mathbb{S}^{\otimes n})}{n}. \quad (8.57)$$

While this formula is not very useful for practical purposes, since it requires optimization over ensembles of arbitrary length, there exist significant examples where  $\chi$  is additive and therefore  $C(\mathbb{S}) = C_1(\mathbb{S}) = \chi(\mathbb{S})$ , see for instance exercise 8.7 for the depolarizing channel.

**Exercise 8.6** Show that the classical capacity of a noiseless quantum state is given by  $C = \max_{\rho} S(\rho)$ .

<sup>3</sup>For any ensemble  $\mathcal{E} = \{\rho_i, p_i\}$  of mixed states, described by the density operator  $\rho = \sum_i p_i \rho_i$ , one can find an ensemble  $\mathcal{E}' = \{|\psi_\alpha\rangle\langle\psi_\alpha|, q_\alpha\}$  of pure states described by same density operator,  $\rho = \sum_\alpha q_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$ , and whose Holevo quantity is at least as large, see Schumacher and Westmoreland (1997). Hence the maximum in Eq. (8.55) can be achieved using ensembles of pure states.

**Exercise 8.7** For the depolarizing channel,  $C = C_1$  and the classical capacity can be achieved by choosing states from the ensemble  $\{\rho_1 = |0\rangle\langle 0|, \rho_2 = |1\rangle\langle 1|\}$ , see for instance Wilde (2013). Show that the classical capacity

$$C = 1 - H_{\text{bin}}(p), \quad (8.58)$$

where we have reparametrized the depolarizing channel of Sec. 7.2.7 as follows:

$$\rho' = \mathbb{S}(\rho) = (1-p)\rho + p \frac{I}{2}. \quad (8.59)$$

**Exercise 8.8** Given the bit-flip channel

$$\rho' = \mathbb{S}(\rho) = (1-p)\rho + p\sigma_x\rho\sigma_x, \quad (8.60)$$

maximize the Holevo information of the ensemble  $\{\rho_1 = |0\rangle\langle 0|, \rho_2 = |1\rangle\langle 1|; p_1, p_2 = 1-p\}$  over  $p_1$ , thus obtaining a lower bound  $C_{\text{lb}}$  to the classical capacity  $C$ . After that, compute  $C$ .

### 8.4.2 Quantum capacity

In this section, we briefly review basic quantities and concepts that are useful to describe the channel capability to transmit quantum information. For this purpose, we consider as quantum information carrier a quantum system  $\mathcal{Q}$  and describe the channel action on  $\mathcal{Q}$  by means of a superoperator  $\mathbb{S}$ , which transforms the (generally mixed) input state  $\rho_{\mathcal{Q}}$  into the output state  $\rho'_{\mathcal{Q}}$ .

#### Entanglement fidelity

A proper way to measure reliability of quantum information transmission is the entanglement fidelity. To define this quantity we look at the system  $\mathcal{Q}$  as a part of a larger quantum system  $\mathcal{R}\mathcal{Q}$  (see Fig. 8.7), initially in a pure entangled state  $|\psi_{\mathcal{R}\mathcal{Q}}\rangle$ . The density operator of system  $\mathcal{Q}$  is then obtained from that of  $\mathcal{R}\mathcal{Q}$  by a partial trace over  $\mathcal{R}$ :  $\rho_{\mathcal{Q}} = \text{Tr}_{\mathcal{R}}[|\psi_{\mathcal{R}\mathcal{Q}}\rangle\langle\psi_{\mathcal{R}\mathcal{Q}}|]$ . The system  $\mathcal{Q}$  is sent through the channel and undergoes the transformation  $\mathbb{S}$ , while  $\mathcal{R}$  is ideally isolated from the environment. The final state of the composite system is:

$$\rho'_{\mathcal{R}\mathcal{Q}} = (\mathbb{I} \otimes \mathbb{S})(|\psi_{\mathcal{R}\mathcal{Q}}\rangle\langle\psi_{\mathcal{R}\mathcal{Q}}|), \quad (8.61)$$

where  $\mathbb{S}$  is the identity superoperator acting on the reference system. Let us consider the following question: how faithfully does the channel preserve the entanglement between the two systems  $\mathcal{Q}$  and  $\mathcal{R}$ ? The answer is the *entanglement fidelity*  $f_e$ , defined as the fidelity  $f$  between the initial pure state  $|\psi_{\mathcal{R}\mathcal{Q}}\rangle$  and the final (generally mixed) state  $\rho'_{\mathcal{R}\mathcal{Q}}$ :

$$f_e(\rho_{\mathcal{Q}}, \mathbb{S}) = \langle\psi_{\mathcal{R}\mathcal{Q}}|(\mathbb{I} \otimes \mathbb{S})(|\psi_{\mathcal{R}\mathcal{Q}}\rangle\langle\psi_{\mathcal{R}\mathcal{Q}}|)|\psi_{\mathcal{R}\mathcal{Q}}\rangle. \quad (8.62)$$

Entanglement fidelity depends only on the initial state  $\rho_{\mathcal{Q}}$  of the system and on the channel action  $\mathbb{S}$ , not on the particular purification chosen. Indeed, using the

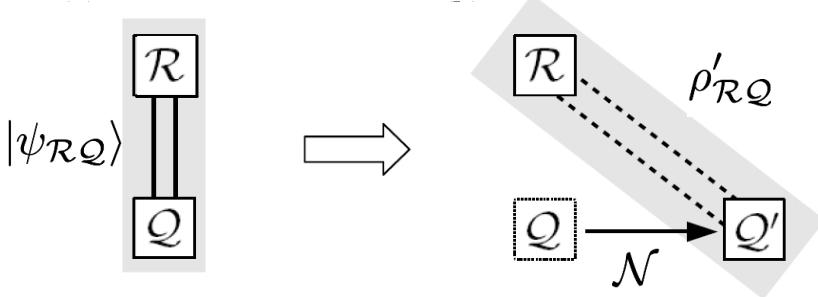


Fig. 8.7 The system  $\mathcal{Q}$  is considered as entangled with a reference system  $\mathcal{R}$ , so that the initial state  $|\psi^{\mathcal{R}\mathcal{Q}}\rangle$  of the overall system  $\mathcal{R}\mathcal{Q}$  is pure; then the system  $\mathcal{Q}$  is sent through the channel which affects both the system  $\mathcal{Q}$  and the entanglement between  $\mathcal{Q}$  and  $\mathcal{R}$ .

operator-sum representation of the superoperator,  $\mathbb{S}(\rho_{\mathcal{Q}}) = \sum_k E_k \rho_{\mathcal{Q}} E_k^\dagger$ , with  $E_k$  Kraus operators, we obtain

$$f_e(\rho_{\mathcal{Q}}, \mathbb{S}) = \sum_k |\text{Tr}(\rho_{\mathcal{Q}} E_k)|^2. \quad (8.63)$$

This latter expression shows that entanglement fidelity is independent of the specific purification  $|\psi_{\mathcal{R}\mathcal{Q}}\rangle$ . Unlike the usual input-output fidelity, entanglement fidelity looks at the same physical process, the transmission of  $\mathcal{Q}$  across a channel, from a different point of view: as a local transformation on a part ( $\mathcal{Q}$ ) of an entangled system ( $\mathcal{R}\mathcal{Q}$ ). This transformation is undesired because it can reduce the amount of entanglement of the overall system.

As an illustrative example, we consider a *dephasing* channel (also referred to as phase-flip or phase damping channel):

$$\mathbb{S}(\rho_{\mathcal{Q}}) = \frac{1-g}{2} \sigma_z \rho_{\mathcal{Q}} \sigma_z + \frac{1+g}{2} \rho_{\mathcal{Q}}, \quad (8.64)$$

with the dephasing factor  $g \in [0, 1]$ . The qubit states before and after the channel transmission read as follows:

$$\rho_{\mathcal{Q}} = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \Rightarrow \rho'_{\mathcal{Q}} = \begin{pmatrix} \rho_{00} & g\rho_{01} \\ g\rho_{10} & \rho_{11} \end{pmatrix}. \quad (8.65)$$

If in particular we consider a completely dephasing channel,  $g = 0$ , and the case when a member of a Bell pair (say,  $|\psi_{\mathcal{R}\mathcal{Q}}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ ) is sent down the channel, then  $\rho_{\mathcal{Q}} = \frac{I}{2}$  and  $\rho'_{\mathcal{Q}} = \mathbb{S}(\rho_{\mathcal{Q}}) = \rho_{\mathcal{Q}}$ , so that the input-output fidelity  $f = 1$ . On the other hand,  $(\mathbb{I} \otimes \mathbb{S})(|\psi_{\mathcal{R}\mathcal{Q}}\rangle\langle\psi_{\mathcal{R}\mathcal{Q}}|) = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$ , and therefore the entanglement fidelity  $f_e = 0$ , as a consequence of the fact that entanglement is completely lost.

**Exercise 8.9** Compute the entanglement fidelity for the dephasing channel (8.64) and for a generic input state  $\rho_{\mathcal{Q}}$ .

### Entropy exchange

The entropy exchange is the entropy that the enlarged system  $\mathcal{R}\mathcal{Q}$  acquires when  $\mathcal{Q}$  undergoes the transformation  $\mathbb{S}$ :

$$S_e(\rho_{\mathcal{Q}}, \mathbb{S}) = S(\rho'_{\mathcal{R}\mathcal{Q}}), \quad (8.66)$$

where  $\rho'_{\mathcal{R}\mathcal{Q}}$  is given by (8.61). We will show below that the entropy exchange is an intrinsic function of the system input  $\rho_{\mathcal{Q}}$  and of the channel  $\mathbb{S}$  and does not depend on the particular purification.

Since  $\mathbb{S}$  is a completely positive trace preserving linear map, it can be represented by a unitary evolution  $U_{\mathcal{Q}\mathcal{E}}$  on a larger system, given by the system  $\mathcal{Q}$  itself plus an ancillary system  $\mathcal{E}$ , that is, a (generally fictitious) environment initially in a pure state  $|0_{\mathcal{E}}\rangle$ :

$$U_{\mathcal{Q}\mathcal{E}} |\phi_{\mathcal{Q}}\rangle |0_{\mathcal{E}}\rangle = \sum_k E_k |\phi_{\mathcal{Q}}\rangle |k_{\mathcal{E}}\rangle, \quad (8.67)$$

where  $|\phi_{\mathcal{Q}}\rangle$  is any pure state for the system  $\mathcal{Q}$  and  $\{|k_{\mathcal{E}}\rangle\}$  is an orthonormal basis for the environment. The final state of the environment is given by

$$\rho'_{\mathcal{E}} = \sum_{k,l} \text{Tr} \left( E_k \rho_{\mathcal{Q}} E_l^\dagger \right) |k\rangle \langle l|. \quad (8.68)$$

Since the overall final state  $\mathcal{R}\mathcal{Q}\mathcal{E}$  is pure, we have  $S(\rho'_{\mathcal{R}\mathcal{Q}}) = S(\rho'_{\mathcal{E}})$ . Therefore

$$S_e(\rho_{\mathcal{Q}}, \mathbb{S}) = S(\rho'_{\mathcal{E}}) \quad (8.69)$$

and  $S_e$  measures the entropy increase of the environment  $\mathcal{E}$  or, equivalently, the entanglement between  $\mathcal{R}\mathcal{Q}$  and  $\mathcal{E}$  after the evolution  $U_{\mathcal{Q}\mathcal{E}}$ . Using (8.68), we obtain

$$S_e(\rho_{\mathcal{Q}}, \mathbb{S}) = S(W) = -\text{Tr}(W \log W), \quad (8.70)$$

where  $W$  is a matrix with matrix elements  $W_{ij} = \text{Tr} \left( E_i \rho_{\mathcal{Q}} E_j^\dagger \right)$ . It is clear from Eq. (8.70) that  $S_e$  does not depend on the particular purification of the state  $\rho_{\mathcal{Q}}$ .

It is intuitive that, when the entropy exchange is large, then the entanglement fidelity must be small. Such expectation is confirmed by the *quantum Fano inequality* (for a proof see Barnum *et al.*, 1998):

$$S_e(\rho_{\mathcal{Q}}, \mathbb{S}) \leq H_{\text{bin}}(f_e(\rho_{\mathcal{Q}}, \mathbb{S})) + (1 - f_e(\rho_{\mathcal{Q}}, \mathbb{S})) \log(d^2 - 1), \quad (8.71)$$

where  $d$  is the dimension of the Hilbert space describing system  $\mathcal{Q}$ . The quantum Fano inequality implies that, if  $f_e \rightarrow 1$ , then  $S_e \rightarrow 0$ .

**Exercise 8.10** Compute the entropy exchange for the dephasing channel (8.64) and for a generic input state  $\rho_{\mathcal{Q}}$ .

### Coherent information

As we have discussed in Sec. 8.1.3, classical channel capacity is given by the maximum (over the input probability distribution) of the input-output mutual information. The quantity analogous to mutual information for quantum information is the coherent information  $\mathcal{I}_c$ , defined as

$$\mathcal{I}_c(\rho_Q, \mathbb{S}) = S(\mathbb{S}(\rho_Q)) - S_e(\rho_Q, \mathbb{S}). \quad (8.72)$$

Using Eqs. (8.66) and (8.69), we can also write

$$\mathcal{I}_c(\rho_Q, \mathbb{S}) = S(\rho'_Q) - S(\rho'_{RQ}) = S(\rho'_Q) - S(\rho'_E). \quad (8.73)$$

From the first equality in this equation, we can conclude that a quantity analogous to coherent information for classical systems can never be positive since the entropy of the joint system  $RQ$  cannot be smaller than the entropy of subsystem  $Q$ . On the other hand, in quantum mechanics coherent information can be either negative or positive and a positive value is a signature of the nonclassicality of the final state  $\rho'_{RQ}$ .

Note that the coherent information is maximal when  $Q$  and  $R$  are maximally entangled and the channel is noiseless. Indeed, in this case  $S(\rho'_Q) = S(\rho_Q)$  is maximal because the input state  $\rho_Q$  is maximally mixed, and  $S(\rho'_{RQ}) = 0$ , since the state  $|\psi_{RQ}\rangle$  remains pure after the transmission. Smaller values of the coherent information are obtained when the state  $|\psi_{RQ}\rangle$  is not maximally entangled or noise affects the channel. This example illustrates the fact that coherent information measures the possibility to convey entanglement through a communication channel.

**Exercise 8.11** Compute the coherent information for the dephasing channel (8.64) and for a generic input state  $\rho_Q$ .

**Theorem 8.6** Quantum capacity: *Given a quantum channel described by a superoperator  $\mathbb{S}$ , the quantum capacity  $Q$  of the channel is given by the maximum of the coherent information over all input states:*

$$Q(\mathbb{S}) = \lim_{n \rightarrow \infty} \frac{Q_n(\mathbb{S}^{\otimes n})}{n}, \quad Q_n(\mathbb{S}^{\otimes n}) = \max_{\rho^n} \mathcal{I}_c(\mathbb{S}^{\otimes n}, \rho^n), \quad (8.74)$$

where  $\rho^n$  is an input state for  $n$  channel uses.

The regularization  $n \rightarrow \infty$  is needed since the coherent information in general is not subadditive.

### Quantum data processing inequality

The similarity between the role played in classical information by the mutual information  $\mathcal{I}$  and in quantum information by the coherent information  $\mathcal{I}_c$  is highlighted by the fact that both quantities fulfill the data processing inequality.

In classical information theory, if the random variable  $X$  characterizes the input message,  $Y$  the output message, and  $Z$  the result of some processing of the output,

then it is possible to prove (see Cover and Thomas, 1991) the data processing inequality

$$H(X) \geq \mathcal{I}(X:Y) \geq \mathcal{I}(X:Z). \quad (8.75)$$

Such equality shows that if some information was lost (if  $\mathcal{I}(X:Y) < H(X)$ ), it is not possible to recover it by any manipulation of the output data.

An analogous result, known as quantum data processing inequality (Schumacher and Nielsen, 1996), holds for the coherent information:

$$S(\rho_Q) \geq \mathcal{I}_c(\rho_Q, \mathbb{S}_1) \geq \mathcal{I}_c(\rho_Q, \mathbb{S}_2 \circ \mathbb{S}_1). \quad (8.76)$$

To prove this result, we use a reference system  $\mathcal{R}$  for the purification of  $\rho_Q$  and two environments  $\mathcal{E}_1$  and  $\mathcal{E}_2$ , associated with the quantum operations  $\mathbb{S}_1$  and  $\mathbb{S}_2$ , respectively. The proof of the first inequality in Eq. (8.76) follows from the subadditivity of the von Neumann entropy,  $S(\rho'_{\mathcal{R}\mathcal{E}}) \leq S(\rho'_R) + S(\rho'_{\mathcal{E}})$ . Hence,

$$\begin{aligned} \mathcal{I}_c(\rho_Q, \mathbb{S}_1) &= S(\rho'_Q) - S(\rho'_{\mathcal{E}}) = S(\rho'_{\mathcal{R}\mathcal{E}}) - S(\rho'_{\mathcal{E}}) \\ &\leq S(\rho'_R) = S(\rho_R) = S(\rho_Q). \end{aligned} \quad (8.77)$$

The second inequality in Eq. (8.76) can be proved using the strong subadditivity of the von Neumann entropy to the compound system  $\mathcal{R}\mathcal{E}_1\mathcal{E}_2$ , after both maps  $\mathbb{S}_1$  and  $\mathbb{S}_2$  have taken place:

$$S(\rho''_{\mathcal{R}\mathcal{E}_1\mathcal{E}_2}) + S(\rho''_{\mathcal{E}_1}) \leq S(\rho''_{\mathcal{R}\mathcal{E}_1}) + S(\rho''_{\mathcal{E}_1\mathcal{E}_2}). \quad (8.78)$$

The purity of the overall state of  $\mathcal{R}\mathcal{Q}\mathcal{E}_1\mathcal{E}_2$  implies that  $S(\rho''_{\mathcal{R}\mathcal{E}_1\mathcal{E}_2}) = S(\rho''_Q)$ . Since quantum operation  $\mathbb{S}_2$  does not affect the systems  $\mathcal{R}$  and  $\mathcal{E}_1$ , we have  $S(\rho''_{\mathcal{E}_1}) = S(\rho'_{\mathcal{E}_1})$  and  $S(\rho''_{\mathcal{R}\mathcal{E}_1}) = S(\rho'_{\mathcal{R}\mathcal{E}_1}) = S(\rho'_Q)$ , where the last equality follows from the purity of system  $\mathcal{R}\mathcal{Q}\mathcal{E}_1$  after the action of quantum operation  $\mathbb{S}_1$ . After inserting the above equalities into Eq. (8.78) and taking into account that  $S(\rho'_{\mathcal{E}_1}) = S_e(\rho_Q, \mathbb{S}_1)$  and  $S(\rho''_{\mathcal{E}_1\mathcal{E}_2}) = S_e(\rho_Q, \mathbb{S}_2 \circ \mathbb{S}_1)$ , we obtain the second inequality in the quantum data processing inequality (8.76).

### Degradable channels

Though coherent information can be superadditive, there exist channels for which it is additive and therefore quantum channel capacity can be computed via the “single-letter” formula, namely  $Q = Q_1$  and the regularization  $n \rightarrow \infty$  of Eq. (8.74) is not needed. This is the case for the class of degradable channels.

To understand what a degradable channel is, we first use the unitary representation of a quantum channel  $\mathbb{S}$ ,

$$\rho'_Q = \mathbb{S}(\rho_Q) = \text{Tr}_{\mathcal{E}} \left[ U_{\mathcal{Q}\mathcal{E}} (\rho_Q \otimes |0_{\mathcal{E}}\rangle\langle 0_{\mathcal{E}}|) U_{\mathcal{Q}\mathcal{E}}^\dagger \right], \quad (8.79)$$

to define the *complementary channel*  $\mathbb{S}_c$ ,

$$\rho'_{\mathcal{E}} = \mathbb{S}_c(\rho_Q) = \text{Tr}_{\mathcal{Q}} \left[ U_{\mathcal{Q}\mathcal{E}} (\rho_Q \otimes |0_{\mathcal{E}}\rangle\langle 0_{\mathcal{E}}|) U_{\mathcal{Q}\mathcal{E}}^\dagger \right]. \quad (8.80)$$

A quantum channel  $\mathbb{S}$  is called degradable when it may be “degraded” to its complementary channel  $\mathbb{S}_c$ , that is, when it exists a map  $\mathbb{T}$ ,

$$\rho''_{\mathcal{E}} = \mathbb{T}(\rho'_{\mathcal{Q}}) = \text{Tr}_{\mathcal{Q}} \left[ V_{\mathcal{Q}\mathcal{E}} (\rho'_{\mathcal{Q}} \otimes |0_{\mathcal{E}}\rangle\langle 0_{\mathcal{E}}|) V_{\mathcal{Q}\mathcal{E}}^\dagger \right], \quad (8.81)$$

such that  $\mathbb{S}_c = \mathbb{T} \circ \mathbb{S}$ , namely  $\rho''_{\mathcal{E}} = \rho'_{\mathcal{E}}$ . Physically, degradability means that one can reconstruct the final state  $\rho'_{\mathcal{E}}$  of the environment from the final state  $\rho'_{\mathcal{Q}}$  of the system, see Fig. 8.8 for a graphical representation of this concept.

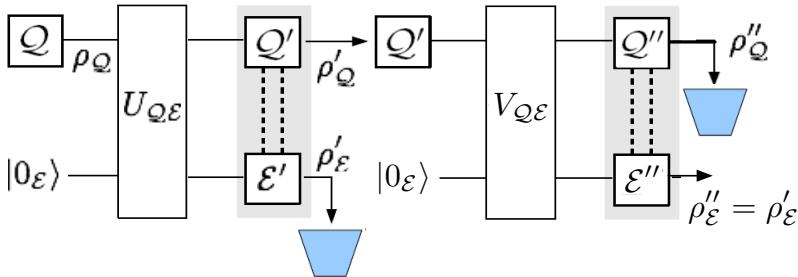


Fig. 8.8 Schematic quantum circuit illustrating degradability.

For degradable channels,

$$S(\rho'_{\mathcal{Q}}) = S(\rho'_{\mathcal{Q}} \otimes |0_{\mathcal{E}}\rangle\langle 0_{\mathcal{E}}|) = S(\rho''_{\mathcal{Q}\mathcal{E}}), \quad (8.82)$$

where the first equality follows from the fact that  $|0_{\mathcal{E}}\rangle\langle 0_{\mathcal{E}}|$  is a pure state and the latter is a consequence of the unitarity of the transformation  $V_{\mathcal{Q}\mathcal{E}}$  for the overall system  $\mathcal{Q}\mathcal{E}$ . Therefore the coherent information

$$I_c(\rho_{\mathcal{Q}}, \mathbb{S}) = S(\rho'_{\mathcal{Q}}) - S(\rho'_{\mathcal{E}}) = S(\rho''_{\mathcal{Q}\mathcal{E}}) - S(\rho''_{\mathcal{E}}) = S(\rho''_{\mathcal{Q}}|\rho''_{\mathcal{E}}), \quad (8.83)$$

where the *conditional von Neumann entropy*  $S(\rho''_{\mathcal{Q}}|\rho''_{\mathcal{E}})$  is known to be subadditive (see Nielsen and Chuang, 2000). We can conclude that the quantum capacity for degradable channels can be computed via the single-letter formula,  $Q = Q_1$ .

Concrete examples of degradable channels are the generalized dephasing channel, which we shall examine in some detail in what follows, and the amplitude damping channel. The unitary representation of the generalized dephasing channel reads

$$U_{\mathcal{Q}\mathcal{E}}|j\rangle|0_{\mathcal{E}}\rangle = |j\rangle|\phi_j\rangle, \quad (8.84)$$

where  $\{|j\rangle\}$  is an orthonormal basis (reference basis) for a  $d$ -dimensional system  $\mathcal{Q}$  and  $|\phi_j\rangle$  are environment states, in general non mutually orthogonal. Map  $\mathbb{S}$  can be written in the Kraus representation as

$$\rho'_{\mathcal{Q}} = \mathbb{S}(\rho_{\mathcal{Q}}) = \sum_{\alpha} E_{\alpha} \rho_{\mathcal{Q}} E_{\alpha}^{\dagger}, \quad (8.85)$$

where the system operators  $(E_\alpha)_{jl} = \langle \alpha_\varepsilon | \phi_j \rangle \delta_{jl}$  are diagonal in the reference basis (here  $\{|\alpha_\varepsilon\rangle\}$  is an orthonormal basis for the environment). For a generic input  $\rho_Q = \sum_{j,l} \rho_{jl} |j\rangle \langle l|$ , equation (8.84) yields

$$\rho'_\varepsilon = \mathbb{S}_c(\rho_Q) = \sum_j \rho_{jj} |\phi_j\rangle \langle \phi_j|. \quad (8.86)$$

Since  $\rho'_\varepsilon$  only depends on the populations  $\rho_{jj}$ , which are conserved by the dephasing channel  $\mathbb{S}$ , we can write as well  $\mathbb{S}_c = \mathbb{S}_c \circ \mathbb{S}$ , thus proving degradability.

The dephasing channel (8.64) for a qubit ( $d = 2$ ) is recovered by setting  $|\phi_0\rangle = |0_\varepsilon\rangle$  and  $|\phi_1\rangle = g|0_\varepsilon\rangle + \sqrt{1-g^2}|1_\varepsilon\rangle$ , with  $0 \leq g \leq 1$ . In this case the coherent information for a single channel use is maximized by the input state  $\rho_Q = \frac{1}{2} I$  (see exercise 8.11) and  $Q = Q_1 = 1 - H_{\text{bin}}\left(\frac{1+g}{2}\right)$ .

## 8.5 \* Quantum memory channels

It is interesting to consider the transmission of information through quantum channels with memory; that is, channels in which correlated noise acts on consecutive uses. This situation occurs in real physical quantum channels, provided the noise is correlated on a time scale larger than the time separation between consecutive uses of the channel. In quantum computation, time correlated noise is important in situations, such as solid state qubits, in which noise has components at frequencies much smaller than the time scales of interest for the system dynamics (we say that in this case the environment is non-Markovian).

In quantum memory channels, the map  $\mathbb{S}_n$  describing  $n$  channel uses is different from  $\mathbb{S}^{\otimes n}$ , where  $\mathbb{S}$  describes a single use of the channel. In this case the regularized coherent and Holevo information in general only provide upper bounds on the channel capacities. However, for the class of *forgetful channels*, for which memory effects decay exponentially with time, a quantum coding theorem showing that this bounds can be saturated exists (see Kretschmann and Werner, 2005).

In this section, we illustrate by means of two relevant examples that memory effects can be useful to enhance the transmission of both classical and quantum information.

### Entanglement-enhanced transmission of classical information

We consider the case of two consecutive uses of a channel with partial memory, following a model introduced by Macchiavello and Palma (2002). Each use of the channel corresponds to the transmission of a qubit and the action of the channel is described by the Kraus operators  $E_k$ , satisfying  $\sum_k E_k^\dagger E_k = I$ . In particular, we assume that

$$E_k = \sqrt{p_k} \sigma_k, \quad (8.87)$$

with  $i = 0, x, y, z$ ,  $\sigma_0 = I$  and  $\sum_k p_k = 1$ . If the state  $\rho$  is sent by Alice through the channel, then Bob receives the state

$$\rho' = \sum_k E_k \rho E_k^\dagger = p_0 \rho + p_x \sigma_x \rho \sigma_x + p_y \sigma_y \rho \sigma_y + p_z \sigma_z \rho \sigma_z. \quad (8.88)$$

Noise has therefore induced a rotation through an angle  $\pi$  about axis  $x, y, z$  of the Bloch sphere with probability  $p_x, p_y, p_z$  or left the state unchanged, with probability  $p_0$ . In the case of two uses of the channel, we assume that the initial two-qubit density matrix  $\rho$  is mapped onto

$$\rho' = \sum_{k_1, k_2} E_{k_1 k_2} \rho E_{k_1 k_2}^\dagger, \quad (8.89)$$

where the Kraus operators have the form

$$E_{k_1 k_2} = \sqrt{p_{k_1 k_2}} \sigma_{k_1} \sigma_{k_2}, \quad (8.90)$$

with  $\sum_{k_1, k_2} E_{k_1 k_2}^\dagger E_{k_1 k_2} = I \otimes I$ , implying  $\sum_{k_1, k_2} p_{k_1 k_2} = 1$ . The two limiting cases of memoryless and perfectly correlated channels are described by

$$E_{k_1 k_2}^{(u)} = \sqrt{p_{k_1}} \sqrt{p_{k_2}} \sigma_{k_1} \sigma_{k_2} \quad (8.91a)$$

and

$$E_{k_1 k_2}^{(c)} = \sqrt{p_{k_1}} \sigma_{k_1} \sigma_{k_2} \delta_{k_1 k_2}, \quad (8.91b)$$

respectively. An intermediate case is described by the Kraus operators

$$E_{k_1 k_2}^{(i)} = \sqrt{p_{k_1}[(1-\mu)p_{k_2} + \mu\delta_{k_1 k_2}]} \sigma_{k_1} \sigma_{k_2}. \quad (8.92)$$

This corresponds to  $p_{k_1 k_2} = p_{k_1} p_{k_2|k_1}$  (*Markov chain* model), with  $p_{k_2|k_1} = (1 - \mu)p_{k_2} + \mu\delta_{k_1 k_2}$ . This means that the channels has partial memory: with probability  $\mu$  the same rotation is applied to both qubits, whereas with probability  $1 - \mu$  the two rotations are uncorrelated. Hence, the parameter  $\mu$  describes the degree of correlation of the channel.

In the following we shall consider the depolarizing channel ( $p_0 = 1 - p$ ,  $p_x = p_y = p_z = \frac{p}{3}$ ) and assume that Alice sends Bob pure orthogonal quantum states drawn with equal *a priori* probabilities  $\pi_i = \frac{1}{4}$  ( $i = 1, \dots, 4$ ) from the ensemble  $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, |\psi_4\rangle\}$ , where

$$\begin{aligned} |\psi_1\rangle &= \cos \theta |00\rangle + \sin \theta |11\rangle, & |\psi_2\rangle &= \sin \theta |00\rangle - \cos \theta |11\rangle, \\ |\psi_3\rangle &= \cos \theta |01\rangle + \sin \theta |10\rangle, & |\psi_4\rangle &= \sin \theta |01\rangle - \cos \theta |10\rangle. \end{aligned} \quad (8.93)$$

Note that these states range from separable ( $\theta = 0$ ) to maximally entangled ( $\theta = \frac{\pi}{4}$ ). We shall maximize, as a function of  $\theta$ , the Holevo information (see Sec. 8.3.1)  $\chi = S(\rho') - \sum_i \pi_i S(\rho'_i)$ , where  $\rho' = \sum_i \pi_i \rho'_i$ ,  $\rho'_i = \sum_{k_1, k_2} E_{k_1 k_2} \rho_i E_{k_1 k_2}^\dagger$  and  $\rho_i = |\psi_i\rangle \langle \psi_i|$ . We shall show that there exists a memory threshold  $\mu_t$  above which the Holevo information is maximal when maximally entangled (Bell) states are transmitted. This demonstrates that the transmission of classical information may be enhanced by sending entangled states.

For this purpose, it is useful to write the input two-qubit state in the Fano representation (see Sec. 7.3.1) as follows:

$$\rho_i = \frac{1}{4} \left( I \otimes I + I \otimes \sum_k \alpha_k^{(i)} \sigma_k + \sum_k \beta_k^{(i)} \sigma_k \otimes I + \sum_{kl} \gamma_{kl}^{(i)} \sigma_k \otimes \sigma_l \right), \quad (8.94)$$

where  $\alpha_k^{(i)} = \text{Tr}[\rho_i(I \otimes \sigma_k)]$ ,  $\beta_k^{(i)} = \text{Tr}[\rho_i(\sigma_k \otimes I)]$  and  $\gamma_{kl}^{(i)} = \text{Tr}[\rho_i(\sigma_k \otimes \sigma_l)]$ . The input state  $\rho_1$  reads

$$\rho_1 = \frac{1}{4}[I \otimes I + \cos 2\theta(I \otimes \sigma_z + \sigma_z \otimes I) + \sigma_z \otimes \sigma_z + \sin 2\theta(\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y)]. \quad (8.95)$$

It can be checked by direct computation that

$$\begin{aligned} \sum_{k_1, k_2} E_{k_1 k_2}^{(u)} (I \otimes I) E_{k_1 k_2}^{(u)\dagger} &= I \otimes I, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(u)} (I \otimes \sigma_i) E_{k_1 k_2}^{(u)\dagger} &= \eta I \otimes \sigma_i, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(u)} (\sigma_i \otimes I) E_{k_1 k_2}^{(u)\dagger} &= \eta \sigma_i \otimes I, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(u)} (\sigma_i \otimes \sigma_j) E_{k_1 k_2}^{(u)\dagger} &= \eta^2 \sigma_i \otimes \sigma_j, \end{aligned} \quad (8.96)$$

where  $\eta = 1 - \frac{4}{3}p$  is the so-called shrinking factor (see Eq. (7.61)). We also obtain

$$\begin{aligned} \sum_{k_1, k_2} E_{k_1 k_2}^{(c)} (I \otimes I) E_{k_1 k_2}^{(c)\dagger} &= I \otimes I, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(c)} (I \otimes \sigma_i) E_{k_1 k_2}^{(c)\dagger} &= \eta I \otimes \sigma_i, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(c)} (\sigma_i \otimes I) E_{k_1 k_2}^{(c)\dagger} &= \eta \sigma_i \otimes I, \\ \sum_{k_1, k_2} E_{k_1 k_2}^{(c)} (\sigma_i \otimes \sigma_j) E_{k_1 k_2}^{(c)\dagger} &= \delta_{ij} \sigma_i \otimes \sigma_j + (1 - \delta_{ij}) \eta \sigma_i \otimes \sigma_j. \end{aligned} \quad (8.97)$$

Taking into account (8.96) and (8.97), we can see that the state  $\rho_1$  is transformed by the depolarizing channel with partial memory (8.92) into the output state

$$\begin{aligned} \rho'_1 &= \frac{1}{4} \left\{ I \otimes I + \eta \cos 2\theta(I \otimes \sigma_z + \sigma_z \otimes I) \right. \\ &\quad \left. + [\mu + (1 - \mu)\eta^2] [\sigma_z \otimes \sigma_z + \sin 2\theta(\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y)] \right\}. \end{aligned} \quad (8.98)$$

The eigenvalues of this density matrix are

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{4}(1 - \mu)(1 - \eta^2), \\ \lambda_{3,4} &= \frac{1}{4} \left\{ 1 + \mu + \eta^2(1 - \mu) \pm 2\sqrt{\eta^2 \cos^2(2\theta) + [\eta^2(1 - \mu) + \mu]^2 \sin^2(2\theta)} \right\}. \end{aligned} \quad (8.99)$$

The same eigenvalues are obtained for the other output states  $\rho'_2$ ,  $\rho'_3$ ,  $\rho'_4$ . As  $\lambda_1$  and  $\lambda_2$  do not depend on  $\theta$ , the von Neumann entropy  $S(\rho_i)$  ( $i = 1, \dots, 4$ ) is minimized (as a function of  $\theta$ ) when the term under the square root in the expression for  $\lambda_3$  and  $\lambda_4$  is maximum. Moreover, we have  $\rho' = \frac{1}{4}(\rho'_1 + \rho'_2 + \rho'_3 + \rho'_4) = \frac{1}{4}(I \otimes I)$ , so that  $S(\rho') = 2$ . Therefore, the Holevo information  $\chi = S(\rho') - \frac{1}{4} \sum_i S(\rho'_i) = 2 - S(\rho'_i)$  is maximal for separable states ( $\theta = 0$ ) when  $\eta^2 > [\eta^2(1 - \mu) + \mu]^2$  and for Bell

states ( $\theta = \frac{\pi}{4}$ ) when  $\eta^2 < [\eta^2(1 - \mu) + \mu]^2$ . This latter condition can be equivalently written as

$$\mu > \mu_t = \frac{\eta}{1 + \eta}. \quad (8.100)$$

Therefore, for states of the form (8.93) the Holevo information is maximal for separable states when  $\mu < \mu_t$  and for Bell states when  $\mu > \mu_t$ . At the threshold value  $\mu = \mu_t$ , the same Holevo information is obtained for any value of  $\theta$  in (8.93). The Holevo information is shown in Fig. 8.9 as a function of  $\theta$ , for different values of the parameter  $\mu$ . The different behaviour below and above the threshold  $\mu_t$  is evident. Note that, for a perfectly correlated noise channel ( $\mu = 1$ ) we have  $\xi = 2$  for Bell states. Indeed, in this case noise does not affect the Bell states:  $\rho'_i = \rho_i$ , and therefore  $S(\rho'_i) = S(\rho_i) = 0$ .

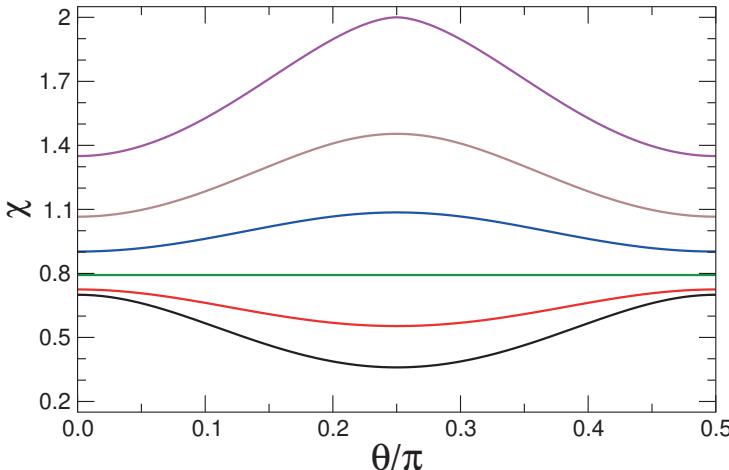


Fig. 8.9 Holevo information  $\chi$  as a function of the parameter  $\theta$ , for  $\eta = \frac{2}{3}$  and, from bottom to top,  $\mu = 0$  (black), 0.2 (red), 0.4 (green), 0.6 (blue), 0.8 (brown) and 1 (magenta). The function  $\chi(\theta)$  has periodicity  $\pi/2$ . At  $\mu = \mu_t = 0.4$ ,  $\chi$  is independent of  $\theta$ .

#### Coding-decoding scheme taking advantage of channel memory

We now consider a Hamiltonian model of a quantum channel. We suppose that information is carried by qubits that transit across a communication channel, modeled as a purely dephasing environment. The Hamiltonian describing the transmission of  $n$  qubits through the channel reads

$$H(t) = H_{\mathcal{E}} - \frac{1}{2} X_{\mathcal{E}} F(t), \quad F(t) = \lambda \sum_{k=1}^n \sigma_k^z f_k(t), \quad (8.101)$$

where  $H_{\mathcal{E}}$  is the environment Hamiltonian and  $X_{\mathcal{E}}$  the environment coupling operator. The  $k$ -th qubit is coupled to the environment via its Pauli operator  $\sigma_k^z$ , the coupling strength being  $\lambda$ . The functions  $f_k(t)$  switch on and off the coupling:

$f_k(t) = 1$  when the  $k$ -th qubit is inside the channel,  $f_k(t) = 0$  otherwise. We call  $\tau_p$  the time each carrier takes to cross the channel and  $\tau$  the time interval that separates two consecutive qubits entering the channel. Note that Hamiltonian (8.101) is expressed in the interaction picture with respect to the qubits.

We call  $w_0$  and  $\rho_Q$  the density operators which represent the initial states of the environment and of the  $n$  qubits, respectively. Assuming that initially the system and the environment are not entangled, we can write the state of the system at time  $t$  as follows:

$$\rho_Q(t) = \text{Tr}_{\mathcal{E}}\{U(t)(\rho_Q \otimes w_0)U^\dagger(t)\}, \quad (8.102)$$

$$U(t) = \mathcal{T}e^{-\frac{i}{\hbar} \int_0^t ds H(s)}, \quad (8.103)$$

where  $\mathcal{T}$  denotes the time-ordering operator. In particular, we are interested in the final state  $\rho'_Q$  after all  $n$  qubits crossed the channels. To treat this problem we choose the factorized basis states  $\{|j\alpha_{\mathcal{E}}\rangle\}$ , where  $\{|j\rangle = |j_1, \dots, j_n\rangle\}$  are the eigenvectors of  $\prod_k \sigma_k^z$ , and  $\{|\alpha_{\mathcal{E}}\rangle\}$  is an orthonormal basis for the environment. The dynamics preserves the basis states  $|j\rangle$  and therefore the evolution operator (8.103) is diagonal in the system indices:

$$\langle j|\alpha_{\mathcal{E}}|U(t)|l\alpha'_{\mathcal{E}}\rangle = \delta_{jl} \langle \alpha_{\mathcal{E}}|U(t|j)|\alpha'_{\mathcal{E}}\rangle, \quad (8.104)$$

where  $U(t|j) = \langle j|U(t)|j\rangle$  expresses the conditional evolution operator of the environment alone. Therefore,

$$(\rho'_Q)_{jl} = (\rho_Q)_{jl} \sum_{\alpha} \langle \alpha_{\mathcal{E}}|U(t|j)w_0 U^\dagger(t|l)|\alpha_{\mathcal{E}}\rangle. \quad (8.105)$$

In this (preferential) basis, the populations are preserved and the environment only changes the off-diagonal elements of  $\rho_Q$ .

Now we model the environment as a bosonic bath with an infinite set of oscillators:

$$H_{\mathcal{E}} = \sum_{\alpha} \omega_{\alpha} b_{\alpha}^{\dagger} b_{\alpha} + H_C, \quad H_C = \sum_{\alpha} \frac{\lambda^2}{4\omega_{\alpha}} \sum_{k=1}^n \sigma_k^z, \quad X_{\mathcal{E}} = \sum_{\alpha} (b_{\alpha}^{\dagger} + b_{\alpha}), \quad (8.106)$$

where  $H_C$  is a counterterm (for a discussion on the meaning of this term, see Weiss, 2012). If the environment is initially in thermal equilibrium,  $w_0 = e^{-\beta H_{\mathcal{E}}}$ , we obtain

$$\begin{aligned} & \sum_{\alpha} \langle \alpha_{\mathcal{E}}|U(t|j)w_0 U^\dagger(t|l)|\alpha_{\mathcal{E}}\rangle \\ &= \exp \left[ -\lambda^2 \int_0^{\infty} \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega\tau_p)}{\omega^2} \left| \sum_{k=1}^n (j_k - l_k) e^{i\omega(k-1)\tau} \right|^2 \right], \end{aligned} \quad (8.107)$$

where  $S(\omega)$  is the *power spectrum* of the coupling operator  $X_{\mathcal{E}}$ , that is the Fourier transform of the bath symmetrized autocorrelation function:  $C(t) = 1/2 \langle X_{\mathcal{E}}(t)X_{\mathcal{E}}(0) + X_{\mathcal{E}}(0)X_{\mathcal{E}}(t) \rangle$ .

For a single channel use, we recover the single-qubit dephasing channel (8.64), with the dephasing factor  $g \in [0, 1]$  deduced from (8.107) for  $n = 1$ :

$$g = \exp \left\{ -\lambda^2 \int_0^\infty \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega\tau_p)}{\omega^2} \right\}. \quad (8.108)$$

We now consider two channel uses. Provided that the time  $\tau$  between two channel uses is smaller than the time scale  $\tau_c$  associated with the decay of environmental correlation functions, the action of the environment on the second qubit is related to the action on the first qubit. Therefore,  $\mathbb{S}_2 \neq \mathbb{S} \otimes \mathbb{S}$ , where the superoperator  $\mathbb{S}_2$  describes the transformation operated by the channel on the overall two-qubit system.

We consider the transmission of two qubits,  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ , initially in the state  $\rho_{\mathcal{Q}_1 \mathcal{Q}_2}$ , with matrix elements  $\rho_{mn}$ ,  $m, n = 0, \dots, 3$ . The final state of the system is

$$\rho'_{\mathcal{Q}_1 \mathcal{Q}_2} = \mathcal{E}_2(\rho_{\mathcal{Q}_1 \mathcal{Q}_2}) = \begin{pmatrix} \rho_{00} & g\rho_{01} & g\rho_{02} & h^+\rho_{03} \\ g\rho_{10} & \rho_{11} & h^-\rho_{12} & g\rho_{13} \\ g\rho_{20} & h^-\rho_{21} & \rho_{22} & g\rho_{23} \\ h^+\rho_{30} & g\rho_{31} & g\rho_{32} & \rho_{33} \end{pmatrix}, \quad (8.109)$$

where the factors  $g$  and  $h^\pm$  describe the channel effects and the noiseless limit is recovered for  $g = h^\pm = 1$ . The last two terms are defined as

$$h^\pm = \exp \left\{ -2\lambda^2 \int_0^\infty \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega\tau_p)}{\omega^2} (1 \pm \cos \omega\tau) \right\} \quad (8.110)$$

and are derived from (8.107) for  $j_1 = j_2 = j$ ,  $l_1 = l_2 \neq j$  ( $h^+$ ) and for  $j_1 = l_2 = j$ ,  $j_2 = l_1 \neq j$  ( $h^-$ ).

In the absence of any memory effects, that is, when the power spectrum  $S(\omega)$  is white and there is no superposition between the time windows when the first or the second qubits are inside the channel ( $\tau \geq \tau_p$ ), we have  $h^\pm = g^2$ ; therefore,  $\mathbb{S}_2 = \mathbb{S}^{\otimes 2}$ . In the opposite limiting case of perfect memory, that is, when  $\tau_c \gg \tau, \tau_p$ , or alternatively the two time windows of the qubit-environment interaction are completely superimposed ( $\tau = 0$ ), we have  $h^+ = g^4$  and  $h^- = 1$ . In this limit the subspace spanned by the basis  $\{|00\rangle, |11\rangle\}$  undergoes a stronger decoherence (with respect to the memoryless case), while the subspace spanned by  $\{|01\rangle, |10\rangle\}$  is *decoherence free*.

It is convenient to measure the memory between two channel uses by introducing the memory coefficient  $\gamma$  defined as follows:

$$\gamma = \int_0^\infty \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega\tau_p)}{\omega^2} \cos \omega\tau / \int_0^\infty \frac{d\omega}{\pi} S(\omega) \frac{1 - \cos(\omega\tau_p)}{\omega^2}. \quad (8.111)$$

For a given power spectrum  $S(\omega)$  and crossing time  $\tau_p$ ,  $\gamma$  only depends on the time interval  $\tau$ , and ranges in the interval  $[0, 1]$ . In particular,  $\gamma = 0$  for a memoryless channel (as it can be checked by letting  $\tau \rightarrow \infty$  in the (8.111)), while  $\gamma = 1$  for perfect memory ( $\tau = 0$  in (8.111)). We can express the dephasing factors  $h^\pm$  by means of the corresponding memoryless value  $g^2$  and the memory factor  $\gamma$ :

$$h^\pm = g^{2(1 \pm \gamma)}. \quad (8.112)$$

In what follows, we show that one can exploit memory effects to design suitable coding-decoding schemes that improve the faithfulness of quantum information transmission (see D'Arrigo *et al.*, 2008). Indeed, memory effects can be used to preserve entanglement in quantum information transmission. In particular, we consider an entanglement sharing protocol: Alice wishes to send one qubit of a Bell pair (qubits  $\mathcal{R}$  and  $\mathcal{Q}$ ) to Bob. The quantum channel (8.101), (8.106) randomizes the phase between the sent qubit ( $\mathcal{Q}$ ) and the reference one ( $\mathcal{R}$ ). In order to take advantage of memory effects, we follow the strategy sketched in Fig 8.10. We encode the sent qubit in a two-qubit system whose state resides in the subspace (spanned by  $\{|01\rangle, |10\rangle\}$ ) resilient to errors in the presence of memory effects.

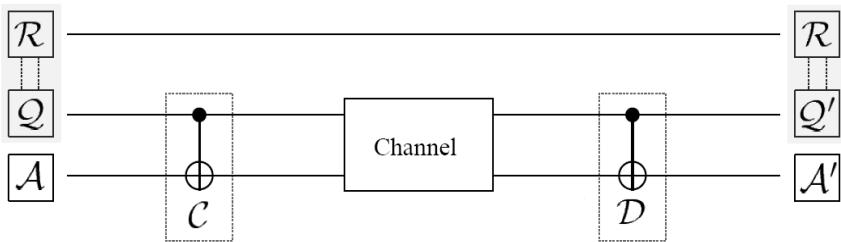


Fig. 8.10 A schematic drawing of a coding-decoding scheme taking advantage of the correlation between two channel uses.

We assume that the initial state of the entangled pair is a Bell pair, say  $|\psi_{\mathcal{R}\mathcal{Q}}\rangle = \frac{1}{2}(|00\rangle + |11\rangle) = |\phi^+\rangle$ . The coding-decoding protocol is performed in the following way:

- (i) We prepare an ancillary qubit  $\mathcal{A}$  in the state  $|1\rangle$ , such that the whole system  $\mathcal{R}\mathcal{Q}\mathcal{A}$  is initially in the state

$$|\psi_{\mathcal{R}\mathcal{Q}\mathcal{A}}\rangle = |\psi_{\mathcal{R}\mathcal{Q}}\rangle \otimes |1\rangle = \frac{1}{2}(|001\rangle + |111\rangle). \quad (8.113)$$

- (ii) The encoding operation  $\mathcal{C}$  is a controlled-not gate acting on the system  $\mathcal{Q}\mathcal{A}$ , where  $\mathcal{Q}$  is the control qubit:

$$|\tilde{\psi}_{\mathcal{R}\mathcal{Q}\mathcal{A}}\rangle = (I_{\mathcal{R}} \otimes \text{CNOT}_{\mathcal{Q}\mathcal{A}})(|\psi_{\mathcal{R}\mathcal{Q}\mathcal{A}}\rangle) = \frac{1}{2}(|001\rangle + |110\rangle). \quad (8.114)$$

As a result, we encode the system  $\mathcal{R}\mathcal{Q}\mathcal{A}$  into a GHZ state, in such a way that the subsystem  $\mathcal{Q}\mathcal{A}$  resides in the subspace spanned by  $\{|01\rangle, |10\rangle\}$ .

- (iii) We send qubits  $\mathcal{Q}$  and  $\mathcal{A}$  through the channel. We call  $\tilde{\rho}'_{\mathcal{R}\mathcal{Q}\mathcal{A}}$  the density operator describing the state that arises from the channel transmission:

$$\tilde{\rho}'_{\mathcal{R}\mathcal{Q}\mathcal{A}} = (I_{\mathcal{R}} \otimes (\mathbb{S}_2)_{\mathcal{Q}\mathcal{A}})(|\tilde{\psi}_{\mathcal{R}\mathcal{Q}\mathcal{A}}\rangle\langle\tilde{\psi}_{\mathcal{R}\mathcal{Q}\mathcal{A}}|). \quad (8.115)$$

- (iv) The decoding operation  $\mathcal{D}$  extracts the state of system  $\mathcal{R}\mathcal{Q}$  from the one of  $\mathcal{R}\mathcal{Q}\mathcal{A}$ . To this aim we apply another controlled-not gate to the system  $\mathcal{Q}\mathcal{A}$ ,

where  $\mathcal{Q}$  is again the control qubit. This operation disentangles systems  $\mathcal{R}\mathcal{Q}$  and  $\mathcal{A}$ :

$$\rho'_{\mathcal{R}\mathcal{Q}\mathcal{A}} = \left( I_{\mathcal{R}} \otimes \text{CNOT}_{\mathcal{Q}\mathcal{A}} \right) \tilde{\rho}'_{\mathcal{R}\mathcal{Q}\mathcal{A}} \left( I_{\mathcal{R}} \otimes \text{CNOT}_{\mathcal{Q}\mathcal{A}} \right) = \rho'_{\mathcal{R}\mathcal{Q}} \otimes |1\rangle, \quad (8.116)$$

where

$$\rho'_{\mathcal{R}\mathcal{Q}} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) + \frac{1}{2}h^-(|00\rangle\langle 11| + |11\rangle\langle 00|). \quad (8.117)$$

The fidelity between the initial and the final state is

$$f = \langle \phi_{\mathcal{R}\mathcal{Q}} | \rho'_{\mathcal{R}\mathcal{Q}} | \phi_{\mathcal{R}\mathcal{Q}} \rangle = \frac{1}{2}(1 + h^-). \quad (8.118)$$

This is also the entanglement fidelity  $f_e^c$  when the initial state of  $\mathcal{Q}$  is  $\rho_{\mathcal{Q}} = \frac{1}{2}I$  and when the above coding-encoding scheme is used. We compare this value with the entanglement fidelity  $f_e = \frac{1+g}{2}$ , obtained when the qubit  $\mathcal{Q}$  is simply sent down the channel (see exercise 8.9). Therefore, the above coding-encoding strategy is useful when memory effects are strong enough, namely when

$$f_e^c \geq f_e \Rightarrow h^- = g^{2(1-\gamma)} \geq g \Rightarrow \gamma \geq \frac{1}{2}. \quad (8.119)$$

While the quantum capacity of the quantum channel (8.101), (8.106) is not known analytically, numerical data for the coherent information for different number of channel uses suggest that memory effects enhance the quantum capacity with respect to the memoryless case (see D'Arrigo *et al.*, 2007).

## 8.6 A guide to the bibliography

Modern information theory started with the work of Shannon (1948); general references are Cover and Thomas (1991) and Gray (1990).

A complete and self-contained presentation of quantum information theory is Wilde (2013).

The quantum noiseless coding theorem is due to Schumacher (1995), see also Barnum *et al.* (1996).

A simplified derivation of the Holevo bound can be found in Fuchs and Caves (1994). The product-state capacity formula for communication of classical information over noisy quantum channels was derived by Holevo (1998) and Schumacher and Westmoreland (1997). Hastings (2009) showed that the Holevo information can be superadditive.

A thorough discussion of the concepts of entanglement fidelity, entropy exchange and coherent information can be found in Barnum *et al.* (1998). For a proof that coherent information is an achievable rate for quantum communication, see Devetak (2005). Degradable quantum channels were introduced by Devetak and Shor (2005), for a review see Cubitt *et al.* (1998).

A review on quantum memory channels is Caruso *et al.* (2014).

# Chapter 9

## Quantum error correction

In this chapter we discuss how to protect quantum information from errors. The use of error-correcting codes to fight the effect of noise is a well developed technique in classical information processing. The key ingredient to protect information against errors is *redundancy*.

To grasp this point, it is useful to consider the following example. Alice wishes to send Bob a classical bit through a classical communication channel; that is, a channel described by the laws of classical mechanics. The effect of noise in the channel is to flip the bit ( $0 \rightarrow 1$  or  $1 \rightarrow 0$ ) with probability  $\epsilon$  ( $0 \leq \epsilon \leq 1$ ), while the bit is transmitted without error with probability  $1 - \epsilon$ . The simplest manner to protect the bit is to send three copies of it: Alice sends 000 instead of just 0, say, or 111 instead of 1. Bob receives the three bits and applies *majority voting*: if, for instance, he receives 010, he assumes that, most probably, there was a single error affecting the second bit ( $0 \rightarrow 1$ ). He therefore concludes that the transmitted bit of information was 0.

We should point out that the underlying hypothesis is that the noisy channel is memoryless; namely, noise acts independently on each bit. Therefore, if Alice sends 000, Bob will receive 000 with probability  $(1 - \epsilon)^3$ . The error-correcting code succeeds if there is a single error; that is, when Bob receives 100, 010, or 001. Each of these messages is received with probability  $\epsilon(1 - \epsilon)^2$ . The code fails if two or more bits have been flipped. This is the case if Bob receives 011, 101, 110 (with probability  $\epsilon^2(1 - \epsilon)$ ), or 111 (with probability  $\epsilon^3$ ). Therefore, the failure probability of the code is  $\epsilon_c = 3\epsilon^2(1 - \epsilon) + \epsilon^3 = 3\epsilon^2 - 2\epsilon^3$ . For just a single bit, the error probability was  $\epsilon$ . Hence, the code improves the probability of successful transmission if  $\epsilon_c < \epsilon$ ; that is, if  $\epsilon < 1/2$ . The improvement is greater for  $\epsilon$  smaller since the error probability is reduced by a factor  $\approx 3\epsilon$ . For instance, for  $\epsilon = 10^{-1}$ ,  $\epsilon_c = 2.8 \times 10^{-2}$ , while, for  $\epsilon = 10^{-2}$ ,  $\epsilon_c = 2.98 \times 10^{-4}$ .

The application of the same redundancy principle to quantum information encounters difficulties directly related to the basic principles of quantum mechanics:

- (1) Owing to the no-cloning theorem (discussed in Sec. 5.2), it is impossible to make copies of an unknown quantum state. Therefore, we cannot mimic the above-

described classical code by sending  $|\psi\rangle|\psi\rangle|\psi\rangle$  to protect an unknown quantum state  $|\psi\rangle$ .

- (2) In order to operate classical error correction, we observe (measure) the output from the noisy channel. In quantum mechanics, we know that, in general, measurements disturb the quantum state under investigation. For instance, if we receive the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  and measure its polarization along the  $z$ -axis, the state will collapse onto  $|0\rangle$  (with probability  $|\alpha|^2$ ) or  $|1\rangle$  (with probability  $|\beta|^2$ ). In either case, the coherent superposition of the states  $|0\rangle$  and  $|1\rangle$  will be destroyed.
- (3) While the only possible classical error affecting a single bit is the bit flip ( $0 \rightarrow 1$  and  $1 \rightarrow 0$ ), the class of possible quantum errors is much richer. For instance, we can have the phase-flip error:  $\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|0\rangle - \beta|1\rangle$ . This error has no classical counterpart. Moreover, a continuum of quantum errors may occur in a single qubit. Given the state  $|\psi\rangle$ , noise may slightly rotate it:  $|\psi\rangle \rightarrow R|\psi\rangle$ , with  $R$  a rotation matrix. Such small errors will accumulate in time, eventually leading to incorrect computations (see Sec. 3.8). At first sight, it might thus appear that infinite resources are required to correct such errors since infinite precision is required to determine a rotation angle exactly.

However, we shall see in this chapter that, in spite of the above difficulties, quantum error correction is possible. We shall first discuss some simple examples: the three-qubit bit-flip and phase-flip codes, the nine-qubit Shor code and the five-qubit code. Then, on more general grounds, we shall discuss quantum codes using the stabilizer formalism. We will start by revisiting the nine-qubit Shor code in this new language, thus setting the ground in order to discuss stabilizers more rigorously in the context of group algebra.

In the remainder of the chapter we shall introduce passive error correction schemes, based on the concept of decoherence-free subspaces, the dynamical decoupling technique, and finally include a discussion of the quantum Zeno effect. Eventually, we shall discuss fault-tolerant quantum computation and show that, under certain hypotheses, if the noise level is below some threshold, then arbitrarily long, but reliable quantum computation is possible.

An important point to evaluate the effectiveness of a realistic error correction protocol, is that all possible error sources cannot be considered in practice. For instance, in general error sources cannot be considered stationary, and memory effects could be relevant. Since quantum computers are intrinsically analog devices, it is reasonable to assume that they need more or less frequent calibration depending on hardware stability. One last point to keep in mind is that quantum computers (beyond small demonstrative systems) require an extrapolation of the validity of the standard postulates of quantum mechanics well beyond the limits of validity so far verified experimentally. In this regard, a large-scale quantum computer would be by far the most sensitive test of the validity of these postulates.

## 9.1 The three-qubit bit-flip code

Let us assume that Alice wishes to send a qubit, prepared in a generic state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , to Bob via a noisy quantum channel. The following hypothesis is made: the noise acts on each qubit independently, leaving the state of the qubit unchanged (with probability  $1 - \epsilon$ ) or applying the Pauli operator  $\sigma_x$  (with probability  $\epsilon$ ). We remind the reader that  $\sigma_x$  produces a bit-flip error since  $\sigma_x|0\rangle = |1\rangle$  and  $\sigma_x|1\rangle = |0\rangle$ . To protect the quantum state  $|\psi\rangle$ , Alice employs the following *encoding*:

$$|0\rangle \rightarrow |0_L\rangle \equiv |000\rangle, \quad |1\rangle \rightarrow |1_L\rangle \equiv |111\rangle. \quad (9.1)$$

The subscript  $L$  indicates that the states  $|0_L\rangle$  and  $|1_L\rangle$  are the *logical*  $|0\rangle$  and  $|1\rangle$  states (also known as *codewords*), encoded by means of three physical qubits. Correspondingly, a generic state is encoded as follows:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|0_L\rangle + \beta|1_L\rangle = \alpha|000\rangle + \beta|111\rangle. \quad (9.2)$$

This encoding is implemented by means of the quantum circuit in Fig. 9.1: the first CNOT gate maps  $(\alpha|0\rangle + \beta|1\rangle)|00\rangle$  into  $(\alpha|00\rangle + \beta|11\rangle)|0\rangle$  and the second CNOT leads to the encoded state  $\alpha|000\rangle + \beta|111\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$ . This state is an entangled three-qubit GHZ state. We should point out that Alice's encoding does not violate the no-cloning theorem since the encoded state is not the same as three copies of the original unknown state:

$$\alpha|000\rangle + \beta|111\rangle \neq |\psi\rangle|\psi\rangle|\psi\rangle = (\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle). \quad (9.3)$$

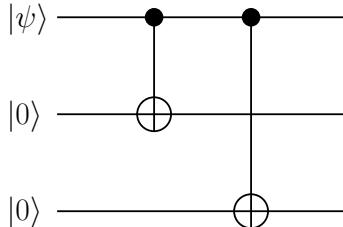


Fig. 9.1 A quantum circuit encoding a single qubit into three.

The three qubits, prepared in the cat state, are sent from Alice to Bob through the noisy channel. As a result, Bob receives one of the following states:

$$\begin{aligned}
 & \alpha|000\rangle + \beta|111\rangle, & (1 - \epsilon)^3, \\
 & \alpha|100\rangle + \beta|011\rangle, & \epsilon(1 - \epsilon)^2, \\
 & \alpha|010\rangle + \beta|101\rangle, & \epsilon(1 - \epsilon)^2, \\
 & \alpha|001\rangle + \beta|110\rangle, & \epsilon(1 - \epsilon)^2, \\
 & \alpha|110\rangle + \beta|001\rangle, & \epsilon^2(1 - \epsilon), \\
 & \alpha|101\rangle + \beta|010\rangle, & \epsilon^2(1 - \epsilon), \\
 & \alpha|011\rangle + \beta|100\rangle, & \epsilon^2(1 - \epsilon), \\
 & \alpha|111\rangle + \beta|000\rangle, & \epsilon^3,
 \end{aligned} \quad (9.4)$$

where in the right-hand column we have written the probabilities of receiving the different states.

In order to correct a single bit-flip error, Bob might be tempted to measure the polarizations  $\sigma_1^z$ ,  $\sigma_2^z$  and  $\sigma_3^z$  of the three qubits. To give a concrete example, let us assume that he receives the state  $\alpha|100\rangle + \beta|011\rangle$ . The three-qubit polarization measurement gives outcome 100 (with probability  $|\alpha|^2$ ) or 011 (with probability  $|\beta|^2$ ). In both case, Bob could apply majority voting and would conclude that the first qubit has been flipped. However, the coherent superposition of the states  $|0\rangle$  and  $|1\rangle$  would then be lost.

The problem may be solved by performing *collective measurements* on two qubits simultaneously. This can be achieved by means of the circuit in Fig. 9.2, which allows Bob to measure  $\sigma_1^z\sigma_2^z$  and  $\sigma_1^z\sigma_3^z$ . Bob employs two ancillary qubits, both prepared in the state  $|0\rangle$ . The first two CNOT gates and the measurement of the polarization  $x_0$  of the first ancillary qubit (by means of the detector  $D_0$ ) tell him the value of  $\sigma_1^z\sigma_2^z$ . Note that  $x_0 = 0$  corresponds to  $\sigma_1^z\sigma_2^z = 1$ , while  $x_0 = 1$  corresponds to  $\sigma_1^z\sigma_2^z = -1$ . In the same manner, the last two CNOT gates and the measurement of the second ancillary qubit provide him with the value of  $\sigma_1^z\sigma_3^z$  ( $x_1 = 0$  when  $\sigma_1^z\sigma_3^z = 1$  and  $x_1 = 1$  when  $\sigma_1^z\sigma_3^z = -1$ ).

As an example, we consider the case in which the first qubit has been flipped. The initial state of the five qubits is then

$$(\alpha|100\rangle + \beta|011\rangle)|00\rangle. \quad (9.5)$$

It is easy to check that the four CNOT gates map this state into

$$(\alpha|100\rangle + \beta|011\rangle)|11\rangle. \quad (9.6)$$

The measurement of the two ancillary qubits gives Bob two classical bits of information,  $x_0$  and  $x_1$ , known as the *error syndrome*, of value  $x_0 = 1$  and  $x_1 = 1$ . Since  $x_0 = 1$  Bob concludes that one of the first two qubits has been flipped. In the same manner, from  $x_1 = 1$  Bob concludes that either the first or the third qubit has been flipped. Put together, the information provided by the values of  $x_0$  and  $x_1$  leads Bob to conclude that the first qubit has been flipped. Therefore, he applies a NOT gate ( $\sigma_x$ ) to this qubit to recover the encoded state  $\alpha|000\rangle + \beta|111\rangle$ .

In general, the measured syndrome and the action taken by Bob are the following (see Fig. 9.3):

$$\begin{aligned} x_0 &= 0, x_1 = 0, && \text{no action,} \\ x_0 &= 0, x_1 = 1, && \text{apply NOT to the third qubit,} \\ x_0 &= 1, x_1 = 0, && \text{apply NOT to the second qubit,} \\ x_0 &= 1, x_1 = 1, && \text{apply NOT to the first qubit.} \end{aligned} \quad (9.7)$$

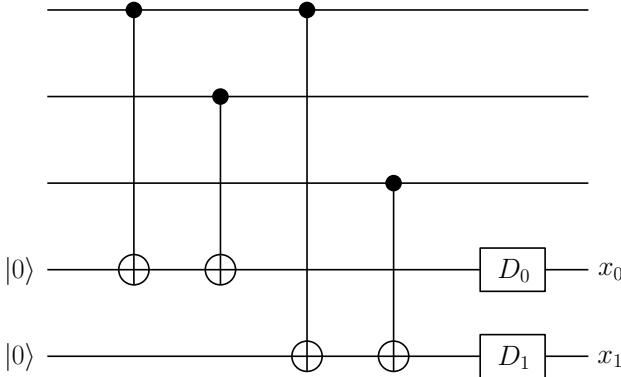


Fig. 9.2 A quantum circuit for extracting the error syndrome in the three-qubit bit-flip code.

After Bob's action, the five-qubit states and their probabilities will be given by

$$\begin{aligned}
 & (\alpha|000\rangle + \beta|111\rangle)|00\rangle, & (1-\epsilon)^3, \\
 & (\alpha|000\rangle + \beta|111\rangle)|11\rangle, & \epsilon(1-\epsilon)^2, \\
 & (\alpha|000\rangle + \beta|111\rangle)|10\rangle, & \epsilon(1-\epsilon)^2, \\
 & (\alpha|000\rangle + \beta|111\rangle)|01\rangle, & \epsilon(1-\epsilon)^2, \\
 & (\alpha|111\rangle + \beta|000\rangle)|01\rangle, & \epsilon^2(1-\epsilon), \\
 & (\alpha|111\rangle + \beta|000\rangle)|10\rangle, & \epsilon^2(1-\epsilon), \\
 & (\alpha|111\rangle + \beta|000\rangle)|11\rangle, & \epsilon^2(1-\epsilon), \\
 & (\alpha|111\rangle + \beta|000\rangle)|00\rangle, & \epsilon^3.
 \end{aligned} \tag{9.8}$$

From now on, we may neglect the ancillary qubits. Finally, to extract the qubit sent by Alice, Bob applies the inverse of the encoding procedure. This *decoding* is shown in Fig. 9.3 and leads the three qubits sent by Alice to the state  $(\alpha|0\rangle + \beta|1\rangle)|00\rangle$  (for the first four states in Eq. (9.8)) or to  $(\alpha|1\rangle + \beta|0\rangle)|00\rangle$  (for the last four states in Eq. (9.8)). Hence, the three-qubit bit-flip code is successful if no more than one qubit has been flipped. This is the most likely possibility if  $\epsilon \ll 1$ . The code fails if more than two qubits have been corrupted by the noisy channel. This takes place with probability  $\epsilon_c = 3\epsilon^2(1-\epsilon) + \epsilon^3$ . Therefore, the encoding improves the transmission of quantum information provided  $\epsilon_c < \epsilon$ ; that is,  $\epsilon < \frac{1}{2}$ . This requirement is the same as in the classical three-bit code discussed at the beginning of this chapter.

A few comments are in order:

- (1) From the syndrome measurement Bob does not learn anything about the quantum state (the values of  $\alpha$  and  $\beta$ ). Hence, quantum coherence is not destroyed. This is possible because a qubit of information is encoded in a many-qubit entangled state and we only measure collective properties of this state.

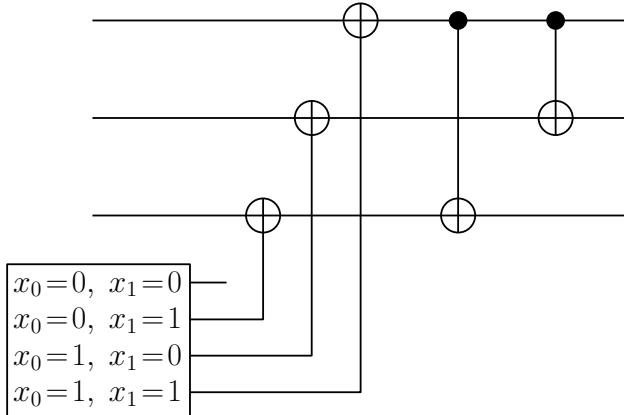


Fig. 9.3 Error correction and decoding in the three-qubit bit-flip code. The values of the classical bits  $x_0$  and  $x_1$  control the application of the NOT gates. The two CNOT gates decode the single-qubit message.

- (2) If we repeat quantum-error correction in the case of several uses of a quantum noisy channel (for instance, if we wish to stabilize the state of a quantum computer, namely the *quantum memory*, against environmental noise), every time we must supply new ancillary qubits or erase them to the  $|0\rangle$  state. This process requires the expenditure of power since, according to Landauer's principle, erasure of information dissipates energy.

**Exercise 9.1** Design a circuit to measure the error syndrome in the three-qubit bit-flip code without using any ancillary qubits.

**Exercise 9.2** Use the quantum circuit for error syndrome extraction of exercise 9.1 and assume that one of the three qubits is subjected to a rotation of angle  $\delta$  about the  $x$ -axis,  $R_x(\delta)$  (see Eq. (3.54)). Show that this error can be corrected after projective measurement of two of the three qubits. This exercise highlights the fact that quantum error correction can be seen as verification of the wave-function collapse after a measurement, as postulated by quantum mechanics.

**Exercise 9.3** Compute the fidelity of a generic pure state sent from Alice to Bob through a bit-flip noisy channel. Compare with the result obtained when the three-qubit bit-flip error-correcting code is applied.

## 9.2 The three-qubit phase-flip code

In this section, we shall show that it is also possible to correct phase errors. These are quantum errors with no classical analogue. The phase-flip error affects the states of the computational basis as follows:

$$|0\rangle \rightarrow \sigma_z |0\rangle = |0\rangle, \quad |1\rangle \rightarrow \sigma_z |1\rangle = -|1\rangle. \quad (9.9)$$

Thus, a generic state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is mapped into  $\sigma_z|\psi\rangle = \alpha|0\rangle - \beta|1\rangle$ . The method developed in Sec. 9.1 cannot correct phase errors. However, we observe that a phase-flip error in the computational basis  $\{|0\rangle, |1\rangle\}$  becomes a bit-flip error in the basis  $\{|+\rangle, |-\rangle\}$ , where

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (9.10)$$

Indeed, we have  $\sigma_z|+\rangle = |-\rangle$  and  $\sigma_z|-\rangle = |+\rangle$ . We may transform the vectors of the computational basis into the new basis vectors (and *vice versa*) by means of the Hadamard gate. Therefore, to correct phase errors we exploit the encoding of Fig. 9.4; that is,

$$|0\rangle \rightarrow |0_L\rangle = |+++ \rangle, \quad |1\rangle \rightarrow |1_L\rangle = |--- \rangle, \quad (9.11)$$

and correct the bit-flip errors in the basis  $\{|+\rangle, |-\rangle\}$  using the method described in Sec. 9.1. The final decoding step is performed simply by implementing the same array of gates as for the encoding (Fig. 9.4) but in the reverse order.

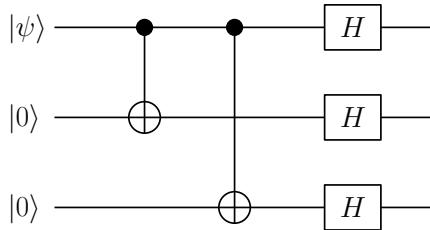


Fig. 9.4 A quantum circuit encoding a single qubit into three for the phase-flip code.

### 9.3 The nine-qubit Shor code

The nine-qubit Shor code corrects the most general possible noise acting on a single qubit.<sup>1</sup> We employ the following encoding:

$$\begin{aligned} |0\rangle \rightarrow |0_L\rangle &\equiv \frac{1}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \\ |1\rangle \rightarrow |1_L\rangle &\equiv \frac{1}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle), \end{aligned} \quad (9.12)$$

so that a generic quantum state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|0_L\rangle + \beta|1_L\rangle$ . The quantum circuit implementing this encoding is shown in Fig. 9.5. The first two CNOT and the Hadamard gates of this circuit implement the three-qubit phase-flip encoding as in Fig. 9.4,

$$|0\rangle \rightarrow |+++ \rangle, \quad |1\rangle \rightarrow |--- \rangle. \quad (9.13)$$

---

<sup>1</sup>In a more accurate analysis of errors, it is necessary to include possible errors in the quantum gates and in the circuits for coding, error correction, and decoding, see Sec. 9.10.

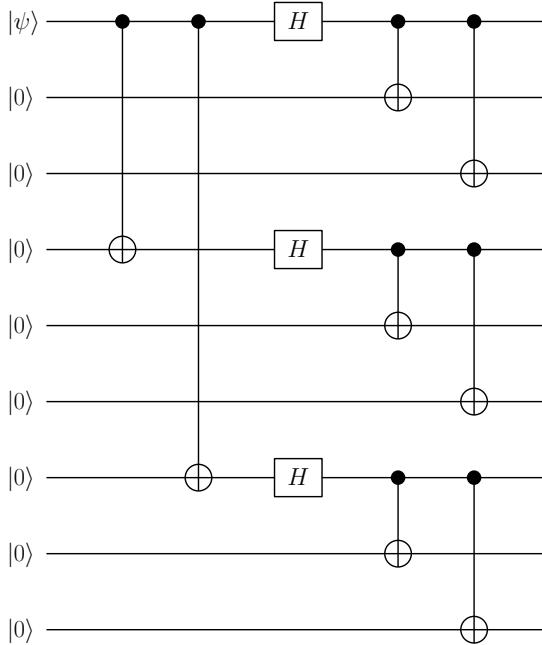


Fig. 9.5 A quantum circuit encoding a single qubit into nine.

Then, the last CNOT gates encode each of these three qubits into a block of three, by means of the three-qubit bit-flip encoding of Fig. 9.1

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \rightarrow \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \\ |-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle). \end{aligned} \quad (9.14)$$

This code can correct both bit- and phase-flip errors. The quantum circuit extracting the error syndrome is shown in Fig. 9.6. In each three-qubit block a single bit-flip error can be detected and corrected following the method described in Sec. 9.1. Moreover, we can deal with phase errors affecting a single qubit. Let us assume that the phase error occurs in the first qubit. As a consequence, the state of the first block of qubits is modified as follows (neglecting the wave function normalization):

$$\begin{aligned} |000\rangle + |111\rangle &\rightarrow |000\rangle - |111\rangle, \\ |000\rangle - |111\rangle &\rightarrow |000\rangle + |111\rangle. \end{aligned} \quad (9.15)$$

In order to detect this phase-flip error without disturbing the encoded quantum state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , we must perform collective measurements. More precisely, we measure

$$y_0 = \sigma_1^x \sigma_2^x \sigma_3^x \sigma_4^x \sigma_5^x \sigma_6^x, \quad \text{and} \quad y_1 = \sigma_1^x \sigma_2^x \sigma_3^x \sigma_7^x \sigma_8^x \sigma_9^x. \quad (9.16)$$

We have

$$\begin{aligned} \sigma_1^x \sigma_2^x \sigma_3^x (|000\rangle + |111\rangle) &= +(|000\rangle + |111\rangle), \\ \sigma_1^x \sigma_2^x \sigma_3^x (|000\rangle - |111\rangle) &= -(|000\rangle - |111\rangle). \end{aligned} \quad (9.17)$$

Therefore, if the phase flip affects the first block of qubits, we obtain  $(y_0, y_1) = (-1, -1)$ . Similarly, the cases  $(y_0, y_1) = (1, 1)$ ,  $(1, -1)$  and  $(-1, 1)$  correspond to no errors, phase error in the third block and phase error in the second block, respectively. To correct a phase error occurring in the first block of qubits, we apply the operator  $\sigma_1^z \sigma_2^z \sigma_3^z$  since

$$\sigma_1^z \sigma_2^z \sigma_3^z (|000\rangle \pm |111\rangle) = (|000\rangle \mp |111\rangle). \quad (9.18)$$

In the same manner, we correct phase errors in the second and third block by applying  $\sigma_4^z \sigma_5^z \sigma_6^z$  and  $\sigma_7^z \sigma_8^z \sigma_9^z$ , respectively.

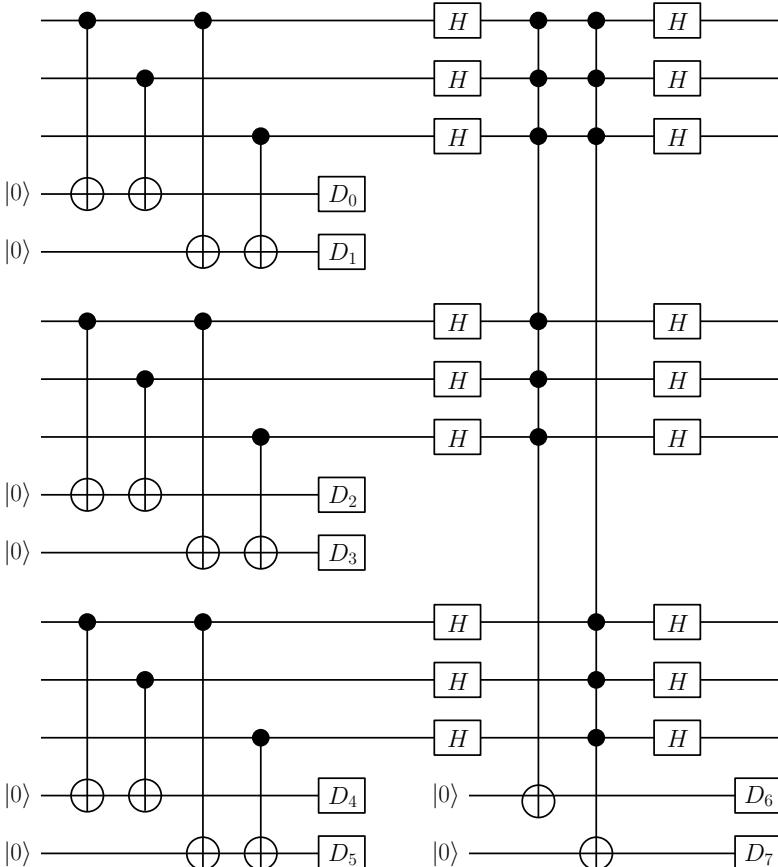


Fig. 9.6 A quantum circuit for extracting the error syndrome for the nine-qubit Shor code. The symbols  $D_i$  ( $i = 0, \dots, 7$ ) denote detectors measuring single-qubit polarizations.

The nine-qubit Shor code not only corrects single-qubit bit and phase-flip errors, but also protects against *arbitrary errors* affecting a single qubit. To understand this crucial point, let us consider a single qubit which interacts with its environment. We know from Chap. 7 that, without loss of generality, we can assume that the

environment is initially in a pure state, which we call  $|0\rangle_E$ . The most general unitary evolution  $U$  of the qubit and its environment may be written as

$$\begin{aligned} U|0\rangle|0\rangle_E &= |0\rangle|e_0\rangle_E + |1\rangle|e_1\rangle_E, \\ U|1\rangle|0\rangle_E &= |0\rangle|e_2\rangle_E + |1\rangle|e_3\rangle_E, \end{aligned} \quad (9.19)$$

where  $|e_0\rangle_E$ ,  $|e_1\rangle_E$ ,  $|e_2\rangle_E$  and  $|e_3\rangle_E$  are states of the environment, not necessarily normalized or mutually orthogonal. For a generic initial state of the system,  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , we have

$$\begin{aligned} U(\alpha|0\rangle + \beta|1\rangle)|0\rangle_E &= \alpha(|0\rangle|e_0\rangle_E + |1\rangle|e_1\rangle_E) + \beta(|0\rangle|e_2\rangle_E + |1\rangle|e_3\rangle_E) \\ &= (\alpha|0\rangle + \beta|1\rangle)\frac{1}{2}(|e_0\rangle_E + |e_3\rangle_E) + (\alpha|0\rangle - \beta|1\rangle)\frac{1}{2}(|e_0\rangle_E - |e_3\rangle_E) \\ &\quad + (\alpha|1\rangle + \beta|0\rangle)\frac{1}{2}(|e_1\rangle_E + |e_2\rangle_E) + (\alpha|1\rangle - \beta|0\rangle)\frac{1}{2}(|e_1\rangle_E - |e_2\rangle_E) \\ &= I|\psi\rangle|e_I\rangle_E + \sigma_z|\psi\rangle|e_z\rangle_E + \sigma_x|\psi\rangle|e_x\rangle_E + \sigma_x\sigma_z|\psi\rangle|e_{xz}\rangle_E, \end{aligned} \quad (9.20)$$

where

$$\begin{aligned} |e_I\rangle_E &\equiv \frac{1}{2}(|e_0\rangle_E + |e_3\rangle_E), & |e_z\rangle_E &\equiv \frac{1}{2}(|e_0\rangle_E - |e_3\rangle_E), \\ |e_x\rangle_E &\equiv \frac{1}{2}(|e_1\rangle_E + |e_2\rangle_E), & |e_{xz}\rangle_E &\equiv \frac{1}{2}(|e_1\rangle_E - |e_2\rangle_E). \end{aligned} \quad (9.21)$$

Therefore, the action of  $U$  can be expanded over the *discrete* set of operators  $\{I, \sigma_x, \sigma_y = i\sigma_x\sigma_z, \sigma_z\}$ . This is because, as can be readily checked, these operators are a basis for the Hilbert space of  $2 \times 2$  matrices. This expansion embodies the fact that arbitrary single-qubit errors can be expressed as a weighted sum of a *finite* number of errors: the bit flip ( $\sigma_x$ ), the phase flip ( $\sigma_z$ ) and the bit–phase flip ( $\sigma_x\sigma_z = -i\sigma_y$ ).

It is a fundamental feature of quantum error correction that a *continuum* of errors may be corrected by correcting only a *discrete* subset of them, namely the bit and phase-flip errors. This is because, by measuring the error syndrome, we project the superposition (9.20) onto one of the four states  $I|\psi\rangle|e_I\rangle_E$ ,  $\sigma_z|\psi\rangle|e_z\rangle_E$ ,  $\sigma_x|\psi\rangle|e_x\rangle_E$ ,  $\sigma_x\sigma_z|\psi\rangle|e_{xz}\rangle_E$ . We can then recover the original state  $|\psi\rangle$  by applying an appropriate error-correcting operation.

To grasp this point, let us give a concrete example: the correction, by means of the three-qubit bit-flip code, of a single-qubit rotation on the first qubit<sup>2</sup> described by the operator

$$U_1(\epsilon) = \cos(\epsilon)I_1 + i\sin(\epsilon)\sigma_1^x. \quad (9.22)$$

The encoded three-qubit state  $\alpha|000\rangle + \beta|111\rangle$  becomes

$$(U_1(\epsilon)\otimes I_2\otimes I_3)(\alpha|000\rangle + \beta|111\rangle) = \cos(\epsilon)(\alpha|000\rangle + \beta|111\rangle) + i\sin(\epsilon)(\alpha|100\rangle + \beta|011\rangle). \quad (9.23)$$

---

<sup>2</sup>Of course, the same error can also be corrected by the Shor code.

If we perform a collective measurement (of  $\sigma_1^z\sigma_2^z$ ) on the first two qubits, then the wave function (9.23) is projected over the undamaged state  $\alpha|000\rangle + \beta|111\rangle$  with probability  $\cos^2(\epsilon)$  or over the state  $\alpha|100\rangle + \beta|011\rangle$  with probability  $\sin^2(\epsilon)$ . In the first case, no further action is needed. In the latter case, we correct the bit-flip error as explained in Sec. 9.1.

Finally, we wish to discuss in more depth the role of encoding in the Shor code. Let us consider the case in which an arbitrary error, described by Eq. (9.20), affects the first qubit. We consider the evolution of the codewords  $|0_L\rangle$  and  $|1_L\rangle$  separately. It is sufficient to write the evolution of only the first three-qubit block since the other blocks are unchanged. We have

$$\begin{aligned} &(|000\rangle + |111\rangle)|0\rangle_E \\ &\rightarrow |000\rangle|e_0\rangle_E + |100\rangle|e_1\rangle_E + |011\rangle|e_2\rangle_E + |111\rangle|e_3\rangle_E \\ &= (|000\rangle + |111\rangle) \frac{1}{2}(|e_0\rangle_E + |e_3\rangle_E) + (|000\rangle - |111\rangle) \frac{1}{2}(|e_0\rangle_E - |e_3\rangle_E) \\ &\quad + (|100\rangle + |011\rangle) \frac{1}{2}(|e_1\rangle_E + |e_2\rangle_E) + (|100\rangle - |011\rangle) \frac{1}{2}(|e_1\rangle_E - |e_2\rangle_E). \end{aligned} \quad (9.24)$$

Similarly, we obtain

$$\begin{aligned} &(|000\rangle - |111\rangle)|0\rangle_E \\ &\rightarrow |000\rangle|e_0\rangle_E + |100\rangle|e_1\rangle_E - |011\rangle|e_2\rangle_E - |111\rangle|e_3\rangle_E \\ &= (|000\rangle - |111\rangle) \frac{1}{2}(|e_0\rangle_E + |e_3\rangle_E) + (|000\rangle + |111\rangle) \frac{1}{2}(|e_0\rangle_E - |e_3\rangle_E) \\ &\quad + (|100\rangle - |011\rangle) \frac{1}{2}(|e_1\rangle_E + |e_2\rangle_E) + (|100\rangle + |011\rangle) \frac{1}{2}(|e_1\rangle_E - |e_2\rangle_E). \end{aligned} \quad (9.25)$$

This implies that the final state of the environment is the same if the system is initially either in the encoded state  $|0_L\rangle$  or in  $|1_L\rangle$  (see exercise 9.4). The deep reason for this result is that the states  $|0_L\rangle$  and  $|1_L\rangle$  are entangled and it is impossible to tell them apart by observing just a single qubit (the state of a single qubit is equal to  $\frac{1}{2}I$  for both  $|0_L\rangle$  and  $|1_L\rangle$ ). Therefore, given an arbitrary state  $\alpha|0_L\rangle + \beta|1_L\rangle$ , the environment cannot learn anything about  $\alpha$  and  $\beta$  through interaction with a single qubit (inducing single-qubit errors). Since quantum information is not destroyed by this interaction, error recovery is possible.

**Exercise 9.4** Compute the final state of the environment when the initial state of system plus environment is described by  $|0_L\rangle|0\rangle_E$  or  $|1_L\rangle|0\rangle_E$  and a generic single-qubit error occurs ( $|0_L\rangle$  and  $|1_L\rangle$  are the codewords of the nine-qubit Shor code).

## 9.4 General properties of quantum error correction

So far, we have described quantum error correction in the case of single-qubit errors. We now discuss how to implement quantum error correction when more general errors occur. First of all, we note that errors affecting  $n$  qubits can be expanded over a set of  $4^n$  operators  $\{E_k\}$ , constructed as tensor products of the single-qubit

operators  $I$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ . The number of qubits on which a given operator differs from the identity is called the *weight* of such operator. As an example, an operator for  $n = 5$  with a weight of three is given by  $I_1 \otimes \sigma_2^y \otimes \sigma_3^x \otimes I_4 \otimes \sigma_5^z$ . The action of an arbitrary unitary operator  $U$  on the  $n$ -qubit system plus the environment is

$$U|\psi\rangle|0\rangle_E = \sum_{k=0}^{4^n-1} E_k |\psi\rangle|e_k\rangle_E, \quad (9.26)$$

where  $|\psi\rangle$  is the initial  $n$ -qubit state. We call  $\mathcal{E} = \{E_0, \dots, E_{4^n-1}\}$  the set of all possible errors affecting  $n$  qubits and  $\mathcal{E}_c$  the subset of errors that can be corrected by a code. The set  $\mathcal{E}$  of all the errors (including the identity), with a possible overall factor of  $\pm 1$  or  $\pm i$ , forms a group  $\mathcal{G}$  under multiplication, and this fact will play an important role in the general stabilizer formalism of Sec. 9.5. Note that, in general, the subset of correctable errors  $\mathcal{E}_c$  is not required to form a group, neither to be closed under multiplication.

Let us discuss what conditions should be satisfied to allow error correction. First of all, correctable errors should map two different codewords  $|i_L\rangle$  and  $|j_L\rangle$  into orthogonal states:

$$\langle i_L | E_a^\dagger E_b | j_L \rangle = 0 \quad \text{for } i \neq j, \quad (9.27)$$

where  $E_a$  and  $E_b \in \mathcal{E}_c$ . If this condition were not satisfied, then the states  $E_a|i_L\rangle$  and  $E_b|j_L\rangle$  could not be distinguished with certainty and therefore perfect error correction would be impossible. The second condition is that, for any correctable errors  $E_a$  and  $E_b$ ,

$$\langle i_L | E_a^\dagger E_b | i_L \rangle = C_{ab}, \quad (9.28)$$

where  $C_{ab}$  does not depend on the state  $|i_L\rangle$ . If this were not the case, we would obtain some information on the encoded state from the measurement of the error syndrome. Therefore, we would inevitably disturb the quantum state. Note that  $C_{ab} = C_{ba}^*$ .

Conditions (9.27) and (9.28) can be put together and it is possible to prove that error correction is possible if and only if

$$\langle i_L | E_a^\dagger E_b | j_L \rangle = C_{ab} \delta_{ij}, \quad (9.29)$$

where  $E_a$  and  $E_b$  belong to the set  $\mathcal{E}_c$  of correctable errors and the matrix  $C_{ab}$  is Hermitian and independent of  $|i_L\rangle$  and  $|j_L\rangle$  (for a proof see, e.g., Preskill, 1998a). If  $C_{ab} = \delta_{ab}$ , the code is known as *non-degenerate*. In this case, it is possible to identify with certainty which error occurred. In contrast, if  $C_{ab} \neq \delta_{ab}$ , we call the code *degenerate*.

**Exercise 9.5** Show that condition (9.29) is fulfilled by the three-qubit bit-flip code.

**Exercise 9.6** Show that the three-qubit bit-flip code is non-degenerate while the nine-qubit Shor code is degenerate.

We note that, in equation (9.29), for any  $E_a$  and  $E_b$  belonging to the group  $\mathcal{G}$ , the product  $E = E_a^\dagger E_b$  is still in the same group. The weight of the smallest error in  $\mathcal{G}$  for which (9.29) does not hold is called the *distance* of the code. A code that is able to correct errors up to weight  $t$  must have distance at least  $2t + 1$ . Every code has distance at least one. In the following we will adopt the shorthand  $[[n, k, d]]$  in order to indicate a quantum error-correction code of distance  $d$ , encoding  $k$  qubits in  $n$  qubits. For example, as we shall see in Sec. 9.5, the nine-qubit Shor code is a  $[[9, 1, 3]]$  code.

It is instructive to describe the error recovery procedure in the simple case of non-degenerate codes. Provided the system has been subjected to correctable errors, the most general system plus environment state is given by

$$\sum_{E_k \in \mathcal{E}_c} E_k |\psi\rangle |e_k\rangle_E. \quad (9.30)$$

To measure the error syndrome, we can attach ancillary qubits, initially in a well known state  $|0\rangle_A$ , to the system, and operate the unitary transformation

$$\sum_{E_k \in \mathcal{E}_c} E_k |\psi\rangle |e_k\rangle_E |0\rangle_A \rightarrow \sum_{E_k \in \mathcal{E}_c} E_k |\psi\rangle |e_k\rangle_E |a_k\rangle_A. \quad (9.31)$$

A projective measurement of the ancillary qubits will then collapse this sum to a single term

$$E_{\bar{k}} |\psi\rangle |e_{\bar{k}}\rangle_E |a_{\bar{k}}\rangle_A. \quad (9.32)$$

Note that the system is now de-entangled from the environment and from the ancillary qubits. Since the operators  $E_k$  are unitary (they are constructed as tensor products of the Pauli matrices, which are unitary), it is sufficient to apply the unitary operator  $E_{\bar{k}}^\dagger = E_{\bar{k}}$  to the system to recover the original state  $|\psi\rangle$ .

#### 9.4.1 The quantum Hamming bound

The quantum Hamming bound only applies to non-degenerate codes. It tells us the minimum number  $n$  of physical qubits required to encode  $k$  logical qubits, in such a manner that errors affecting at most  $t$  qubits can be corrected. If  $j$  errors occur, there are  $\binom{n}{j}$  possible locations for these errors. For instance, if  $n = 3$  and  $j = 2$ , the  $\binom{3}{2}$  possibilities are: (i) errors in the first and second qubit, (ii) in the first and third qubit, and (iii) in the second and third qubit. Each qubit may be subjected to three possible errors (bit flip  $\sigma_x$ , phase flip  $\sigma_z$ , and bit–phase flip  $\sigma_x \sigma_z = -i\sigma_y$ ). Hence, there are  $3^j$  possible errors for each error location. The total number of possible errors affecting  $t$  or less qubits is therefore given by

$$\sum_{j=0}^t \binom{n}{j} 3^j. \quad (9.33)$$

Note that the sum over  $j$  starts from zero to include the error-free case too. To encode  $k$  qubits by means of a non-degenerate code, each of these errors must correspond to a  $2^k$ -dimensional subspace. These subspaces must be mutually orthogonal

and belong to the  $2^n$ -dimensional Hilbert space for  $n$  qubits. Therefore, we can write the quantum Hamming bound

$$\sum_{j=0}^t \binom{n}{j} 3^j 2^k \leq 2^n. \quad (9.34)$$

For non-degenerate codes correcting a single error ( $t = 1$ ), the quantum Hamming bound reduces to  $(1 + 3n) 2^k \leq 2^n$ . Let us call  $n_{\min}$  the smallest value of  $n$  satisfying this bound. For codes encoding a single qubit ( $k = 1$ ) and correcting arbitrary single-qubit errors,  $n_{\min} = 5$  qubits.<sup>3</sup> Note that the ratio  $n_{\min}/k$  decreases with  $k$ . For example,  $n_{\min} = 12$  for  $k = 6$ . Therefore, the encoding of quantum information is more efficient for large  $k$ . The price to pay is a greater complexity of the corresponding quantum error-correcting codes.

## 9.5 Stabilizer coding

We are now in the position to introduce the basic concepts of stabilizer codes, a group-theoretical general structure that has proved particularly useful in order to understand the structure of some classes of error-correction codes and also to produce new codes. This represents nowadays the standard formalism to deal with quantum error correction. In order to do that, we will first go back to the nine-qubit Shor code and work out its features from a more general perspective. At a second stage, we will describe the stabilizer formalism.

### 9.5.1 The nine-qubit Shor code revisited

We have seen before that, according to the nine-qubit code described in Sec. 9.3, in order to detect a bit-flip error in any of the first three qubits, one needs to perform collective measurements of two qubits simultaneously. Specifically they correspond to measuring the eigenvalues of  $\sigma_1^z \sigma_2^z$  and  $\sigma_1^z \sigma_3^z$  (see also Sec. 9.1). If the two qubits are the same, the eigenvalue of the corresponding operator will be  $+1$ ; if the two qubits are different, the eigenvalue will be  $-1$  and thus an error is witnessed. An analogous procedure has to be adopted to detect bit-flip errors in the second and the third group of three qubits (with two collective measurements per group). Conversely, a sign (phase-flip) error can be detected by measuring the eigenvalues of  $y_0$  and  $y_1$  in Eq. (9.16), which enable to compare the signs of the first and second blocks of three, and the first and third blocks of three, respectively. If the signs agree, the eigenvalues will be  $+1$ ; if they disagree, the eigenvalues will be  $-1$  thus again signalling an error.

Summarizing all the above detailed procedure, in order to totally correct the codeword, one needs to measure the eigenvalues of a total of eight operators, which are listed in Table 9.1. The two valid codewords  $|0_L\rangle$  and  $|1_L\rangle$  of Shor's code in

---

<sup>3</sup>It is possible to prove that  $n_{\min} = 5$  also in the case of degenerate codes, see Knill and Laflamme (1997).

Eq. (9.12) are eigenvectors of all eight of these operators, with eigenvalue +1. Then all the error operators in the group  $\mathcal{G}$  that fix both  $|0_L\rangle$  and  $|1_L\rangle$  can be written as the product of these eight operators, and form themselves a group, named the *stabilizer  $\mathcal{S}$  of the code*. The operators  $\{M_i\}_{i=1,\dots,8}$  represent the generators of the stabilizer group for the nine-qubit Shor code.

When measuring the eigenvalue of  $M_1$ , we determine whether a bit flip error has occurred on qubit one ( $\sigma_1^x$ ) or on qubit two ( $\sigma_2^x$ ). Note that both these errors anti-commute with  $M_1$ . Similarly,  $M_2$  detects either  $\sigma_1^x$  or  $\sigma_3^x$ , which anti-commute with it, and so on, up to  $M_6$ . The operator  $M_7$  detects  $\sigma_1^z$  through  $\sigma_6^z$  (commuting with them), and similarly for the operator  $M_8$ . In general we have that, for a given linear superposition  $|\psi\rangle$  of the two codewords  $|0_L\rangle$  and  $|1_L\rangle$ ,  $M \in \mathcal{S}$ , and for any single-qubit error  $E$  such that  $\{M, E\} = 0$ ,

$$ME|\psi\rangle = -EM|\psi\rangle = -E|\psi\rangle, \quad (9.35)$$

so that  $E|\psi\rangle$  is eigenvector of  $M$  with eigenvalue  $-1$ . To detect an error, it is thus sufficient to measure the generator  $M$  of the stabilizer that anti-commutes with it.

**Exercise 9.7** Show explicitly that any single-qubit operator like  $\sigma_i^x$ ,  $\sigma_i^y$ , or  $\sigma_i^z$  anti-commutes with one or more elements  $M_j \in \mathcal{S}$ . This immediately proves that we have found a way to correct any error  $E$  of weight one.

Since states  $|\psi\rangle$  with different eigenvalues are orthogonal, condition (9.29) is always satisfied, whenever  $E_a$  has weight one and  $E_b = I$ . It is also possible to see that every two-qubit operator  $E$  anti-commutes with some element of  $\mathcal{S}$ , except for those of the form  $\sigma_a^z \sigma_b^z$ , where (a) and (b) are the positions of two qubits in the same block of three. However, operators of such form actually belong to the stabilizer. This means that  $\sigma_a^z \sigma_b^z |\psi\rangle = |\psi\rangle$  for any linear combination of codewords, and thus  $\langle \psi | \sigma_a^z \sigma_b^z | \psi \rangle = \langle \psi | \psi \rangle = 1$ , and Eq. (9.29) is satisfied as well. This argument shows that the distance of the nine-qubit Shor code is at least three. Conversely, operators of weight three do not satisfy Eq. (9.29), thus the distance of the nine-qubit Shor code is exactly three, and we say that it is a  $[[9, 1, 3]]$  code.

**Exercise 9.8** Show that an example of weight-three error operator which does not satisfy Eq. (9.29) is  $O_{3w} = \sigma_1^x \sigma_2^x \sigma_3^x$ .

Table 9.1 Stabilizer for the nine-qubit Shor code.

$M_1$	$\sigma_1^z$	$\sigma_2^z$	$I_3$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$I_9$
$M_2$	$\sigma_1^z$	$I_2$	$\sigma_3^z$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$I_9$
$M_3$	$I_1$	$I_2$	$I_3$	$\sigma_4^z$	$\sigma_5^z$	$I_6$	$I_7$	$I_8$	$I_9$
$M_4$	$I_1$	$I_2$	$I_3$	$\sigma_4^z$	$I_5$	$\sigma_6^z$	$I_7$	$I_8$	$I_9$
$M_5$	$I_1$	$I_2$	$I_3$	$I_4$	$I_5$	$I_6$	$\sigma_7^z$	$\sigma_8^z$	$I_9$
$M_6$	$I_1$	$I_2$	$I_3$	$I_4$	$I_5$	$I_6$	$\sigma_7^z$	$I_8$	$\sigma_9^z$
$M_7$	$\sigma_1^x$	$\sigma_2^x$	$\sigma_3^x$	$\sigma_4^x$	$\sigma_5^x$	$\sigma_6^x$	$I_7$	$I_8$	$I_9$
$M_8$	$\sigma_1^x$	$\sigma_2^x$	$\sigma_3^x$	$I_4$	$I_5$	$I_6$	$\sigma_7^x$	$\sigma_8^x$	$\sigma_9^x$

### 9.5.2 \* General formalism for stabilizer codes

Let us come back to the definition of the group  $\mathcal{G}$ , formed by all the possible tensor products of the single-qubit operations  $I$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ . We will denote by  $\mathcal{G}_n$  the group made out of  $n$  of such matrices, with a possible overall factor of  $\pm 1$  or  $\pm i$ :

$$\mathcal{G}_n = \{\pm 1, \pm i\} \otimes \{I, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n}, \quad (9.36)$$

and refer to it as the  $n$ -qubit Pauli group. The following properties hold:

- (1) each  $M \in \mathcal{G}_n$  is unitary:  $M^{-1} = M^\dagger$ ;
- (2) for each element  $M \in \mathcal{G}_n$ , we have that  $M^2 = \pm I^{\otimes n}$ ;
- (3) any two elements  $M, N \in \mathcal{G}_n$  either commute or anti-commute:  $MN = \pm NM$ .

We point out that the space of Pauli matrices is  $2n$ -dimensional, since  $\sigma_y$  can be seen as a product of  $\sigma_x$  and  $\sigma_z$  (up to overall  $\pm 1, \pm i$  factors). Indeed each element of  $\mathcal{G}_n$  can be always expressed as a product  $\otimes_{i=1}^n (\sigma_z^{\alpha_i} \sigma_x^{\beta_j})$ , where  $\alpha$  and  $\beta$  are binary strings of length  $n$ .

Suppose now that  $\mathcal{S}$  is a given Abelian subgroup of  $\mathcal{G}_n$  (that is, all the elements of  $\mathcal{S}$  do commute between them). Thus all the elements of  $\mathcal{S}$  acting on the  $2^n$  dimensional Hilbert space  $\mathcal{H}_{2^n}$  of  $n$  qubits can be simultaneously diagonalized. The stabilizer code  $\mathcal{H}_{\mathcal{S}} \subseteq \mathcal{H}_{2^n}$  associated with  $\mathcal{S}$  is defined as the simultaneous eigenspace with eigenvalue 1 of all the elements of  $\mathcal{S}$ :

$$|\psi\rangle \in \mathcal{H}_{\mathcal{S}} \iff M|\psi\rangle = |\psi\rangle, \quad \forall M \in \mathcal{S}. \quad (9.37)$$

The group  $\mathcal{S}$  is called the stabilizer of the code, since it preserves all the codewords. It is possible to define the generators  $\{M_i\}$  of the stabilizer  $\mathcal{S}$  as a set of independent operators (no one can be expressed as a product of others) such that each element of  $\mathcal{S}$  can be expressed as a product of elements of  $\{M_i\}$ . It is possible to show that if  $\mathcal{S}$  has  $n - k$  generators, the code space  $\mathcal{H}_{\mathcal{S}}$  has dimension  $2^k$ , that is, there are  $k$  encoded qubits (for a proof see, e.g., Preskill, 1998a).

The stabilizer formalism provides a simple way to characterize errors that the code can detect and correct. We may think of the  $n - k$  generators  $\{M_1, \dots, M_{n-k}\}$  as the check operators of the code, or, in other words, the observables that one needs to measure in order to diagnose the errors. As a matter of fact, if the encoded information in  $|\phi\rangle \in \mathcal{H}_{2^n}$  is undamaged, then  $\langle \phi | M_i | \phi \rangle = 1$  for all the generators; on the other hand, if  $\langle \phi | M_i | \phi \rangle = -1$  for some  $i$ , then  $|\phi\rangle$  is orthogonal to the code subspace and an error has been detected.

Let us now consider a generic error  $E$ , which can be expanded in terms of elements of the Pauli group  $\mathcal{G}_n$ . Due to the properties of  $\mathcal{G}_n$ , we have that  $E$  either commutes or anti-commutes with a given  $M \in \mathcal{S}$ . If  $E$  and  $M$  commute, then for any  $|\psi\rangle \in \mathcal{H}_{\mathcal{S}}$  we have:

$$ME|\psi\rangle = EM|\psi\rangle = E|\psi\rangle, \quad (9.38)$$

so that the error preserves the outcome  $M = 1$ . Conversely, if  $E$  and  $M$  anti-commute, then

$$ME|\psi\rangle = -EM|\psi\rangle = -E|\psi\rangle, \quad (9.39)$$

so that the error can be detected by measuring  $M = -1$ . In general, for a given stabilizer generator  $M_j \in \mathcal{S}$  and an error  $E_a \in \mathcal{G}_n$ , one can write:

$$M_j E_a = (-1)^{s_{ja}} E_a M_j. \quad (9.40)$$

The numbers  $s_{ja}$ , with  $j = 1, \dots, n - k$  are called the syndrome for the error  $E_a$ , since  $(-1)^{s_{ja}}$  expresses the result of the measure of  $M_i$  if the error  $E_a$  occurs. In the case of a non-degenerate code, the syndrome will be distinct for all the errors  $E_a \in \mathcal{E}_c$  that can be corrected by the code, so that measuring the  $n - k$  stabilizer generators will identify the error completely.

More generally, it is possible to write a condition to be satisfied by the stabilizer, that is sufficient to ensure that error recovery is possible. We recall the condition for error correction of Eq. (9.29), stating that for each  $E_a, E_b \in \mathcal{E}_c$  and  $|\psi\rangle \in \mathcal{H}_{\mathcal{S}}$  one has:

$$\langle \psi | E_a^\dagger E_b | \psi \rangle = C_{ab}, \quad (9.41)$$

where  $C_{ab}$  is independent of  $|\psi\rangle$ . This condition is satisfied provided that, for each  $E_a, E_b \in \mathcal{E}_c$ , one of the following holds:

- (1)  $E_a^\dagger E_b \in \mathcal{S}$ ;
- (2) There is an  $M \in \mathcal{S}$  that anti-commutes with  $E_a^\dagger E_b$ .

**Exercise 9.9** Show that (9.41) is satisfied provided that either (1) or (2) holds.

In conclusion we have shown that a stabilizer code that is able to correct errors in  $\mathcal{E}_c$  is a space  $\mathcal{H}_{\mathcal{S}}$  fixed by an Abelian subgroup  $\mathcal{S}$  of the Pauli group  $\mathcal{G}_n$ , where either (1) or (2) is satisfied by every  $E_a^\dagger E_b$  with  $E_a, E_b \in \mathcal{E}_c$ . The code is non-degenerate if condition (1) is not satisfied for any  $E_a^\dagger E_b$ .

Two stabilizer codes are said to be equivalent if they differ only according to how the qubits are labeled, and how the basis for each single-qubit Hilbert space is chosen. This amounts to say that the stabilizer of one code can be transformed into the stabilizer of the other by a permutation of the qubits, together with a tensor product of single-qubit transformations.

Recovery may fail if there is an operator  $E_a^\dagger E_b$  that commutes with the stabilizer, but does not lie in the stabilizer. Such operator preserves the code subspace  $\mathcal{H}_{\mathcal{S}}$ , but may act non-trivially in that space; thus it can modify the encoded information. A stabilizer code with distance  $d$  has the property that each  $E \in \mathcal{G}_n$  of weight less than  $d$  either lies in the stabilizer, or anti-commutes with some element of  $\mathcal{S}$ . The code is non-degenerate if  $\mathcal{S}$  contains no elements of weight less than  $d$ . A distance  $d = 2t + 1$  code can correct  $t$  errors, and a distance  $s + 1$  code can detect  $s$  errors or correct  $s$  errors at known locations.

### 9.5.3 \* Logical operators for stabilizer codes

We now shed light on those Pauli operators which commute with the stabilizer, but lie outside it. Using the group algebra formalism, we will denote with  $\mathcal{S}^\perp$  the space

of vectors that are orthogonal to each vector in the stabilizer  $\mathcal{S}$ . This space is called the *normalizer group* of  $\mathcal{S}$ , also referred to as  $N(\mathcal{S})$ , that is a subgroup containing all the elements that commute with each element of  $\mathcal{S}$ . Since  $\mathcal{S}$  is Abelian, all the vectors contained in it are mutually orthogonal, and thus every stabilizer is contained in its normalizer:  $\mathcal{S} \subseteq \mathcal{S}^\perp$ .

In the 2 dimensional space of the Pauli matrices  $\mathcal{G}_n$ , the normalized subspace  $\mathcal{S}^\perp$ , containing all the vectors that are orthogonal to each of the  $n - k$  linearly independent vectors of  $\mathcal{S}$ , has dimension  $2n - (n - k) = n + k$ . Of the  $n + k$  vectors that span this space,  $n - k$  can be chosen to be the generators of the stabilizer itself. The remaining  $2k$  generators preserve the code subspace, because they commute with the stabilizer, but act non-trivially on the  $k$  encoded qubits. They can be regarded as *logical operations* that act on the encoded data that is protected by the code. They can be chosen to be single-qubit operators  $\Sigma_i^z$ ,  $\Sigma_i^x$ , ( $i = 1, \dots, k$ ), corresponding to Pauli operators  $\sigma_z$  and  $\sigma_x$  acting on the  $i$ -th encoded qubit. The simultaneous eigenstates of  $\Sigma_1^z \dots \Sigma_k^z$  (in the code subspace  $\mathcal{H}_{\mathcal{S}}$ ) can be seen as the logical basis states  $|\bar{z}_1, \dots, \bar{z}_k\rangle$ , with  $\bar{z}_j = 0$  corresponding to  $\Sigma_j^z = 1$  and  $\bar{z}_j = 1$  corresponding to  $\Sigma_j^z = -1$ . The remaining  $k$  generators  $\Sigma_i^x$  of  $\mathcal{S}^\perp$  may be chosen to be mutually commuting and to commute with the stabilizer, but they will not commute with any of the  $\Sigma_i^z$ . In such case, the effect of each  $\Sigma_j^x$  would be that of flipping the eigenvalue of the corresponding  $\Sigma_j^z$ .

As a practical example, following the discussion in exercise 9.8, one logical operator for the nine-qubit Shor code is  $\Sigma_x = \sigma_1^x \sigma_2^x \sigma_3^x$ . The other one can be shown to be  $\Sigma_z = \sigma_1^z \sigma_4^z \sigma_5^z$ . Indeed these two operators anti-commute with one another, commute with the stabilizer generators, and are independent of the generators (no element of the stabilizer contains three  $\sigma_x$  or three  $\sigma_z$ ).

## 9.6 \* The five-qubit code

In this section, we describe a quantum error-correcting code which protects a qubit of information against arbitrary single-qubit errors. To accomplish this, it is sufficient to encode a single logical qubit into five physical qubits, the minimum number required for this task. Specifically, one can construct the following  $[[5, 1, 3]]$  code, given by the four stabilizer generators of Table 9.2.

Table 9.2 Stabilizer for the five-qubit  $[[5, 1, 3]]$  code.

$M_1$	$\sigma_1^x$	$\sigma_2^z$	$\sigma_3^z$	$\sigma_4^x$	$I_5$
$M_2$	$I_1$	$\sigma_2^x$	$\sigma_3^z$	$\sigma_4^z$	$\sigma_5^x$
$M_3$	$\sigma_1^x$	$I_2$	$\sigma_3^x$	$\sigma_4^z$	$\sigma_5^z$
$M_4$	$\sigma_1^z$	$\sigma_2^x$	$I_3$	$\sigma_4^z$	$\sigma_5^z$

As can be immediately observed from the table, each generator is obtained from the previous one by performing a cyclic permutation of the qubits, and all the generators do commute between them. The missing fifth operator  $M_5 = \sigma_1^z \sigma_2^z \sigma_3^x I_4 \sigma_5^x$

can be constructed as a linear combination of the other four ones, therefore it is not independent of them. One can quickly check that each Pauli operator of weight 1 or 2 anti-commutes with at least one generator. Moreover it is easy to find an operator of weight 3 which commutes with all the generators, but lies outside the stabilizer (for example, one can take  $\sigma_1^y \sigma_2^z \sigma_3^y$ ). Therefore the distance of the code is 3, and the code is non-degenerate.

Because of the cyclic property of the code, there is an immediate way to characterize all the 15 non-trivial elements of its stabilizer. Aside from  $M_1$  and the four operators obtained from it by cyclic permutations, the stabilizer will also contain  $M_1 M_2$  plus its cyclic permutations, and  $M_1 M_3$  plus its cyclic permutations. For the logical operators, one can choose  $\Sigma_z = \sigma_1^z \sigma_2^z \sigma_3^z \sigma_4^z \sigma_5^z$  and  $\Sigma_x = \sigma_1^x \sigma_2^x \sigma_3^x \sigma_4^x \sigma_5^x$ , since they commute with all the  $\{M_i\}_{i=1,\dots,4}$ , square to the identity, and anti-commute with one another.

Let us now construct a basis of codewords for this code. In view of the cyclic property of the stabilizer, it is easy to see that a cyclic permutation of a codeword will be again a codeword. The easiest way to construct a codeword is the following:

$$\begin{aligned}
|0\rangle \rightarrow |0_L\rangle &\equiv \sum_{M \in \mathcal{S}} M|00000\rangle \\
&= [I + (M_1 + \text{perms.}) + (M_1 M_2 + \text{perms.}) + (M_1 M_3 + \text{perms.})] |00000\rangle \\
&= |00000\rangle + (|10010\rangle + |01001\rangle + |10100\rangle + |01010\rangle + |00101\rangle) \\
&\quad - (|11011\rangle + |11101\rangle + |11110\rangle + |01111\rangle + |10111\rangle) \\
&\quad - (|00110\rangle + |00011\rangle + |10001\rangle + |11000\rangle + |01100\rangle),
\end{aligned} \tag{9.42}$$

such that  $M'|0_L\rangle = |0_L\rangle$  for each  $M' \in \mathcal{S}$ . Indeed the multiplication by an element of the stabilizer merely permutes the terms in the sum. The other codeword can be found by simply applying the logical operator  $\Sigma_x$  to  $|0_L\rangle$ , that is by flipping all the five qubits:

$$\begin{aligned}
|1\rangle \rightarrow |1_L\rangle &\equiv \Sigma_x |0_L\rangle \\
&= |11111\rangle + (|01101\rangle + |10110\rangle + |01011\rangle + |10101\rangle + |11010\rangle) \\
&\quad - (|00100\rangle + |00010\rangle + |00001\rangle + |10000\rangle + |01000\rangle) \\
&\quad - (|11001\rangle + |11100\rangle + |01110\rangle + |00111\rangle + |10011\rangle).
\end{aligned} \tag{9.43}$$

Quite remarkably, it can be shown that it is possible to devise a simpler encoding of the two codewords, for the  $[[5, 1, 3]]$  code:

$$\begin{aligned}
|0\rangle \rightarrow |0_L\rangle &\equiv \frac{1}{\sqrt{8}} (|00000\rangle - |01111\rangle - |10011\rangle + |11100\rangle \\
&\quad + |00110\rangle + |01001\rangle + |10101\rangle + |11010\rangle), \\
|1\rangle \rightarrow |1_L\rangle &\equiv \frac{1}{\sqrt{8}} (|11111\rangle - |10000\rangle + |01100\rangle - |00011\rangle \\
&\quad + |11001\rangle + |10110\rangle - |01010\rangle - |00101\rangle),
\end{aligned} \tag{9.44}$$

which is implemented by the circuit in Fig. 9.7.

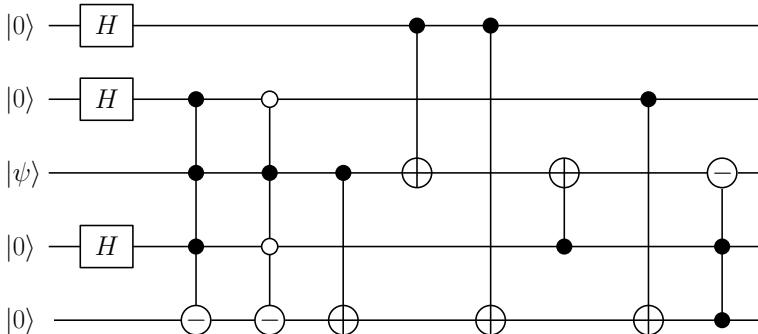


Fig. 9.7 A quantum circuit encoding a single qubit into five. The circles with a minus sign correspond to a phase shift of  $\pi$ . The qubit  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is encoded into the five-qubit state  $\alpha|0_L\rangle + \beta|1_L\rangle$ .

A remarkable feature of the five-qubit code is that the circuit for detecting the error syndrome, drawn in Fig. 9.8, is exactly the same as that for encoding, but run backwards. There are 15 possible single-qubit errors, 3 for each of the five qubits (bit flip, phase flip and bit–phase flip). The four measurements in Fig. 9.8 provide the 4 classical bits  $a, b, c$  and  $d$  allowing us to distinguish the 15 possible errors plus the case without errors. Table 9.3 exhibits all possibilities. For instance, the case  $a = b = c = d = 0$  corresponds to no errors. If instead the outcomes of the measurements are  $a = 1$  and  $b = c = d = 0$ , then the bit-flip error affected the first qubit. Different outcomes (error syndromes) are associated with different errors, as shown in Table 9.3. The post-measurement state  $|\psi'\rangle$  of the qubit carrying the quantum information is shown in the same table. It is easy to see that the original state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  is recovered by a unitary transformation  $U$  that depends upon the results  $a, b, c$  and  $d$  of the measurements. For examples, if  $a = b = c = 0$  and  $d = 1$ , then  $|\psi'\rangle = \alpha|0\rangle - \beta|1\rangle$  and we restore  $|\psi\rangle$  by means of  $U = \sigma_z$ . Indeed,  $\sigma_z|\psi'\rangle = |\psi\rangle$ .

### Exercise 9.10 Verify Table 9.3.

We point out that the five-qubit code does not require any ancillary qubits. In any case, the code is dissipative: to apply the code again we must first of all encode the state  $|\psi\rangle$  onto the five-qubit state  $\alpha|0_L\rangle + \beta|1_L\rangle$ . For this purpose, we must supply four new ancillary qubits prepared in the state  $|0\rangle$ . Alternatively, if we wish to recycle the ancillary qubits, we must first map their state  $|abcd\rangle$  into  $|0000\rangle$ . This means that the information contained in the classical bits  $a, b, c$  and  $d$  is erased. As we know from Landauer's principles, erasure is a dissipative process. Therefore, expenditure of power is required to correct errors.

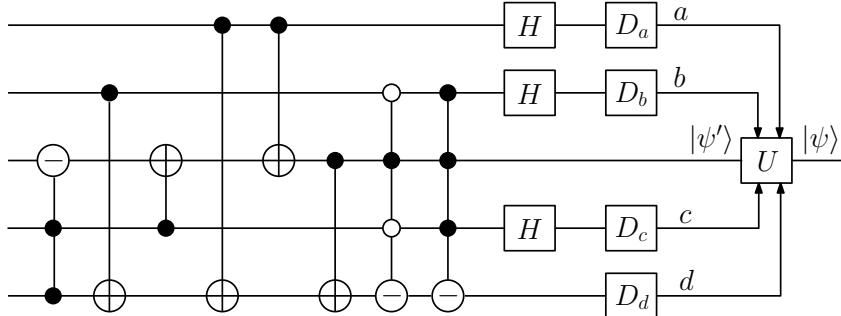


Fig. 9.8 A quantum circuit extracting the error syndrome and recovering the correct state  $|\psi\rangle$  in the five-qubit code. The four detectors  $D_a$ ,  $D_b$ ,  $D_c$  and  $D_d$  measure single qubit polarizations. The resulting classical bits  $a$ ,  $b$ ,  $c$  and  $d$  drive the unitary operator  $U$ , which maps  $|\psi'\rangle$  onto the original state  $|\psi\rangle$ .

Table 9.3 Error, syndrome and resulting state in the five-qubit code.

Error	$abcd$	$ \psi'\rangle$
None	0000	$\alpha 0\rangle + \beta 1\rangle$
$\sigma_3^x \sigma_3^z$	1101	$-\alpha 1\rangle + \beta 0\rangle$
$\sigma_5^x \sigma_5^z$	1111	$-\alpha 0\rangle + \beta 1\rangle$
$\sigma_2^x$	0001	$\alpha 0\rangle - \beta 1\rangle$
$\sigma_3^z$	1010	$\alpha 0\rangle - \beta 1\rangle$
$\sigma_5^z$	1100	$\alpha 0\rangle - \beta 1\rangle$
$\sigma_2^x \sigma_2^z$	0101	$\alpha 0\rangle - \beta 1\rangle$
$\sigma_5^x$	0011	$-\alpha 0\rangle - \beta 1\rangle$
$\sigma_1^z$	1000	$-\alpha 0\rangle - \beta 1\rangle$
$\sigma_2^z$	0100	$-\alpha 0\rangle - \beta 1\rangle$
$\sigma_4^z$	0010	$-\alpha 0\rangle - \beta 1\rangle$
$\sigma_1^x$	0110	$-\alpha 1\rangle + \beta 0\rangle$
$\sigma_3^x$	0111	$-\alpha 1\rangle + \beta 0\rangle$
$\sigma_4^x$	1011	$-\alpha 1\rangle + \beta 0\rangle$
$\sigma_1^x \sigma_1^z$	1110	$-\alpha 1\rangle + \beta 0\rangle$
$\sigma_4^x \sigma_4^z$	1001	$-\alpha 1\rangle + \beta 0\rangle$

## 9.7 Decoherence-free subspaces

In this section, we shall discuss *passive* quantum error-avoiding codes, in which no measurements or recovery operations are performed to detect and correct errors. The basic idea of passive codes is to encode the information in decoherence-free subspaces. This is possible if the system–environment interaction has certain *symmetries*.

An example will help us clarify this concept. Let us assume that a system of  $n$  qubits is coupled to the environment in a symmetric manner and undergoes a

dephasing process, defined as

$$|0\rangle_j \rightarrow |0\rangle_j, \quad |1\rangle_j \rightarrow e^{i\phi}|1\rangle_j, \quad (j = 1, \dots, n). \quad (9.45)$$

In this model, we suppose that the phase  $\phi$  has no dependence on the qubit  $j$ ; that is, the dephasing process is invariant under qubit permutations. This is an example of *collective decoherence*: several qubits couple identically to the environment. A concrete example of collective decoherence is obtained when an  $n$ -qubit register is implemented by a solid state system and the main source of errors is the coupling of each qubit with phonons whose wavelength is much larger than the distance between the qubits.

It is instructive to consider what happens for  $n = 2$  qubits. We have

$$|00\rangle \rightarrow |00\rangle, \quad |01\rangle \rightarrow e^{i\phi}|01\rangle, \quad |10\rangle \rightarrow e^{i\phi}|10\rangle, \quad |11\rangle \rightarrow e^{2i\phi}|11\rangle. \quad (9.46)$$

Since the states  $|01\rangle$  and  $|10\rangle$  acquire the same phase, a simple encoding allows us to avoid phase errors:

$$|0_L\rangle \equiv |01\rangle, \quad |1_L\rangle \equiv |10\rangle. \quad (9.47)$$

Then the state  $|\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$  evolves under the dephasing process as follows:

$$|\psi_L\rangle \rightarrow \alpha e^{i\phi}|01\rangle + \beta e^{i\phi}|10\rangle = e^{i\phi}|\psi_L\rangle. \quad (9.48)$$

The overall phase factor  $e^{i\phi}$  acquired due to the dephasing process has no physical significance. Therefore, the two-dimensional subspace spanned by the states  $|01\rangle$  and  $|10\rangle$  is a decoherence-free subspace, which can be used to encode a single qubit. It is interesting that, if the dephasing angles are different for the two qubits ( $|1\rangle_1 \rightarrow e^{i\phi_1}|1\rangle_1$  and  $|1\rangle_2 \rightarrow e^{i\phi_2}|1\rangle_2$ ), then the dephasing angle affecting the state  $|\psi_L\rangle$  is  $\phi_1 - \phi_2$ . Indeed, we have

$$|\psi_L\rangle \rightarrow e^{i\phi_2}(\alpha|01\rangle + \beta e^{i(\phi_1-\phi_2)}|10\rangle). \quad (9.49)$$

For  $n = 3$ , it is easy to check that the subspaces spanned by  $\{|000\rangle\}$ ,  $\{|100\rangle, |010\rangle, |001\rangle\}$ ,  $\{|011\rangle, |101\rangle, |110\rangle\}$  and  $\{|111\rangle\}$  are decoherence-free. Indeed, the states residing in these subspaces acquire global phases 1,  $e^{i\phi}$ ,  $e^{2i\phi}$  and  $e^{3i\phi}$ , respectively.

More generally, in the  $n$ -qubit case any subspace spanned by the states of the computational basis with an equal number of 1's and 0's (say,  $k$  1's and  $n - k$  0's) is decoherence-free. These subspaces have dimension  $d_k = \binom{n}{k}$  and may be used to encode  $\log_2 d_k$  qubits.

**Exercise 9.11** Is it possible to find decoherence-free subspaces in the case in which amplitude (bit-flip) errors act identically on every qubit of a quantum register?

### 9.7.1 \* Conditions for decoherence-free dynamics

Let us establish the conditions for decoherence-free dynamical evolution. By definition, a subspace  $\tilde{\mathcal{H}}$  of a Hilbert space  $\mathcal{H}$  is decoherent-free if the evolution inside  $\tilde{\mathcal{H}}$  is unitary. We point out that this definition of decoherence-free subspace does not rule out the possible presence of unitary errors in quantum computation. Such errors may result from inaccurate implementations of quantum logic gates. For instance, in ion-trap quantum processors laser pulses are used to implement sequences of quantum gates and fluctuations in the duration of each pulse induce unitary errors, which accumulate during a quantum computation.

We first formulate the conditions for decoherence-free dynamics in terms of the Hamiltonian description for a system in interaction with a reservoir. Following Sec. 7.4.1, we can write the most general Hamiltonian describing such a situation as

$$H = H_S \otimes I_R + I_S \otimes H_R + H_{SR} = H_S \otimes I_R + I_S \otimes H_R + \sum_i \sigma_i \otimes B_i, \quad (9.50)$$

where  $H_S$  and  $H_R$  describe the system and the reservoir and the operators  $\sigma_i$  and  $B_i$  act on the system and on the reservoir, respectively.

A decoherence-free subspace is found by assuming that there exists a set of eigenvectors  $\{|\tilde{k}\rangle\}$  of the operators  $\sigma_i$  such that

$$\sigma_i |\tilde{k}\rangle = c_i |\tilde{k}\rangle, \quad (9.51)$$

for any  $i, |\tilde{k}\rangle$ . Note that the eigenvalues  $c_i$  are *degenerate* since they depend only on the index  $i$  of the operator  $\sigma_i$  and not on  $\tilde{k}$ . If we limit ourselves to considering the subspace  $\tilde{\mathcal{H}}$  spanned by the states  $\{|\tilde{k}\rangle\}$ , we can write the Hamiltonian as

$$\tilde{H} = H_S \otimes I_R + I_S \otimes \left[ H_R + \sum_i c_i B_i \right], \quad (9.52)$$

where  $\tilde{H}$  is the restriction of  $H$  to  $\tilde{\mathcal{H}}$ . If we assume that the system Hamiltonian  $H_S$  leaves the Hilbert subspace  $\tilde{\mathcal{H}}$  invariant and if the initial state resides in  $\tilde{\mathcal{H}}$ , then the evolution of the system is decoherence-free.

To show this, we assume that at time  $t = 0$  the system and the environment are not entangled. Then, the initial system plus environment state may be written as

$$\rho_{SR}(0) = \rho_S(0) \otimes \rho_R(0), \quad (9.53)$$

where  $\rho_S(0)$  and  $\rho_R(0)$  are the system and environment density matrices at time  $t = 0$ . Moreover, we assume that the initial density matrix describes a state residing in  $\tilde{\mathcal{H}}$ ; that is,

$$\rho_S(0) = \sum_{\tilde{i}, \tilde{j}} s_{\tilde{i}\tilde{j}} |\tilde{i}\rangle \langle \tilde{j}|, \quad (9.54)$$

with  $|\tilde{i}\rangle, |\tilde{j}\rangle \in \tilde{\mathcal{H}}$ . We can also write

$$\rho_R(0) = \sum_{\mu, \nu} r_{\mu\nu} |\mu\rangle \langle \nu|, \quad (9.55)$$

where  $\{|\mu\rangle\}$  is a basis for the Hilbert space of the environment. It is easy to see that temporal evolution does not take the system out of the subspace  $\tilde{\mathcal{H}}$ . Indeed, we have

$$U_{SR}(t)(|\tilde{i}\rangle \otimes |\mu\rangle) = U_S(t)|\tilde{i}\rangle \otimes U_R(t)|\mu\rangle, \quad (9.56)$$

where  $U_{SR}(t) = \exp(-i\tilde{H}t/\hbar)$ ,  $U_S(t) = \exp(-iH_S t/\hbar)$  and  $U_R(t) = \exp[-i(H_R + \sum_j c_j B_j)t/\hbar]$ . Hence, the evolution of the state (9.53) is given by

$$\rho_{SR}(t) = \sum_{\tilde{i}, \tilde{j}} s_{\tilde{i}\tilde{j}} U_S(t)|\tilde{i}\rangle\langle\tilde{j}|U_S^\dagger(t) \otimes \sum_{\mu, \nu} r_{\mu\nu} U_R(t)|\mu\rangle\langle\nu|U_R^\dagger(t). \quad (9.57)$$

It follows that

$$\rho_S(t) = \text{Tr}_R[\rho_{SR}(t)] = U_S(t)\rho_S(0)U_S^\dagger(t), \quad (9.58)$$

and therefore the evolution of the system is unitary (and decoherence-free).

The conditions for decoherence-free dynamics can also be expressed in the framework of the Kraus representation. As we saw in Sec. 7.1, the Kraus operator  $E_\mu$  is defined as  $E_\mu = {}_B\langle\mu|U_{SR}|0\rangle_R$ . The matrix representation of  $E_\mu$  in the basis in which the first states span  $\mathcal{H}$  is given by

$$E_\mu = \begin{bmatrix} g_\mu \tilde{U}_S & 0 \\ 0 & C_\mu \end{bmatrix}, \quad (9.59)$$

where  $g_\mu = {}_R\langle\mu|U_R|0\rangle_R$ ,  $\tilde{U}_S$  is the restriction of  $U_S$  to  $\tilde{\mathcal{H}}$  and  $C_\mu$  is a block matrix acting on the subspace  $\tilde{\mathcal{H}}^\perp$  orthogonal to  $\tilde{\mathcal{H}}$ . Therefore, all Kraus operators  $E_\mu$ , when restricted to a decoherence-free subspace  $\tilde{\mathcal{H}}$ , have an identical unitary representation  $\propto \tilde{U}_S$ , up to a multiplicative constant  $g_\mu$ . The normalization constraint  $\sum_\mu E_\mu^\dagger E_\mu = I_S$  implies  $\sum_\mu |g_\mu|^2 = 1$ . If the initial state  $\rho_S$  resides in the subspace  $\tilde{\mathcal{H}}$ ; that is,

$$\rho_S = \begin{bmatrix} \tilde{\rho}_S & 0 \\ 0 & 0 \end{bmatrix}, \quad (9.60)$$

then the final state  $\rho'_S$  also resides in  $\tilde{\mathcal{H}}$ , and the system's evolution is unitary since

$$\rho'_S = \begin{bmatrix} \sum_\mu |g_\mu|^2 \tilde{U}_S \tilde{\rho}_S \tilde{U}_S^\dagger & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \tilde{U}_S \tilde{\rho}_S \tilde{U}_S^\dagger & 0 \\ 0 & 0 \end{bmatrix}. \quad (9.61)$$

### 9.7.2 \* The spin-boson model

A nice example of decoherence-free dynamics is the spin-boson model, which describes  $n$  spin- $\frac{1}{2}$  particles (the system) interacting with a bosonic field (the reservoir). The interaction Hamiltonian is

$$H_{SR} = \sum_{j=1}^n \sum_k \left[ g_{jk}^+ \sigma_j^+ \otimes b_k + g_{jk}^- \sigma_j^- \otimes b_k^\dagger + g_{jk}^z \sigma_j^z \otimes (b_k + b_k^\dagger) \right], \quad (9.62)$$

where  $\sigma_j^\pm = \sigma_j^x \mp i\sigma_j^y$  and  $\sigma_j^z$  are Pauli operators acting on the  $j$ -th spin,  $b_k$  ( $b_k^\dagger$ ) is the annihilation (creation) operator for the  $k$ -th mode of the bosonic field and  $g_{jk}^\pm$ ,  $g_{jk}^z$  are coupling constants (note that the requirement of Hermitian  $H_{SR}$  implies that  $(g_{jk}^+)^* = g_{jk}^-$ ). This model describes the interaction between a system of qubits (spins) and a bosonic environment, including both dissipative coupling (the terms  $\sigma_j^+ \otimes b_k$  and  $\sigma_j^- \otimes b_k^\dagger$  describe energy exchanges between the system and the environment) and phase-damping processes (through the  $\sigma_j^z \otimes (b_k + b_k^\dagger)$  term).

Let us assume that the coupling constants are independent of the qubit index; that is,  $g_{jk}^\pm \equiv g_k^\pm$  and  $g_{jk}^z \equiv g_k^z$ . This collective decoherence situation is relevant in solid-state systems, provided the coupling to a phononic bath is the dominant source of decoherence and that the wavelength of the relevant phonon modes is much larger than the qubit spacing.<sup>4</sup>

Given the collective decoherence assumption, a decoherence-free subspace exists. Indeed, we can define the total spin operators

$$S_\alpha \equiv \sum_{j=1}^n \sigma_j^\alpha, \quad (9.63)$$

with  $\alpha = +, -, z$ , so that the coupling Hamiltonian becomes

$$H_{SR} = \sum_{\alpha=+,-,z} S_\alpha \otimes B_\alpha, \quad (9.64)$$

where  $B_+ \equiv \sum_k g_k^+ b_k$ ,  $B_- \equiv B_+^\dagger$  and  $B_z \equiv \sum_k g_k^z (b_k + b_k^\dagger)$ . The condition (9.51) for decoherence-free dynamics is fulfilled if we encode the quantum information in *singlet states* ( $S = 0$ ); that is, in states  $|\tilde{k}\rangle$  satisfying

$$S_\alpha |\tilde{k}\rangle = 0, \quad (9.65)$$

for  $\alpha = +, -, z$ .

For the case  $n = 2$ , the only singlet state is

$$\frac{1}{\sqrt{2}} (|01\rangle - |10\rangle). \quad (9.66)$$

For  $n = 4$ , the (singlet) decoherence-free subspace has dimension two and is spanned by the states<sup>5</sup>

$$\begin{aligned} |0_L\rangle &= \frac{1}{2} (|0101\rangle + |1010\rangle - |0110\rangle - |1001\rangle), \\ |1_L\rangle &= \frac{1}{\sqrt{12}} (2|0011\rangle + 2|1100\rangle - |0101\rangle - |1010\rangle - |0110\rangle - |1001\rangle). \end{aligned} \quad (9.67)$$

---

<sup>4</sup>Another physical situation in which the spin-boson model is relevant is the coupling of  $n$  identical two-level atoms to a single mode of the electromagnetic field (for instance, we can consider  $n$  ions in a trap coupled to a microwave field). In this case, provided the wavelength of the radiation field is much longer than the distance between the atoms, Hamiltonian (9.62) reduces to

$$H_{SR} = \sum_{j=1}^n (g^+ \sigma_j^+ \otimes b + g^- \sigma_j^- \otimes b^\dagger).$$

<sup>5</sup>These states can be computed using standard methods for the addition of angular momenta. In general, given two angular momenta  $\mathbf{j}_1$  and  $\mathbf{j}_2$  and the total angular momentum  $\mathbf{J} = \mathbf{j}_1 + \mathbf{j}_2$ , we have

$$|j_1 j_2; JM\rangle = \sum_{m_1, m_2} |j_1 m_1; j_2 m_2\rangle \langle j_1 m_1; j_2 m_2 | j_1 j_2; JM\rangle,$$

Hence, this subspace can be used to encode a single-qubit state. In this manner we can construct the singlet states for progressively higher numbers of qubits. Group theory tells us (see, *e.g.*, Lidar *et al.*, 2000) that the dimension of the (singlet) decoherence-free subspace for  $n$  qubits is

$$\dim[\text{DFS}(n)] = \frac{n!}{(n/2 + 1)! (n/2)!}. \quad (9.68)$$

This subspace can be used to encode  $n_L$  logical qubits, with

$$n_L = \log_2\{\dim[\text{DFS}(n)]\} \approx n - \frac{3}{2} \log_2 n, \quad (9.69)$$

where the right-hand expression is obtained after application of Stirling's formula  $n! \sim \sqrt{2\pi} n^{(n+1/2)} e^{-n}$  for large  $n$ . This means that the *encoding efficiency*  $\epsilon \equiv n_L/n$ , defined as the number of logical qubits  $n_L$  per number of physical qubits  $n$ , tends to unity as  $n \rightarrow \infty$ .

## 9.8 \* Dynamical decoupling

There are alternative strategies to fight against decoherence and dissipation, whose working principle goes under a common framework: the idea is to perform state steering and certain quantum operations in such a way to require the insensitivity of the control strategy with respect to the uncontrolled interactions of the system with the environment.

Here we present the basic principles of the so-called *dynamical decoupling* strategy, which has been proposed by Viola *et al.* (1999). We will discuss a model which can decouple a generic open quantum system from any environmental interaction through simpler “open-loop” control techniques. Specifically, the system can be manipulated by means of a periodic control strategy that relies on a high-frequency, unbounded, control approximation. In this way, it may undergo a wide range of dynamical behaviours, while still eliminating the effects of the environment.

Let us recall once more the standard formalism in order to deal with the most general situation where a given quantum system interacts with an external bath (see Sec. 7.4.1). Our goal is to suppress the interaction  $H_{SR}$  in the global Hamiltonian  $H$  of Eq. (9.50). To control the system evolution, we add to the Hamiltonian  $H$  a time-dependent Hamiltonian  $H_1(t)$  that acts on the system space alone. Let us suppose  $H_1$  to be cyclic in time, that is, the associated propagator

$$U_1(\tau) = U_1(\tau + T_c), \quad \text{with} \quad U_1(t) = \mathcal{T}e^{-\frac{i}{\hbar} \int_0^t H_1(s) ds}, \quad (9.70)$$

is periodic in time, with a period  $T_c > 0$ . This implies that  $U_1(kT_c) = I_S, \forall k \in \mathbb{N}$ .

---

where  $|j_1 j_2; JM\rangle$  and  $|j_1 m_1; j_2 m_2\rangle$  are eigenstates of  $j_1^2, j_2^2, J^2, J_z$  and  $j_1^2, j_{1z}, j_2^2, j_{2z}$ , while the matrix elements  $\langle j_1 m_1; j_2 m_2 | j_1 j_2; JM \rangle$ , usually denoted as  $\langle j_1 j_2 m_1 m_2 | JM \rangle$ , are known as the Clebsch–Gordan coefficients. Note that the conditions  $M = m_1 + m_2$ ,  $|M| \leq J$ ,  $|j_1 - j_2| \leq J \leq j_1 + j_2$  must be fulfilled. The four-qubit singlet states of (9.67) are obtained as combination of the two-qubit singlet ( $|s\rangle \equiv |j=0, m=0\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ ) and triplet ( $|t_+\rangle \equiv |j=1, m=1\rangle = |00\rangle$ ,  $|t_0\rangle \equiv |j=1, m=0\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ ,  $|t_-\rangle \equiv |j=1, m=-1\rangle = |11\rangle$ ) states, with the correct Clebsch–Gordan coefficients:  $|0_L\rangle \equiv |j_1=0, j_2=0; J=0, M=0\rangle = |s\rangle_{12} \otimes |s\rangle_{34}$  and  $|1_L\rangle \equiv |j_1=1, j_2=1; J=0, M=0\rangle = \frac{1}{\sqrt{3}}(|t_+\rangle_{12} \otimes |t_-\rangle_{34} - |t_0\rangle_{12} \otimes |t_0\rangle_{34} + |t_-\rangle_{12} \otimes |t_+\rangle_{34})$ .

It is now convenient to switch to the interaction picture associated with  $H_1$  and consider the time-dependent change of basis in  $H_S$  induced by  $U_1(t)$ . Then, the total Hamiltonian reads

$$\tilde{H}(t) = \sum_i U_1^\dagger(t) \sigma_i U_1(t) \otimes B_i. \quad (9.71)$$

It is possible to show (see Viola *et al.*, 1999) that the associated propagator  $\tilde{U}(t)$  for  $t = T_c$  can be conveniently written as

$$\tilde{U}(T_c) \approx e^{-\frac{i}{\hbar} \bar{H} T_c}, \quad \text{where} \quad \bar{H} = \frac{1}{T_c} \int_0^{T_c} \tilde{H}(s) ds. \quad (9.72)$$

This approximation is valid in the limit where a large number  $N \rightarrow \infty$  of control cycles are applied in  $[0, T]$ , with  $T = NT_c$ , given the time horizon  $T$  for decoupling. In this way, the motion of the system under the time-dependent Hamiltonian  $H_1(t)$  has been replaced by a stroboscopic development under an effective average Hamiltonian  $\bar{H}$ .

Suppose now that we choose a piecewise constant control propagator generated, for example, by a discrete temporal sequence of impulsive control Hamiltonian:

$$U_1(t) = G_j, \quad j \Delta t \leq t \leq (j+1) \Delta t, \quad (9.73)$$

with  $\mathcal{G} = \{G_j\}_{j=1}^{n_g}$  being a finite group made of a set of unitary operators and  $\Delta t = T_c/n_g$ . Thus we have:

$$\bar{H} = \frac{1}{T_c} \int_0^{T_c} \left( \sum_i U_1^\dagger(t) \sigma_i U_1(t) \otimes B_i \right) ds = \sum_i \left( \frac{1}{n_g} \sum_{j=1}^{n_g} G_j^\dagger \sigma_i G_j \right) \otimes B_i. \quad (9.74)$$

The expression in brackets defines the projection of every  $\sigma_i$  onto the commutant of the group algebra associated with  $\mathcal{G}$ , that is the set of the operators commuting with every operator in the group algebra. If such group algebra coincides with the whole space of operators defined on the system's Hilbert space, the commutant only consists of scalar matrices  $\lambda I_S$ ,  $\lambda \in \mathbb{R}$ . This means that the effective, average Hamiltonian  $\bar{H}$  on a control cycle is reduced to a scalar matrix. Moreover, at every integer multiple of  $T_c$  the interaction frame coincides with the initial frame, since  $U_1(kT_c) = e^{i\lambda} I_S$ . In the rotating frame introduced here, this leaves the initial state of the system unchanged (up to an irrelevant global phase factor), thus stroboscopically stabilizing a given arbitrary initial state.

In conclusion, this decoherence control strategy enables to average out the interactions with the environment in a way such that the mean effects on the system of interest are negligible in the fast control cycle approximation. The dynamical decoupling strategy can be applied not only to remove system-bath interactions, but also undesired components of the free system dynamics. It suffices to consider the system Hamiltonian terms  $H_S \otimes I_B$  as the disturbance to be rejected.

### 9.8.1 \* Explicit form of control Hamiltonian

On the basis of what learned above, we now describe a procedure to find a general control Hamiltonian which is able to completely decouple the system from the environment. We follow the approach adopted by Ticozzi and Ferrante (2006). To this aim, let us start from a standard system-bath coupling of the form in Eq. (9.50):

$$H_{SR} = \sum_i \sigma_i \otimes B_i, \quad \text{with constraint } [\sigma_i, \sigma_j] = 0 \quad \forall i, j. \quad (9.75)$$

Then there exists a control Hamiltonian of the form  $u(t)H_d$  that is sufficient to ensure the decoupling, provided  $u(t)$  is a given periodic, impulsive control function. In order to construct that, we notice that each  $\sigma_i$  can be always decomposed as

$$\sigma_i = \frac{\text{Tr}(\sigma_i)}{n} I_S + Z_i \quad \text{with } \text{Tr}(Z_i) = 0, \quad (9.76)$$

where  $n$  is the dimension of the system. It is sufficient to restrict our attention to  $Z_i$  and move to a basis in which each  $Z_i$  is simultaneously diagonal. Choosing without loss of generality  $Z_0 = I_S$ , from an inspection of Eq. (9.74) we can see that an  $n$ -step decoupling strategy must guarantee that

$$Z_i + G_1^\dagger Z_i G_1 + G_2^\dagger Z_i G_2 + \dots + G_{n-1}^\dagger Z_i G_{n-1} = 0. \quad (9.77)$$

The most general zero-trace diagonal  $Z$  is of the form:

$$Z = \begin{bmatrix} \alpha_1 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \alpha_{n-1} & 0 \\ 0 & 0 & \dots & 0 & -\sum_i \alpha_i \end{bmatrix} \equiv \text{diag}\left\{\alpha_1, \dots, \alpha_{n-1}, -\sum_i \alpha_i\right\}. \quad (9.78)$$

Thus we can consider the  $n - 1$  unitary operations  $G_j$  corresponding to cyclic permutations of the eigenvalues, namely, a set of circulant matrices like

$$G_1 = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \quad G_2 = (G_1)^2, \quad \dots, \quad G_{n-1} = (G_1)^{n-1}. \quad (9.79)$$

With these unitary control actions and considering a general  $Z$ , Eq. (9.77) becomes

$$\begin{aligned} & \text{diag}\left\{\alpha_1, \dots, \alpha_{n-1}, -\sum_i \alpha_i\right\} + \text{diag}\left\{-\sum_i \alpha_i, \alpha_1, \dots, \alpha_{n-1}\right\} + \\ & \dots + \text{diag}\left\{\alpha_2, \dots, \alpha_{n-1}, -\sum_i \alpha_i, \alpha_1\right\} = 0. \end{aligned} \quad (9.80)$$

Having made this choice for  $Z_i$ , it can be seen that the desired unitary control actions  $G_i$  can be realized by choosing  $H_d$  as control Hamiltonian, and a impulsive, periodic control function of the form:

$$u(t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left[ \Theta\left(t - \frac{(k-1)T_c}{n}\right) - \Theta\left(t - \frac{(k-1)T_c}{n} - \tau\right) \right], \quad (9.81)$$

for  $t \in [t_0 + (k-1)T_c/n, t_0 + kT_c/n]$ , with  $\Theta(\cdot)$  being the Heaviside function.

Let us now explicitly provide a simple construction of the desired unitary circulant operation as a sequence of standard two qubits unitary gates, in the simplest scenario of a qubit system that is coupled to an environment constituted of another two-level quantum system. We are going to show that the solution emerging from the proposed strategy becomes equivalent to the so-called spin-echo technique. We assume that the Hamiltonian driving the whole system is given by

$$H = \frac{\omega_1}{2} \sigma_1^z \otimes I_2 + \frac{\omega_2}{2} I_1 \otimes \sigma_2^z + J_z \sigma_1^z \otimes \sigma_2^z, \quad (9.82)$$

where the subscripts  $_1$  and  $_2$  respectively refer to the system and the environment, while  $\omega_j$ ,  $J_z$  are real coupling parameters.

The aim of the control strategy is to freeze the dynamics by averaging out the total Hamiltonian  $H$ . Since the latter is already diagonal in the computational basis, one can implement the four circulant matrices generated by the powers of  $G_1$ . In this two-qubit setting,  $G_1$  can be generated easily by applying a bit-flip  $\sigma_x$  gate to both qubits, followed by a generalized CNOT gate (see Eq. (3.68)):

$$G_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} (\sigma_1^x \otimes \sigma_2^x). \quad (9.83)$$

It is thus easy to show that, for  $G_i = (G_1)^i$ , one has  $\sum_{i=1}^4 G_i^\dagger H G_i = 0$ . We point out that this strategy is intrinsically robust with respect to imperfect knowledge of the Hamiltonian parameters  $\omega_j$  and  $J_z$ .

## 9.9 \* The Zeno effect

The Zeno effect, in its simplest instance, refers to the freezing of the evolution of a quantum state due to frequent measurements. However, as we shall discuss below, the Zeno effect also takes place in systems in which a strong disturbance dominates the temporal evolution of the quantum systems. In general, there is no need to invoke the collapse of the wave function. Even more importantly from the viewpoint of quantum computation, the Zeno effect does not necessarily freeze the dynamics. The system can evolve away from its initial state, although it remains in a “decoherence-free” subspace, which can in principle be appropriately engineered. These issues are discussed in the present section, following the presentation of Facchi and Pascazio (2003).

We first consider a simple example where the Zeno phenomenon is induced by frequent projective measurements. Let  $H$  be the total, time-independent Hamiltonian of a quantum system and  $|\psi(0)\rangle = |a\rangle$  its initial state at time  $t = 0$ . The survival probability  $p(t)$ ; that is, the probability to find the system in the same state  $|a\rangle$  at time  $t$ , is given by

$$p(t) = |\langle a|\psi(t)\rangle|^2 = |\langle a|\exp(-\frac{i}{\hbar}Ht)|a\rangle|^2. \quad (9.84)$$

A short-time expansion yields a quadratic behaviour:

$$p(t) \sim |\langle a|(I - \frac{i}{\hbar}Ht - \frac{1}{2\hbar^2}H^2t^2)|a\rangle|^2 \sim 1 - \frac{1}{\hbar^2}(\langle a|H^2|a\rangle - \langle a|H|a\rangle^2)t^2 = 1 - \frac{t^2}{t_Z^2} \quad (9.85)$$

where

$$t_Z = \frac{\hbar}{\sqrt{\langle a|H^2|a\rangle - \langle a|H|a\rangle^2}} \quad (9.86)$$

is the so-called Zeno time.<sup>6</sup>

If  $N$  projective measurements are performed at time intervals  $\tau = \frac{t}{N}$ ,<sup>7</sup> then the survival probability at time  $t$  is

$$\begin{aligned} p(t) &= [|\langle a|\psi(\tau)\rangle|^2]^N = [p(\tau)]^N = \left[p\left(\frac{t}{N}\right)\right]^N \\ &\sim \left(1 - \frac{t^2}{N^2 t_Z^2}\right)^N \sim \exp\left(-\frac{t^2}{N t_Z^2}\right). \end{aligned} \quad (9.87)$$

If  $N \rightarrow \infty$ , then  $p(t) \rightarrow 1$ ; namely, the evolution is completely frozen. Note that the decay in time of the survival probability (9.87) is exponential: for a given  $\tau$  and  $t = N\tau$  ( $N$  integer),

$$p(t) \sim \exp[-\gamma(\tau)t], \quad (9.88)$$

with the decay rate  $\gamma(\tau) \sim \tau/t_Z^2$ .

We remark that the quantum Zeno effect is a direct consequence of the following mathematical property of the Schrödinger equation (sketched in Fig. 9.9): in a short time  $\delta\tau = t/N = O(1/N)$ , the phase of the wave function evolves as  $O(\delta\tau)$ , while the probability changes by  $O((\delta\tau)^2)$ , so that  $p(t) \sim [1 - O(1/N^2)]^N \rightarrow 1$  when  $N \rightarrow \infty$ .

**Exercise 9.12** Discuss the Zeno effect for a two-level system driven by the Hamiltonian  $H = H_0 + H_{\text{int}}$ , with  $H_0 = \frac{1}{2}\hbar\omega\sigma_z$  and  $H_{\text{int}} = \frac{1}{2}\hbar\Omega\sigma_x$ , the initial state of the system being  $|0\rangle$ .

<sup>6</sup>Note that, if the Hamiltonian  $H$  is divided into free and interaction parts,  $H = H_0 + H_{\text{int}}$ , with the initial state  $|a\rangle$  eigenstate of the free Hamiltonian ( $H_0|a\rangle = \omega_a|a\rangle$ ) and the interaction part off-diagonal in the basis of the eigenstates of  $H_0$ , so that  $\langle a|H_{\text{int}}|a\rangle = 0$ , then the Zeno time is given by  $t_Z = \hbar/\sqrt{\langle a|H_{\text{int}}^2|a\rangle}$  and only depends on the interaction Hamiltonian.

<sup>7</sup>In this example we assume that the measurement is *selective*: we select only the survived component ( $|\psi(\tau)\rangle \rightarrow |a\rangle\langle a|\psi(\tau)\rangle$ ) and stop the others. Note that, as we shall discuss below, the Zeno effect also takes place for non-selective measurements.

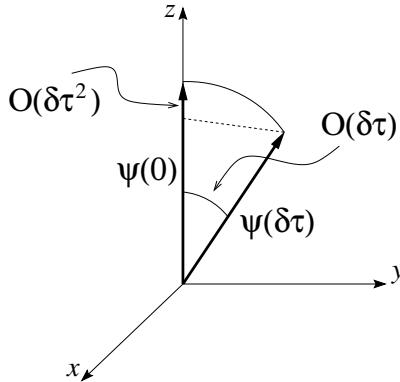


Fig. 9.9 A schematic drawing of the short-time evolution of phase and probability for a wave function whose evolution is governed by the Schrödinger equation.

Note that the collapse of the wave function (an inherently non-unitary and irreversible process) is not necessarily required for the quantum Zeno effect. To illustrate this concept we consider a three-level system governed by the Hamiltonian

$$H = \Omega_1(|0\rangle\langle 1| + |1\rangle\langle 0|) + \Omega_2(|1\rangle\langle 2| + |2\rangle\langle 1|). \quad (9.89)$$

In the basis  $\{|0\rangle, |1\rangle, |2\rangle\}$  this Hamiltonian reads

$$H = \begin{bmatrix} 0 & \Omega_1 & 0 \\ \Omega_1 & 0 & \Omega_2 \\ 0 & \Omega_2 & 0 \end{bmatrix}. \quad (9.90)$$

If the system is prepared at time  $t = 0$  in the state  $|\psi(0)\rangle = |0\rangle$ , then the survival probability is given by (see exercise 9.13)

$$p(t) = |\langle 0|\psi(t)\rangle|^2 = \frac{1}{(\Omega_1^2 + \Omega_2^2)^2} \left[ \Omega_2^2 + \Omega_1^2 \cos\left(\frac{\sqrt{\Omega_1^2 + \Omega_2^2} t}{\hbar}\right) \right]^2. \quad (9.91)$$

Note that for large values of the ratio  $\Omega_2/\Omega_1$  the system is in practice frozen in the level  $|0\rangle$ . In this case, as soon as the system makes a transition from  $|0\rangle$  to  $|1\rangle$  it undergoes a very fast Rabi oscillation to level  $|2\rangle$ . Therefore, we can say that level  $|2\rangle$  acts as a measuring apparatus: when the ratio  $\Omega_2/\Omega_1$  is large, then a better observation of the state of the system is performed, thus hindering the transition  $|0\rangle \rightarrow |1\rangle$ . We stress that the measurement performed by level  $|2\rangle$  is continuous (there is no wave-function collapse) and Hermitian (the model is purely Hamiltonian).

### Exercise 9.13 Prove Eq. (9.91).

The following theorem provides a very general formulation of the quantum Zeno effect. Let us consider a quantum system whose states reside in the Hilbert space  $\mathcal{H}$ . The evolution of the system density matrix  $\rho$  is described by the superoperator

$$\rho(t) = \mathcal{U}_t \rho(0) = U(t) \rho(0) U^\dagger(t), \quad U(t) = \exp\left(-\frac{i}{\hbar} H t\right), \quad (9.92)$$

where  $H$  is a time-independent bounded Hamiltonian. We also introduce a set of projectors  $P_i$  such that  $P_i P_j = \delta_{ij} P_i$  and  $\sum_i P_i = I$ . The subspaces relative to the operators  $P_i$  are denoted by  $\mathcal{H}_i = P_i \mathcal{H}$  ( $\mathcal{H} = \bigoplus_i \mathcal{H}_i$ ). We consider a non-selective measurement (the measuring apparatus does not select the different outcomes) described by the superoperator

$$\mathcal{P}\rho = \sum_n P_n \rho P_n. \quad (9.93)$$

The evolution of the system after  $N$  such measurements performed in a time  $t$  is determined by the superoperator

$$\mathcal{S}_t^{(N)} = (\mathcal{P} \mathcal{U}_{t/N})^N \mathcal{P}, \quad (9.94)$$

where we have also operated a first measurement at time  $t = 0$ , which prepares the state  $\mathcal{P}\rho(0) = \sum_i P_i \rho(0) P_i$ . The evolution of the system density matrix reads

$$\rho(t) = \sum_i W_i(t) \rho_0 W_i^\dagger(t), \quad (9.95)$$

with  $W_i^\dagger(t) W_i(t) = P_i$ . The probability to find the system in the subspace  $\mathcal{H}_i$  is

$$p_i(t) = \text{Tr}[\rho(t) P_i] = \text{Tr}[W_i(t) \rho_0 W_i^\dagger(t)] = \text{Tr}[\rho_0 P_i] = p_i(0). \quad (9.96)$$

It is clear from Eqs. (9.95) and (9.96) that any interference term between the different subspaces  $\mathcal{H}_i$  is destroyed and that the probability is conserved in each subspace. Each operator  $W_i(t)$  is unitary within the subspace  $\mathcal{H}_i$  and has the form

$$W_i(t) = P_i \exp\left(-\frac{i}{\hbar} P_i H P_i t\right). \quad (9.97)$$

The above theorem is interesting from the viewpoint of quantum computation, as it suggests strategies to contrast decoherence. A simple example will help clarify this concept. We consider again Hamiltonian (9.90) and the projectors

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (9.98)$$

satisfying  $P_1 + P_2 = I$ . The subspace  $\mathcal{H}_1 = P_1 \mathcal{H}$  is the two-dimensional subspace (qubit) of interest for quantum computation, while the coupling  $\Omega_2$  mimics decoherence. In the limit  $N \rightarrow \infty$  the operators (9.97) become

$$\begin{aligned} W_1(t) &= P_1 \exp\left(-\frac{i}{\hbar} P_1 H P_1 t\right) = P_1 \exp\left\{-\frac{i}{\hbar} \begin{bmatrix} 0 & \Omega_1 t & 0 \\ \Omega_1 t & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right\} \\ &= \begin{bmatrix} \cos\left(\frac{\Omega_1 t}{\hbar}\right) & -i \sin\left(\frac{\Omega_1 t}{\hbar}\right) & 0 \\ -i \sin\left(\frac{\Omega_1 t}{\hbar}\right) & \cos\left(\frac{\Omega_1 t}{\hbar}\right) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned} \quad (9.99)$$

$$W_2(t) = P_2 \exp\left(-\frac{i}{\hbar} P_2 H P_2 t\right) = P_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the qubit evolves according to the Hamiltonian

$$P_1 H P_1 = \begin{bmatrix} 0 & \Omega_1 & 0 \\ \Omega_1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (9.100)$$

The subspaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  decouple; that is, the evolution of the qubit becomes decoherence-free. Finally, we point out that other Zeno strategies based on unitary disturbances (instead of projective measurements) of the system that we wish to protect are also possible (see Facchi and Pascazio, 2003).

## 9.10 Fault-tolerant quantum computation

So far, our discussion of quantum error correction has assumed that encoding, decoding of quantum information and error recovery operations can be achieved perfectly. However, these are complex quantum computations subject to errors. Moreover, quantum logic gates performed in quantum information processing may propagate errors in the quantum computer. In spite of these difficulties, we shall show that, under certain assumptions, arbitrarily long quantum computation can, in principle, be performed reliably, provided the noise in individual quantum gates is below a critical threshold. A quantum computer that performs reliably even in the presence of imperfections is said to be *fault-tolerant*. Sophisticated techniques have been developed for the construction of fault-tolerant quantum circuits (for a review see, e.g., Preskill, 1998b). In the following, we shall limit ourselves to illustrate the basic principles of fault-tolerant quantum computation.

### 9.10.1 Avoidance of error propagation

If an error affects one qubit and this qubit interacts with another in order to perform a two-qubit gate, then the error is likely to propagate to the second qubit. To grasp this point, it is sufficient to consider the CNOT gate. If a bit-flip error affects the control qubit, then the error also spreads to the target qubit. For instance, we consider  $\text{CNOT}(|0\rangle|0\rangle) = |0\rangle|0\rangle$ . If there is a bit-flip error affecting the control qubit ( $|0\rangle \leftrightarrow |1\rangle$ ), then  $\text{CNOT}(|1\rangle|0\rangle) = |1\rangle|1\rangle$ , so that both the control and the target qubit are flipped. A more subtle, purely quantum effect, is the *backward sign propagation*, discussed in Sec. 3.5 (see exercise 3.14): a phase error affecting the target qubit is also transferred, after application of the CNOT gate, to the control qubit.

The backward sign propagation problem spoils the efficiency of the error-correcting quantum circuits shown earlier in this chapter. If we assume that the probabilities of errors affecting one and two qubits are  $O(\epsilon)$  and  $O(\epsilon^2)$ , respectively,<sup>8</sup> then a single-qubit error-correcting code is useful when it lowers the error probability from  $O(\epsilon)$  to  $O(\epsilon^2)$ . This is not the case, for example, for the circuit drawn

---

<sup>8</sup>This is the case, for instance, when errors affecting different qubits are completely uncorrelated with one another.

in Fig. 9.6 once phase errors affecting the ancillary qubits are taken into consideration. The problem is that we use a single ancillary qubit for more than one CNOT gate. This is clear from Fig. 9.10, which contains the basic building block for error extraction. If, with probability  $O(\epsilon)$ , a phase error affects the bottom qubit in the left-hand circuit of Fig. 9.10 before the application of the two CNOT gates, then this error spreads to two of the qubits used to encode the quantum information. Therefore, a code able to correct a single-qubit error (such as Shor's nine-qubit code) fails with  $O(\epsilon)$  probability. This problem is avoided by the right-hand circuit in Fig. 9.10, which employs each ancillary qubit only once. Therefore, a phase error affecting a single ancillary qubit is propagated only to a single qubit. Then, the probability of having two phase errors transferred from the ancillary qubits to the qubits used for encoding is  $O(\epsilon^2)$ . We say that the right-hand circuit in Fig. 9.10 is fault-tolerant, while the left-hand circuit is not. More generally, a quantum code correcting up to  $t$  errors is said to be fault-tolerant if its failure probability is  $O(\epsilon^{t+1})$ .

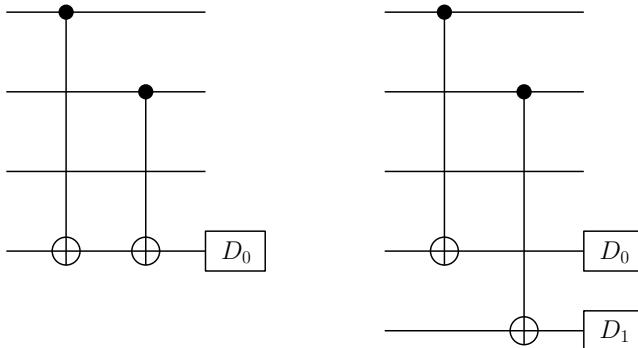


Fig. 9.10 Quantum circuits for extracting the error syndrome. The left-hand circuit employs the same ancillary qubit twice and is therefore not fault-tolerant. In contrast, the right-hand circuit is fault-tolerant.

We point out that the ancillary qubits must be prepared in an appropriate initial state. If we prepare them in the usual state  $|00\rangle$ , consider a generic encoded initial state  $\alpha|000\rangle + \beta|111\rangle$ , and assume that a bit-flip error has corrupted the first qubit, then the fault-tolerant circuit in Fig. 9.10 maps the initial state

$$(\alpha|100\rangle + \beta|011\rangle)|00\rangle \quad (9.101)$$

onto

$$\alpha|10010\rangle + \beta|01101\rangle. \quad (9.102)$$

Therefore, the measurements of the two ancillary qubits projects the five-qubit state onto  $|10010\rangle$  (with probability  $|\alpha|^2$ ) or  $|01101\rangle$  (with probability  $|\beta|^2$ ). In both case, since one of the two ancillary qubits changed its state from  $|0\rangle$  to  $|1\rangle$ , we may conclude that a bit-flip error affected the first or the second qubit. However, this procedure is not adequate, as we have destroyed the quantum information encoded in the superposition of the states  $|000\rangle$  and  $|111\rangle$ .

To solve this problem, we prepare the ancillary qubits in the equally weighted superposition of the states  $|00\rangle$  and  $|11\rangle$ . Therefore, the initial state

$$(\alpha|100\rangle + \beta|011\rangle) \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (9.103)$$

is mapped by the right-hand circuit in Fig. 9.10 onto

$$(\alpha|100\rangle + \beta|011\rangle) \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle). \quad (9.104)$$

The measurement of the ancillary qubits gives us, with equal probabilities, outcome 01 or 10. In both cases, we can conclude that a bit-flip error affected the first or the second qubit, without destroying the quantum superposition  $\alpha|100\rangle + \beta|011\rangle$ .

### 9.10.2 Fault-tolerant quantum gates

In order to implement a reliable quantum computation, we must apply fault-tolerant quantum gates. This is possible if we perform quantum logic operations directly on encoded states.

A fault-tolerant CNOT gate is shown in Fig. 9.11. In this quantum circuit, the first three physical qubits encode the control and the last three physical qubits the target, according to the rule  $|0_L\rangle = |000\rangle$  and  $|1_L\rangle = |111\rangle$ . It is easy to show that, if the CNOT gates are applied *transversally* (that is, bitwise), as shown in Fig. 9.11, then the truth table of the CNOT gate is verified for the logical qubits. Indeed, starting from the six-qubit state  $|x_L\rangle|y_L\rangle$ , with  $x_L, y_L = 0, 1$ , we obtain at the end of the circuit  $|x_L\rangle|x_L \oplus y_L\rangle$ . We point out that the CNOT gate is implemented fault-tolerantly, because each qubit in each code block is involved in a single quantum gate. Therefore, errors in one block can propagate at most to one qubit in the other block, not inside the same block and this construction of the CNOT gate is thus fault-tolerant.

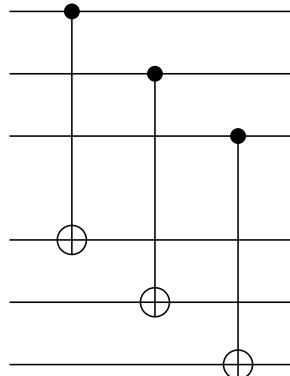


Fig. 9.11 A quantum circuit implementing a transversal CNOT gate between two logical qubits encoded in three-qubit blocks.

We note that it is possible to find a universal set of fault-tolerant quantum gates, in terms of which any quantum computation may be expressed.

### 9.10.3 The noise threshold for quantum computation

The threshold theorem for quantum computation tells us that, given certain assumptions about the noise model (in the simplest case, we consider random and uncorrelated errors) and provided the noise affecting individual quantum gates is below a certain threshold, then it is in principle possible to efficiently implement arbitrarily long quantum computations.

The key ingredient for this result is the use of *concatenated codes*. To understand this concept, let us consider the encoding of a single logical qubit in a block of  $n = 7$  qubits.<sup>9</sup> In a concatenated code each qubit of the block is itself a 7-qubit block, and so on (see Fig. 9.12). If there are  $L$  levels of concatenation, then a single logical qubit is encoded into  $n^L = 7^L$  physical qubits.

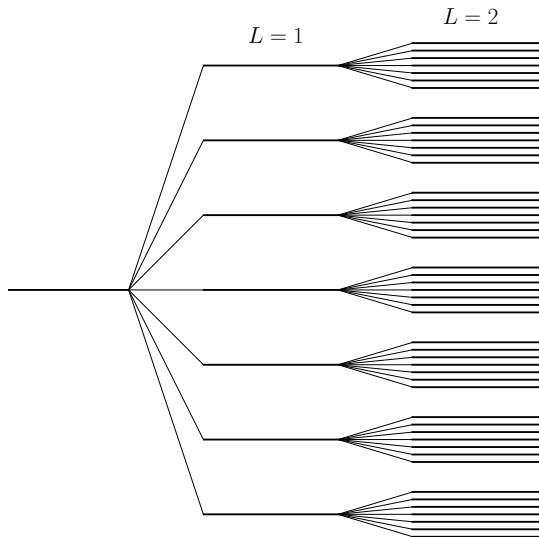


Fig. 9.12 Concatenation of a 7-qubit code up to the  $L = 2$  level.

Let us call  $\epsilon$  the error probability per qubit per appropriate unit of time (for instance, the time required to implement a single elementary quantum gate) and  $\alpha$  the number of locations in the quantum circuit where an error can affect a single qubit before that error correction is applied. Typically, for quantum gates such as the fault-tolerant CNOT and for codes correcting a single error, assuming that error correction is applied after each fault-tolerant quantum gate, we obtain  $\alpha \sim 10^2$ . Error correction at the first level of encoding ( $L = 1$ ) fails if at least two qubits

<sup>9</sup>This is the number of qubits required in the CSS quantum code, developed by Calderbank and Shor (1996) and Steane (1996b).

have been corrupted. Therefore, the failure probability is

$$p_1 \approx c\epsilon^2 \approx \alpha^2\epsilon^2, \quad (9.105)$$

where  $c \approx \alpha^2$  is the number of ways in which a fault-tolerant circuit can introduce at least two errors. At the second level of encoding ( $L = 2$ ), we employ  $n^2$  qubits and error correction fails if at least two of the subblocks of size  $n$  fail. Thus, the failure probability is

$$p_2 \approx cp_1^2 \approx \alpha^2(\alpha^2\epsilon^2)^2. \quad (9.106)$$

We can iterate this procedure. The failure probability at level- $L$  concatenation is

$$p_L \approx cp_{L-1}^2 \approx \frac{(\alpha^2\epsilon)^{2^L}}{\alpha^2}. \quad (9.107)$$

If we wish to implement a computation of length  $T$  ( $T$  denotes the number of logic quantum gates) with accuracy  $\epsilon_0$ , then the error probability per logic gate must be  $\leq \epsilon_0/T$ . Thus, we must concatenate our code a number of times  $L$  such that

$$p_L \approx \frac{(\alpha^2\epsilon)^{2^L}}{\alpha^2} \leq \frac{\epsilon_0}{T}. \quad (9.108)$$

Provided  $\epsilon < \epsilon_{\text{th}} \equiv 1/\alpha^2$ , this inequality is fulfilled for

$$L > \bar{L} \approx \log \left[ \frac{\log(T/\alpha^2\epsilon_0)}{\log(1/\alpha^2\epsilon)} \right]. \quad (9.109)$$

The number of physical qubits  $\bar{n}_{\text{tot}} = n^{\bar{L}}$  required to achieve this level of accuracy is

$$\bar{n}_{\text{tot}} \approx \left[ \frac{\log(T/\alpha^2\epsilon_0)}{\log(1/\alpha^2\epsilon)} \right]^{\log n}. \quad (9.110)$$

Note that  $\bar{n}_{\text{tot}}$  grows only polylogarithmically with  $T$  and  $1/\epsilon$ .

We stress that the above results assume that the quantum computer hardware is such that many quantum gates can be executed in parallel in a single time step. Otherwise, errors in concatenated codes would accumulate too quickly to allow successful error correction.

Finally, we note that, for  $\alpha \sim 10^2$ , the noise threshold is  $\epsilon_{\text{th}} \sim 10^{-4}$ . Various sophisticated calculations found in the literature give different results  $\epsilon_{\text{th}} \sim 10^{-6} - 10^{-4}$ . The numerical value of the noise threshold depends on the assumed characteristics of the quantum computer hardware.

## 9.11 A guide to the bibliography

Quantum error correction was invented by Shor (1995) and Steane (1996a). Tutorials on quantum error correction are Gottesman (2000), Knill *et al.* (2002) and Steane (2006). A very readable introduction is Preskill (1999).

The most complete references on the stabilizer formalism are contained in Preskill (1998a) and in the PhD thesis by Gottesman (1997). The five-qubit code is discussed in Laflamme *et al.* (1996) and Bennett *et al.* (1996b).

A review on decoherence-free subspaces is Lidar and Whaley (2003), while the dynamical decoupling strategy has been first proposed by Viola and Lloyd (1998) and subsequently formalized in Viola *et al.* (1999). A useful reference on the link between the quantum Zeno and decoherence-free subspaces is Facchi *et al.* (2004). Both passive quantum error-avoiding codes and active quantum error correction can be treated in a unified picture based on the so-called noiseless subsystems, see Knill *et al.* (2000) and Viola *et al.* (2001).

A review on fault-tolerant quantum computation is Preskill (1998b).

## Chapter 10

# Principles of experimental implementations of quantum protocols

The great challenge of quantum computation is to experimentally realize a large scale quantum computer. The requirements that must be fulfilled to achieve this imposing objective are summarized in DiVincenzo (2000a):

- (1) A scalable system with well characterized qubits.
- (2) The ability to initialize ('reset') the state of the qubits to a fiducial state (such as  $|0 \cdots 0\rangle$ ).
- (3) Long significant decoherence times, much longer than the gate operation time.
- (4) An experimentally feasible universal set of quantum gates.
- (5) A high-fidelity readout method.

It should be remarked that these requirements are to some extent conflicting: we desire the quantum computer to be well isolated from the environment to preserve its coherence and at the same time we must interact with it strongly to prepare the initial state, realize the desired unitary evolution and measure the final state. The problem here is that external control operations typically introduce noise into the computer, thus disturbing the programmed coherent evolution. An important question is how large should the ratio be between the decoherence time  $\tau_d$  and the "clock time" of the quantum computer; that is, the time  $\tau_g$  for the execution of a quantum gate. The answer is that the ratio  $\tau_d/\tau_g$  should be large enough to allow quantum error correction. As discussed in Sec. 9.10.3, the threshold value for fault-tolerant quantum computation depends on the characteristics of the quantum hardware. However, optimistic estimates require  $\tau_d/\tau_g > 10^4$ , an extremely demanding requirement, corresponding to less than one error in  $10^4$  quantum gate operations.

The five requirements above are sufficient for quantum computation. However, we are also interested in the implementation of quantum communication protocols. For this purpose, two more items must be added to the list of requirements:

- (1) The ability to interconvert "stationary" and "flying" qubits.
- (2) Faithful transmission of flying qubits between specified locations.

Using the terms “stationary” and “flying” qubits we emphasize the fact that the physical systems (in practice, photons) used to transmit qubits from place to place are very different from the qubits used for reliable local computation (for instance, two-level atoms or ions). The development of interfaces between quantum information carriers and quantum information storage and processors is an important objective in the development of quantum technologies.

On the other hand, the last requirement alone is sufficient for quantum cryptography, which deals with one qubit (or one Bell state) at a time and not with complex many-qubit systems as in the case of quantum computation. For this reason, quantum cryptography is the first quantum-information protocol that is having commercial applications.

In this chapter, we shall discuss the basic principles guiding the physical realizations of few-qubit quantum computers in different physical systems (cavity quantum electrodynamics, trapped ions and solid-state qubits) as well as quantum cryptography with photons. We shall not dwell on the technical aspects of the implementations but present instead the basic physical ideas underlying the development of these first quantum machines.

## 10.1 Cavity quantum electrodynamics

The wording cavity quantum electrodynamics (cavity QED) denotes a set of techniques allowing the interaction of single atoms and single photons inside a resonating cavity. Here we focus on experiments performed with *Rydberg atoms*; that is, atoms whose valence electrons are in states with a very large principal quantum number  $n$ . More precisely, we consider alkali atoms, which have a single valence electron which is highly excited up to  $n \sim 20 - 50$ . In such conditions, the valence electron is very far from the atomic nucleus and therefore its electric dipole moment is very high (see Table 10.1). As a consequence, the intensity of the interaction with an applied electromagnetic field is very high. It is therefore possible to achieve the so-called *strong-coupling regime* in which the coherent evolution of a single atom coupled to a single photon stored in a high-quality cavity overwhelms the incoherent dissipative processes.<sup>1</sup> This allows for atom-photon entanglement to be produced before decoherence dominates. Moreover, the energy separation  $E_n - E_{n-1}$  between two consecutive atomic levels is very low (corresponding to a frequency  $\sim 10 - 50$  GHz, to be compared with optical frequencies  $O(10^{15} \text{ Hz})$  relevant when  $n \sim 1$ ). This entails two important consequences:

- (1) these (radio)frequencies are available in laboratories, so that resonant cavities can be excited and then used to manipulate the atoms;

---

<sup>1</sup>The quality factor  $Q$  is a measure of the rate at which a vibrating system dissipates its energy (a higher  $Q$  indicates a lower rate of energy dissipation). By definition,  $Q$  is  $2\pi$  times the ratio of the energy stored divided by the energy lost per cycle. For instance, a damping time of 115 ms, corresponding to a quality factor  $Q \sim 3 \times 10^{10}$  is reported in Haroche and Raimond (2006). Within this damping time a single photon bounces  $1.2 \times 10^9$  times on the cavity mirrors, travelling over 34000 kilometres, a journey which is of the order of the Earth’s circumference.

- (2) the lifetime of Rydberg atoms is very long (much longer, as shown in Table 10.1, than for atoms at  $n \sim 1$ ).

To give some relevant figures, let us note that the lifetime of a Rydberg atom with  $n \sim 50$  and high angular momentum  $l \sim n$  can be as large as 30 ms and the transition frequency between states with principal quantum numbers  $n$  and  $n - 1$  is in the microwave range.<sup>2</sup> This property is very useful as it allows one to employ resonant cavities of centimetre size, which are very convenient for the experimental manipulation (obviously this applies to systems with a small number of qubits).

Table 10.1 Scaling of physical quantities for Rydberg states.

Physical quantity	Scales as
Binding energy $E_n$	$1/n^2$
$E_n - E_{n-1}$	$1/n^3$
Size	$n^2$
Electric dipole moment	$n^2$
Lifetime (low angular momentum)	$n^3$
Lifetime (high angular momentum)	$n^5$
Critical electric field for ionization	$1/n^4$

The typical experimental setup for cavity QED experiments is sketched in Fig. 10.1. Alkali atoms leave the oven  $O$  and are excited into the desired Rydberg state by means of appropriately tuned laser pulses  $L$ . It is possible to select atoms having a well defined velocity using the Doppler effect. Even though the source emits atoms randomly, pulsed lasers allow one to select the incoming atoms and to know the preparation time for the circular Rydberg states within  $O(\mu\text{s})$  interval. The position of each atom flying inside the apparatus is then known with  $O(\text{mm})$  precision. It is therefore possible to address and control individual atoms. The prepared Rydberg atom crosses one or more cavities (usually, microwave superconducting cavities)  $R_1$ ,  $C$ , and  $R_2$ , resonant with the transition between two atomic levels  $|g\rangle$  and  $|e\rangle$ . The two cavities  $R_1$  and  $R_2$  implement microwave Rabi pulses and are used to prepare the initial state in the desired superposition of the states  $|g\rangle$  and  $|e\rangle$  and to analyze the final state, respectively. Note that the atom can be treated as a two-level system (qubit) since it is prepared in  $R_1$  in a superposition  $\alpha|g\rangle + \beta|e\rangle$  and the cavities are resonant with the  $|g\rangle \leftrightarrow |e\rangle$  transition. The relevant Hilbert space for the atom is therefore spanned by the  $\{|g\rangle, |e\rangle\}$  basis. The cavity  $C$  can be prepared in the vacuum state  $|0\rangle$  with no photons (the mean photon number can be reduced to less than 0.1) and can evolve to the one-photon state  $|1\rangle$  after interaction with the atom. Note that the photon storage time is  $O(\text{ms})$ , much larger than the atom–cavity interaction time, a few tens of  $\mu\text{s}$ , thus allowing the coherent manipulation of entangled atom–photon states. The up/down ( $|g\rangle/|e\rangle$ ) state of the

<sup>2</sup>Note that the states with maximum angular momentum  $l = n - 1$  (known as circular Rydberg states) are the quantum counterpart of classical circular orbits.

atom is finally measured using the two detectors  $D_g$  and  $D_e$ : the atom is ionized by means of a static electric-field ( $O(10^2)$  V/cm) and the resulting electron is counted. This procedure is very effective as the static electric field threshold for ionization strongly depends on the principal quantum number  $n$ , as shown in Table 10.1. The detectors  $D_g$  and  $D_e$  are state selective: if the atom is in the state  $|g\rangle$  it is ionized by the static field in  $D_g$ , if instead it is in  $|e\rangle$  it is ionized by the field in  $D_e$ . As an example of this technique, the circular Rydberg states for rubidium atoms with  $n = 50$  and  $n = 51$  can be distinguished.

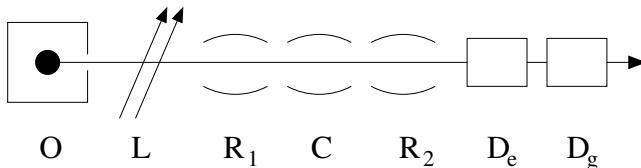


Fig. 10.1 A sketch of a cavity quantum electrodynamics apparatus: the atoms leaving the oven  $O$  are excited into the desired Rydberg state by pulsed lasers  $L$ , enter the cavities  $R_1$ ,  $C$  and  $R_2$  and are finally detected using state selective field ionization in  $D_e$  and  $D_g$ .

We point out that the experimental apparatus sketched in Fig. 10.1 can be seen as the actual implementation of the theoretical procedure described in Sec. 7.2.2 for the measurement of the quantum operation acting on a qubit. The preparation of the initial density matrix  $\rho$  involves  $O$ ,  $L$  and  $R_1$  in Fig. 10.1, while the density matrix  $\rho'$  obtained after interaction with the cavity  $C$  is analyzed through  $R_2$ ,  $D_e$  and  $D_g$ . It is therefore possible to measure the 12 parameters determining the mapping (quantum operation) of the single-qubit density matrix  $\rho$  into  $\rho'$ .

### 10.1.1 Interaction of a two-level atom with a classical field

Note that the fields applied in  $R_1$  and  $R_2$  have relaxation times  $O(\text{ns})$  and therefore do not produce any entanglement between the atom and the microwave radiation field. Indeed, the time required to induce  $|g\rangle \leftrightarrow |e\rangle$  Rabi oscillations is of the order of  $10\ \mu\text{s}$ , much longer than the relaxation time. Hence, we describe the electromagnetic fields in  $R_1$  and  $R_2$  as classical fields. It can be shown (see exercise 10.1) that the action of such a field on a two-level atoms is described, in the  $\{|g\rangle, |e\rangle\}$  basis, by the unitary matrix

$$U = \begin{bmatrix} \cos \frac{\theta}{2} & -ie^{i\phi} \sin \frac{\theta}{2} \\ -ie^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix}, \quad (10.1)$$

where  $\theta$  is proportional to the amplitude of the radiation field and to the atom–field interaction time while  $\phi$  is the phase of the field. Since  $U = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} (\mathbf{n} \cdot \boldsymbol{\sigma})$ , where the unit vector  $\mathbf{n} = (\cos \phi, -\sin \phi, 0)$  and  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ , then  $U$  represents a rotation of the Bloch sphere through an angle  $\theta$  about the axis directed along  $\mathbf{n}$  [see Eq. (3.60)] This axis lies in the  $(x, y)$  plane of the Bloch sphere and forms an

angle  $-\phi$  with the  $x$ -axis. Starting from a given initial state, say  $|\psi_0\rangle = |g\rangle$ , we can obtain the generic state of a qubit,  $|\psi\rangle = U|\psi_0\rangle = \cos\frac{\theta}{2}|g\rangle - ie^{-i\phi}\sin\frac{\theta}{2}|e\rangle$ . Note that, when the atom interacts with a classical field, its state remains pure.

**Exercise 10.1** The evolution of a single alkali atom in a classical electromagnetic field is governed, in the dipole approximation, by the Schrödinger equation

$$\begin{aligned} i\hbar\frac{d}{dt}|\psi(t)\rangle &= (H_0 + H_I)|\psi(t)\rangle, \\ H_0 &= \frac{p^2}{2m} + V(r), \\ H_I &= -ezE(t), \end{aligned} \quad (10.2)$$

where the first term in  $H_0$  is the kinetic energy of the valence electron of the atom and  $V(r)$  the effective potential acting on such electron, generated by the atomic nucleus and the other electrons, and  $H_I$  is due to the interaction of the electron with the electric field generated by a wave linearly polarized along the  $z$ -axis. Solve the Schrödinger equation and in particular derive (10.1) when only two atomic levels are relevant ( $\phi_g(\mathbf{r})$  and  $\phi_e(\mathbf{r})$ , the corresponding energies being  $E_g = \hbar\omega_g$  and  $E_e = \hbar\omega_e$ ) and the electric field is given by

$$E(t) = E_0 \cos(\omega t + \phi). \quad (10.3)$$

To solve this exercise, expand the wave function  $\psi(\mathbf{r}, t) = \langle \mathbf{r} | \psi(t) \rangle$  on the basis of the eigenfunctions of the two-level atom,

$$\psi(\mathbf{r}, t) = \sum_{i=g,e} c_i(t) \phi_i(\mathbf{r}) \exp\left(-i\frac{E_i}{\hbar}t\right). \quad (10.4)$$

Then derive the equations

$$\begin{cases} i\hbar\dot{c}_g(t) = c_e(t) \left[ D\alpha e^{i(\omega-\omega_0)t} + D^\star\alpha^* e^{-i(\omega+\omega_0)t} \right], \\ i\hbar\dot{c}_e(t) = c_g(t) \left[ D^\star\alpha e^{i(\omega+\omega_0)t} + D\alpha^* e^{-i(\omega-\omega_0)t} \right], \end{cases} \quad (10.5)$$

where we have defined

$$D = -e \int d\mathbf{r} \phi_g^*(\mathbf{r}) z \phi_e(\mathbf{r}), \quad (10.6)$$

$\alpha = \frac{1}{2}E_0 e^{i\phi}$ , and  $\omega_0 = \omega_e - \omega_g$ . Then neglect the rapidly oscillating terms depending on  $\omega + \omega_0$  (*rotating wave approximation* (RWA)) to finally derive (10.1).

### 10.1.2 The Jaynes–Cummings model

The electromagnetic field in the cavity  $C$  must be considered as a quantum object. Within the dipole approximation, the interaction between a two-level atom and a single mode of the quantized electromagnetic field is modelled by the Rabi Hamiltonian (7.111) (see, *e.g.*, Meystre and Sargent III, 2007). After neglecting the counter-rotating terms proportional to  $a^\dagger\sigma_+$  (which simultaneously excites the

two-level atom and creates a photon), and to  $a\sigma_-$  (which de-excites the two-level atom and annihilates a photon), we obtain the *Jaynes-Cummings Hamiltonian*

$$H_{JC} = -\frac{1}{2}\hbar\omega_0\sigma_z + \hbar\omega(a^\dagger a + \frac{1}{2}) + \hbar(\lambda\sigma_+a + \lambda^*\sigma_-a^\dagger), \quad (10.7)$$

where the Pauli matrices are written in  $\{|g\rangle, |e\rangle\}$  basis of the eigenstates of the two-level atom, corresponding to the energies  $E_g$  and  $E_e$ ,  $\hbar\omega_0 = E_e - E_g$ , the raising and lowering operators  $\sigma_+$  and  $\sigma_-$  are such that  $\sigma_+|g\rangle = |e\rangle$ ,  $\sigma_+|e\rangle = 0$ ,  $\sigma_-|g\rangle = 0$ ,  $\sigma_-|e\rangle = |g\rangle$ ,  $\hbar\omega$  the single-photon energy,  $\hbar\omega(a^\dagger a + \frac{1}{2})$  the Hamiltonian describing a single mode of the electromagnetic field,  $a^\dagger$  and  $a$  the photon creation and annihilation operators:  $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$ ,  $a|n\rangle = \sqrt{n}|n-1\rangle$ ,  $n$  being the number of photons in the cavity. Therefore, the operator  $\sigma_-a^\dagger$  de-excites the atom ( $|e\rangle \rightarrow |g\rangle$ ) and creates a photon by means of the operator  $a^\dagger$  while  $\sigma_+a$  represents the excitation of the atom by absorption of a photon.

To understand why the counter-rotating terms can be neglected, we can consider the free evolution ( $\lambda = 0$ ) of these operators in the Heisenberg picture. Since for the field annihilation and creation operators we have

$$a(t) = a(0)e^{-i\omega t}, \quad a^\dagger(t) = a^\dagger(0)e^{i\omega t}, \quad (10.8)$$

and similarly for the qubit lowering and raising operators

$$\sigma_\pm(t) = \sigma_\pm(0)e^{\pm i\omega_0 t}, \quad (10.9)$$

we obtain for the counter-rotating terms

$$\sigma_-(t)a(t) = \sigma_-(0)a(0)e^{-i(\omega_0+\omega)t}, \quad \sigma_+(t)a^\dagger(t) = \sigma_+(0)a^\dagger(0)e^{i(\omega_0+\omega)t}. \quad (10.10)$$

Hence, these terms evolve “rapidly”, at frequency  $\omega_0 + \omega$ . In contrast, the terms  $\sigma_+a$  and  $\sigma_-a^\dagger$  obey

$$\sigma_+(t)a(t) = \sigma_+(0)a(0)e^{i(\omega_0-\omega)t}, \quad \sigma_-(t)a^\dagger(t) = \sigma_-(0)a^\dagger(0)e^{-i(\omega_0-\omega)t}, \quad (10.11)$$

and therefore vary slowly near resonance, that is, for  $\omega \approx \omega_0$ . The rapidly oscillating counter-rotating terms can be dropped if the time scale relevant for the dynamics of the system are long enough to average out their effect. In cavity QED, a typical time scale of interest is the inverse of the Rabi frequency of the oscillations  $|g, 1\rangle \leftrightarrow |e, 0\rangle$ , which is given (see below) by  $\Omega = |\lambda|$ . For an atom residing in a resonant cavity  $\Omega$  is typically  $10^{-6}$  of the atomic frequency  $\omega_0$  and of the cavity frequency  $\omega \approx \omega_0$ , and therefore the RWA yields a good description of the system. We shall discuss later in this chapter that this is not necessarily the case for circuit QED experiments, and in that case the Rabi model rather than the Jaynes-Cummings model must be used.

### 10.1.3 Rabi oscillations

As we know, Rabi oscillations consist in the variation of the population of levels (the eigenstates of the system Hamiltonian) induced by an external field, which can be either classical or quantized. The theory of Rabi oscillations is developed in

details in exercises 10.1 (for a classical field) and 10.2 (for a quantized field). Here we are interested in the case in which the electromagnetic field is quantized, so that quantum information can be transferred from the atom to the field and *vice versa*. This is possible in cavity QED experiments where a two-level atom interacts with a cavity field prepared with a given (small) number of photons. Such states  $|n\rangle$  with fixed photon number are known as *Fock states* or *number states*. In particular, the ground state  $|0\rangle$  of the quantum field is the so-called *vacuum state*, in which the photon number is equal to zero. Rabi oscillations between the atom–cavity states  $|g, n\rangle$  (*i.e.*, atomic state  $|g\rangle$  and  $n$  photons in the cavity) and  $|e, n-1\rangle$  are described by the Jaynes–Cummings model. The frequency of these oscillations is proportional to the coupling constant  $|\lambda|$  and to  $\sqrt{n}$  (see exercise 10.2). For instance, the Rabi oscillations between the states  $|g, 1\rangle$  and  $|e, 0\rangle$  take place at the Rabi frequency  $\Omega = |\lambda|$ , those between  $|g, 2\rangle$  and  $|e, 1\rangle$  at frequency  $\sqrt{2}\Omega$ , and so on.

**Exercise 10.2** Solve the Schrödinger equation for the Jaynes–Cummings model.<sup>3</sup> In particular, show that Rabi oscillations between the states  $|g, n\rangle$  and  $|e, n-1\rangle$  take place at frequency  $\sqrt{n}|\lambda|$ .

**Exercise 10.3** Discuss the temporal evolution of the atom–field entanglement for the Jaynes–Cummings model at resonance ( $\omega = \omega_0$ ), when the initial state is  $|\psi_0\rangle = |g, n\rangle$ .

**Exercise 10.4** Discuss the temporal evolution of the state of a two-level atom interacting with a single mode of the electromagnetic field according to the resonant Jaynes–Cummings model, when the initial state is  $|\psi_0\rangle = |g, \alpha\rangle$ , with  $|\alpha\rangle$  coherent state corresponding to a large average number of photons,  $\bar{n} = |\alpha|^2 \gg 1$ . Study the temporal evolution of the Bloch-sphere coordinates and of the von Neumann entropy of the atomic state. In particular, with the help of numerical simulations discuss the *collapse* of the Bloch vector to the center of the Bloch sphere and *revivals* at longer times.

#### 10.1.4 Entanglement generation

Atom–atom entanglement is obtained by sending two atoms, one after the other, through the cavity (see Hagley *et al.*, 1997). The two atoms and the cavity are initially prepared in the state

$$|\psi_i\rangle = |e_1, g_2, 0\rangle, \quad (10.12)$$

where the index 1 refers to the first atom, the index 2 to the second and the third quantum number gives the number of photons in the cavity, initially in the vacuum state  $|0\rangle$ . The interaction of the first atom with the cavity corresponds to a  $\theta = \pi/2$  Rabi pulse. Therefore, with probability  $\frac{1}{2}$  the atom emits a photon and evolves into the state  $|g\rangle$ , whereas with probability  $\frac{1}{2}$  it remains in  $|e\rangle$ . In the first case, the

---

<sup>3</sup>Note that the Jaynes–Cummings model is one of the few exactly solvable models in quantum field theory.

cavity is left in the state  $|1\rangle$ , in the latter it stays in  $|0\rangle$ . Therefore, the combined state of the two atoms and the cavity is given by

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|e_1, g_2, 0\rangle - |g_1, g_2, 1\rangle). \quad (10.13)$$

Note that the first atom is now maximally entangled with the cavity field, while there is no entanglement with the second atom. The second atom then enters the cavity and the angle  $\theta$  is set equal to  $\pi$ , corresponding to a complete population reversal ( $|g_2, 1\rangle \rightarrow |e_2, 0\rangle$ ). If instead the cavity is in the vacuum state, the second atom stays in  $|g_2\rangle$  without affecting the cavity field. In both cases the cavity ends up in the vacuum state and the overall state is now given by

$$|\psi_f\rangle = \frac{1}{\sqrt{2}}(|e_1, g_2\rangle - |g_1, e_2\rangle)|0\rangle. \quad (10.14)$$

Therefore, the two atoms are in a maximally entangled state, while the cavity state  $|0\rangle$  is factorized; that is, no atom–cavity entanglement remains.

A  $\pi/2$  pulse is then applied to both atoms in the analyzing cavity ( $R_2$  in Fig. 10.1), and finally the two detectors  $D_e$  and  $D_g$  measure the state of the two atoms. Since the EPR state  $\frac{1}{\sqrt{2}}(|e_1, g_2\rangle - |g_1, e_2\rangle)$  is rotationally invariant, it can be written in the same manner also in the basis rotated by the  $\pi/2$  pulse. Therefore, the joint probability  $p_{ge}$  of finding the first state in  $|g\rangle$  and the second in  $|e\rangle$  is  $\frac{1}{2}$ . Similarly,  $p_{eg} = 1/2$  and  $p_{gg} = p_{ee} = 0$ ; that is, perfect anticorrelation is expected (in the ideal case without experimental imperfections) in the outcomes of the measurements.

Finally, we point out that entanglement has been established between two atoms separated by a *macroscopic* distance (of the order of 1 cm in Hagley *et al.*, 1997).

**Exercise 10.5** In the entanglement generation protocol described above, the two atoms cross the analyzing cavity  $R_2$  at different times. If the frequency of  $R_2$  is detuned from the atomic resonance, the two atoms experience different phases,  $\phi_1$  and  $\phi_2$ , of the field. The phase difference  $\phi_2 - \phi_1$  accumulated between the microwave source and the atom is given by the  $\Delta T$ , where  $\Delta = \omega - \omega_0$  is the detuning; that is, the difference between the field frequency  $\omega$  and the Bohr frequency  $\omega_0$  associated with the  $|g\rangle \leftrightarrow |e\rangle$  transition, while  $T$ s is the interval separating the times at which the two atoms reach the cavity  $R_2$ . As a consequence the joint probabilities oscillate as a function of the phase difference  $\phi_1 - \phi_2$ . Show that

$$p_{eg} = p_{ge} = \frac{1}{4}[1 + \cos(\phi_2 - \phi_1)], \quad p_{gg} = p_{ee} = \frac{1}{4}[1 - \cos(\phi_2 - \phi_1)]. \quad (10.15)$$

## 10.2 The ion-trap quantum computer

The basic idea behind the use of trapped ions for quantum computation is to have a string of ions trapped in well controlled positions and to individually address each ion by means of laser pulses. The ion-trap quantum computer takes advantage of impressive experimental progress made in the field of quantum optics, which has

rendered *quantum state engineering* possible, *i.e.*, on-demand preparation and manipulation of quantum states with a very high degree of fidelity. Thanks to progress in laser technology, the degree of control over the states of trapped ions is continuously increasing, so that generation and coherent manipulation of entangled states with several qubits has been achieved. In this section, we shall first describe the main ingredients of ion-trap quantum computation, from the Paul trap mechanism to laser cooling. After this, we shall discuss the operations required to realize a universal set of one- and two-ion quantum gates.

### 10.2.1 The Paul trap

In the *Paul trap*, ions are confined by a spatially varying time-dependent radiofrequency (RF) field. We are interested in the case in which the trapped ions line up along the trap axis ( $z$ ). This is obtained by means of an oscillating field with a quadrupole geometry in two dimensions, providing confinement along the radial direction ( $r = \sqrt{x^2 + y^2}$ ), while trapping along the  $z$ -axis is provided by a static electric field (see Fig. 10.2 and exercise 10.6).

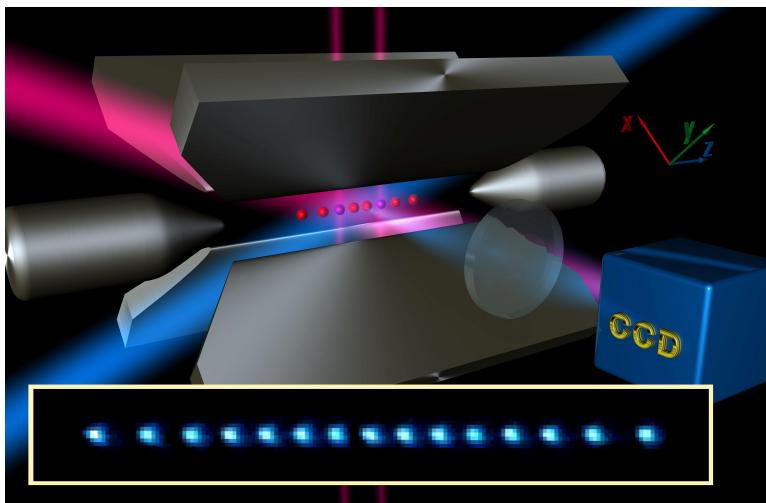


Fig. 10.2 Main figure: a schematic drawing of a linear ion trap setup with a trapped ion string. The four blades are at high voltage (neighbouring blades with opposite potential), oscillating at radio frequency, thus providing confinement in the radial directions. The tip electrodes are at positive high voltage and trap the ions axially (the  $z$  direction according to the notations used in the text). A laser addresses the ions individually and manipulates their quantum state. The resonance fluorescence of the ions is imaged onto a CCD (charge-coupled device) camera. Inset: the CCD image of a string of fifteen ions is shown. The distance between the outer ions is approximately  $70\text{ }\mu\text{m}$ . Drawing courtesy of Rainer Blatt, Innsbruck.

Let us first consider a single ion in a trap. By averaging over the fast oscillatory motion at (radio)frequency  $\omega_{RF}$ , an effective harmonic potential is obtained, with frequencies  $\omega_x, \omega_y, \omega_z$  along the three principal axes of the trap. Note that the

trapping frequency  $\omega_t \equiv \omega_z \ll \omega_x, \omega_y$ , so that we can limit our considerations to motion along the  $z$ -axis. Typical experimental parameters are trap size of approximately 1 mm, applied voltages of 100 – 500 V and RF field of a few tens of MHz leading to harmonic motion of the trapped ion in the  $z$  direction with frequency  $\frac{\omega_t}{2\pi} \sim 1 - 5$  MHz.

**Exercise 10.6** An ion with charge  $q$  and mass  $M$  is confined in a linear trap by the quadrupolar electric potential

$$\Phi(x, y, z; t) = \frac{1}{2} \frac{U_0}{R^2} (x^2 - y^2) \cos(\omega_{RF} t) + \frac{1}{2} \frac{V_0}{R^2} [z^2 - \epsilon x^2 - (1 - \epsilon)y^2], \quad (10.16)$$

with  $R$  and  $\epsilon$  geometric factors.

(i) Show that the equations of motion for the ion lead to harmonic confinement along  $z$ , while for both  $\xi = x$  and  $\xi = y$  we have a Mathieu differential equation, of the form

$$\frac{d^2\xi}{d\tau^2} + [a_\xi + 2q_\xi \cos(2\tau)]\xi = 0, \quad (10.17)$$

where  $\tau = \omega_{RF}t/2$ . Find numerically the region of stability of this equation for parameter values  $a_\xi$  and  $q_\xi$  around zero.

(ii) Compare the exact numerical solution of the Mathieu equation with the approximate analytic solution

$$\xi = \xi_0 \cos(\beta_\xi \tau) [1 + \frac{1}{2} q_\xi \cos(2\tau)], \quad (10.18)$$

where  $\beta_\xi = \sqrt{a_\xi + \frac{1}{2} q_\xi^2}$  and the initial conditions  $\xi(t = 0) = \xi_0(1 + q_\xi/2)$ ,  $\dot{\xi}(t = 0) = 0$  are assumed.

We are interested in both the vibrational motion of the ion in the trap and in the internal electronic motion. The electronic motion has frequencies  $O(10^{15})$  Hz and the motion relative to the hyperfine structure frequencies in the GHz range, while the motion of the ion in the trap is in the MHz range. Therefore, we can employ the Born–Oppenheimer approximation and separate the fast electronic motion from the slow motion of the ion. States relevant to quantum information processing can be written as  $|i\rangle|n\rangle$ , where  $|i\rangle$  refers to the electronic levels  $|g\rangle$  and  $|e\rangle$  (the computational basis states for a qubit) and  $n = 0, 1, 2, \dots$  denotes the harmonic oscillator states of the vibrational motion of the ion.

Let us now consider a string of  $N$  trapped ions (qubits). In this case, there are  $3N$  normal modes of vibration ( $2N$  radial and  $N$  axial modes). We are only interested here in the two lowest frequency axial modes, the centre-of-mass mode, where all ions oscillate together along  $z$  as a rigid body, and the stretch mode, where the oscillation amplitude of each ion is proportional to its distance from the centre of the trap. The frequencies of the centre-of-mass and stretch modes are  $\omega_c = \omega_t$  and  $\omega_s = \sqrt{3}\omega_c$ , where  $\omega_t$  is the frequency of the motion along  $z$  for a single ion. Note that the frequencies  $\omega_c$  and  $\omega_s$  are independent of the number  $N$  of ions in the trap (see exercise 10.7 for  $N = 2$  and  $N = 3$ ). The vibrational modes at

higher frequencies are essentially “frozen” during quantum information processing experiments and therefore we ignore them.

**Exercise 10.7** The Hamiltonian governing the motion of  $N$  ions in a harmonic linear trap is

$$H = \sum_{i=1}^N \frac{p_i^2}{2M} + \sum_{i=1}^N \frac{1}{2} M \omega_z^2 z_i^2 + \sum_{i=1}^{N-1} \sum_{j>i} \frac{q^2}{4\pi\epsilon_0 |z_j - z_i|}, \quad (10.19)$$

where  $q$  and  $M$  are the charge and mass of each ion,  $\epsilon_0$  is the electric permittivity of free space and the last term in (10.19) represents the Coulomb repulsion between the ions. Compute the equilibrium positions and the normal modes of vibration for  $N = 2$  and  $N = 3$ .

### 10.2.2 Laser pulses

Resonant interaction with laser light is used in all stages of ion-trap quantum computations, from state preparation by means of laser cooling techniques to controlled qubit manipulation to state measurement by the quantum-jump technique. The Hilbert space for  $N$  ions in a trap is spanned by the states  $|i_1, \dots, i_N; n\rangle$ , where  $i_1, \dots, i_N = g, e$  refer to the internal states of the ions, while  $n$  determines the collective vibrational motion of the ions. We assume that only one vibrational mode is relevant, say the centre-of-mass mode at frequency  $\omega_t$ . In  $|n\rangle$ , the string is in the  $n$ -th excited state for the (harmonic oscillator) motion at frequency  $\omega_t$  and we say that  $n$  *phonons* are excited. Let us consider a laser beam addressing the ion  $j$  ( $1 \leq j \leq N$ ), with the laser frequency  $\omega$  tuned in such a manner that  $\omega = \omega_0 + (n' - n)\omega_t$ , where  $\hbar\omega_0 = E_e - E_g$  is the energy difference between the ground state  $|g\rangle$  and the excited state  $|e\rangle$ . A resonant transition between the states  $|i_j = 0, n\rangle$  and  $|i_j = 1, n'\rangle$  is induced (we do not write the states of the other ions in the trap since they are not modified by the laser). As shown in Fig. 10.3, it is possible to combine two laser pulses in order to change only the vibrational state of the string and not the internal state of the ions. It is evident that, with an appropriate combination of laser pulses, we can build the generic motional superposition state  $\sum_n c_n |n\rangle$  starting from the ground state  $|0\rangle$  (see exercise 10.8). We can also build a generic superposition  $\alpha|g\rangle + \beta|e\rangle$  for the internal state of each ion. Indeed, a classical resonant field with  $\omega = \omega_0$  induces Rabi oscillations given by Eq. (10.1). It is then clear that a generic single-ion (-qubit) state can be obtained from the ground state  $|g\rangle$  by applying a resonant laser pulse of appropriate duration and phase.

The following three resonant interactions are of special importance for ion-trap quantum computation.

- (1) *Carrier resonance:* we have  $\omega = \omega_0$  and, keeping only the resonant terms, the Hamiltonian describing the trapped ion-laser interaction is given by

$$H_c = \frac{1}{2}\hbar\Omega(\sigma_+ e^{-i\phi} + \sigma_- e^{i\phi}), \quad (10.20)$$

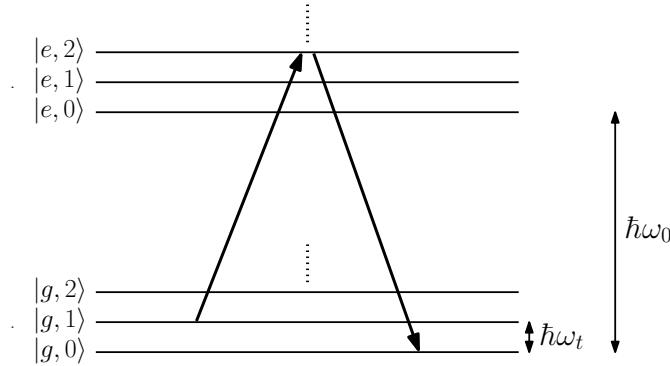


Fig. 10.3 Energy levels of a trapped ion. The global effect of the two transitions shown in the figure ( $|g, 1\rangle \rightarrow |e, 2\rangle$  and  $|e, 2\rangle \rightarrow |g, 0\rangle$ ) is to induce a transition between the quantized levels of the harmonic trapping potential, leaving unchanged the electronic state of the ion.

where  $\Omega$  is the Rabi frequency measuring the strength of the ion-laser coupling,  $\phi$  is the phase of the laser and the operators  $\sigma_+ = |e\rangle\langle g|$ ,  $\sigma_- = |g\rangle\langle e|$ . This Hamiltonian gives rise to transitions of the type  $|g, n\rangle \leftrightarrow |e, n\rangle$ . Indeed, the temporal evolution governed by Hamiltonian (10.20) in a time interval  $t$  leads to the unitary evolution operator

$$R_c(\theta, \phi) = e^{-\frac{i}{\hbar}H_c t} = \begin{bmatrix} \cos \frac{\theta}{2} & -ie^{i\phi} \sin \frac{\theta}{2} \\ -ie^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix}, \quad (10.21)$$

where  $\theta \equiv \Omega t$  and the matrix is written in the  $\{|e, n\rangle, |g, n\rangle\}$  basis. In particular, the transition  $|g, n\rangle \leftrightarrow |e, n\rangle$  is obtained when  $\theta = \pi$ , for any  $\phi$  (up to a phase factor determined by  $\phi$ ). More generally, Eq. (10.21) describes Rabi oscillations between the states  $|g, n\rangle$  and  $|e, n\rangle$ .

- (2) *First red sideband:* in this case  $\omega = \omega_0 - \omega_t$  (red detuned laser) and the trapped ion-laser resonant interaction Hamiltonian is

$$H_- = \frac{1}{2}\hbar\Omega\eta(a\sigma_+e^{-i\phi} + a^\dagger\sigma_-e^{i\phi}), \quad (10.22)$$

where  $a$  and  $a^\dagger$  are lowering and raising operators for the harmonic trapping potential and  $\eta = 2\pi z_0/\lambda$  is the *Lamb-Dicke parameter*, with  $z_0 = \langle 0|z^2|0\rangle^{1/2}$  spatial extension of the motional ground state and  $\lambda$  laser wavelength. Note that  $z_0 = \sqrt{\hbar/(2NM\omega_t)}$ , where  $M$  is the ion mass and  $N$  the number of ions in the string. The width of the ground state oscillations scales  $\propto 1/\sqrt{NM}$  since the effective mass of the collective centre-of-mass motion is  $NM$ . Hamiltonian (10.22) generates the unitary evolution

$$R_-(\theta, \phi) = e^{-\frac{i}{\hbar}H_- t} = \begin{bmatrix} \cos \frac{\theta}{2} & -ie^{i\phi} \sin \frac{\theta}{2} \\ -ie^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix}, \quad (10.23)$$

where  $\theta = \eta\Omega\sqrt{nt}$  and the matrix is written in the  $\{|g, n\rangle, |e, n-1\rangle\}$  basis. Therefore, Hamiltonian (10.22) gives rise to  $|g, n\rangle \leftrightarrow |e, n-1\rangle$  transitions with

Rabi frequency  $\eta\Omega\sqrt{n}$ . Note that (10.22) is formally equivalent to the resonant Jaynes–Cummings Hamiltonian. There is, however, a different physical interpretation: a phonon and not a photon is absorbed while the ion goes to the excited state. Moreover, the electromagnetic field is not quantized as in the Jaynes–Cummings model.

- (3) *First blue sideband:* we have  $\omega = \omega_0 + \omega_t$  (blue detuned laser) and resonant interaction Hamiltonian

$$H_+ = \frac{1}{2}\hbar\Omega\eta(a^\dagger\sigma_+e^{-i\phi} + a\sigma_-e^{i\phi}). \quad (10.24)$$

The unitary evolution  $R_+(\theta, \phi) = e^{-\frac{i}{\hbar}H_+t}$ , with  $\theta = \eta\Omega\sqrt{n+1}t$  has the same matrix representation as  $R_c$  and  $R_-$  but with respect to the  $\{|g, n\rangle, |e, n+1\rangle\}$  basis. Therefore, Hamiltonian (10.24) induces  $|g, n\rangle \leftrightarrow |e, n+1\rangle$  oscillations with frequency  $\eta\Omega\sqrt{n+1}$ . Such oscillations have no direct analogue in the cavity QED realm since a process in which the atom transits to an excited state while at the same time a photon is emitted would violate energy conservation. Hamiltonian (10.24) is known as the resonant *anti*-Jaynes–Cummings Hamiltonian.

**Exercise 10.8** Give a quantum protocol to build a generic motional superposition state  $\sum_{n=0}^N c_n|g, n\rangle$  starting from the ground state  $|g, 0\rangle$ .

The derivation of Hamiltonians (10.20), (10.22) and (10.24) can be found, for instance, in Leibfried *et al.* (2003a), see also exercise 10.9. Three important conditions must be fulfilled: (i) the Lamb–Dicke parameter  $\eta \ll 1$  (values of  $\eta \sim 0.2$  are typical in experiments); (ii) the laser must be on resonance to avoid undesired excitations of phonons; more precisely, we need  $|\omega - \omega_0| \ll \omega_t$  for the carrier transition,  $|\omega - (\omega_0 - \omega_t)| \ll \omega_t$  for the first red sideband transition and  $|\omega - (\omega_0 + \omega_t)| \ll \omega_t$  for the first blue sideband transition; (iii) the pulses must be longer than  $1/\omega_t$ , so that their Fourier spectrum does not extend over the sidebands.

**Exercise 10.9** The Hamiltonian describing the interaction of a trapped two-level ion with a laser field is

$$H_I = \hbar\Omega\left(\sigma_+e^{i(kz-\omega t+\phi)} + \sigma_-e^{-i(kz-\omega t+\phi)}\right), \quad (10.25)$$

where  $\Omega$  is the Rabi frequency,  $\phi$  the phase of the laser,  $\sigma_+ = |e\rangle\langle g|$ ,  $\sigma_- = \sigma_+^\dagger$ , and the motion is restricted to one dimension, the position of the ion in the harmonic trap being  $z = z_0(a^\dagger + a)$ , with  $z_0 = \sqrt{\hbar/(2M\omega_t)}$  ( $M$  is the mass of the ion and  $\omega_t$  the angular frequency of the trap, typically of the order of 10 MHz). The harmonic motion is quantized and described by the Hamiltonian  $H_{\text{osc}} = \hbar\omega_t(a^\dagger a + \frac{1}{2})$ . Study the effect of the interaction (10.25). In particular:

- (i) find the Rabi frequency for the resonant transitions  $|g, n\rangle \leftrightarrow |e, n'\rangle$ , with  $|n\rangle, |n'\rangle$  eigenstates of the Hamiltonian  $H_{\text{osc}}$ ;
- (ii) derive Hamiltonians (10.20), (10.22) and (10.24) from (10.25) in the limit in which the Lamb–Dicke parameter  $\eta = kz_0 \ll 1$ .

### Laser cooling

Laser cooling relies on the mechanical effect of light in a photon–ion scattering process, that is on the fact that photons carry not only energy, but also momentum  $p = h/\lambda$ , where  $h$  is the Planck constant and  $\lambda$  the wavelength of the light. If an ion is moving along the light beam, it sees a Doppler-shifted light frequency, the frequency being higher if the ion moves towards the laser beam and lower if the atom moves away from the beam. The physical principle of *Doppler cooling* is to compensate the Doppler shift for ions approaching the laser beam by means of a red detuned laser. Then these ions are slowed down owing to the photons kicking them. Typically, Doppler cooling allows cooling down to an average motional quantum state  $\langle n \rangle \sim 10$  for trap frequencies in the MHz range. The ultimate limit for Doppler cooling is due to the fact that the ions are excited to a strong (usually dipole) transition with natural linewidth (spontaneous emission rate)  $\Gamma > \omega_t$ .

The motional ground state  $|n = 0\rangle$  can then be prepared by *sideband cooling*; that is, by exciting the ions to a narrow transition ( $\Gamma \ll \omega_t$ ) (a forbidden optical line or a Raman transition). The laser is tuned into the  $|g, n\rangle$  to  $|e, n - 1\rangle$  first red sideband transition. Subsequent spontaneous emission occurs predominantly at the carrier frequency if the recoil energy of the atom is negligible compared with the vibrational quantum energy (this is the case if the Lamb–Dicke parameter  $\eta \ll 1$ ). In this case, spontaneous emission induces the transition  $|e, n - 1\rangle \rightarrow |g, n - 1\rangle$ . The red detuned laser then leads to  $|g, n - 1\rangle \rightarrow |e, n - 2\rangle$ , and so on. At the end of the process, the state  $|g, 0\rangle$  is reached with a high probability (preparation of the motional ground state for the centre-of-mass mode has been achieved with ground state occupation  $> 99.9\%$ ).

### Quantum gates

Single-qubit gates are obtained by tuning the laser to the carrier resonance. Indeed, it is clear from Eq. (10.21) that, starting from the ground state  $|g\rangle$ , a generic single-qubit state is obtained by means of a laser pulse of appropriate duration and phase.

The CNOT gate can be obtained following the proposal of Cirac and Zoller (1995). The basic idea is to employ the motional state of the string of ions as a “bus” to transfer quantum information between two qubits (ions). Therefore, the qubit–qubit interaction, which is necessary to implement controlled two-qubit operations, is mediated by the collective vibrational motion of the trapped ions.

Let us describe the Cirac–Zoller CNOT quantum gate between ions  $l$  (control qubit) and  $m$  (target qubit). We start from the initial state  $|i_1, \dots, i_l, \dots, i_m, \dots, i_N; n = 0\rangle$ , which we simply write as  $|i_l, i_m; 0\rangle$  since the other qubits are not affected by the quantum protocol described in what follows. The use of an auxiliary level ( $|a\rangle$ ) helps in performing the CNOT gate.<sup>4</sup> The following sequence of laser pulses is applied:

---

<sup>4</sup>The auxiliary level could be a third level (in addition to  $|g\rangle$  and  $|e\rangle$ ) in the hyperfine structure of the ground state.

- (1) A red detuned laser acts on ion  $l$ . The unitary evolution  $R_-(\theta = \pi, \phi = 0)$  changes the states  $\{|g_l, g_m\rangle, |g_l, e_m\rangle, |e_l, g_m\rangle, |e_l, e_m\rangle\}$  of the two-qubit computational basis as follows:

$$\begin{cases} |g_l, g_m; 0\rangle \rightarrow |g_l, g_m; 0\rangle, \\ |g_l, e_m; 0\rangle \rightarrow |g_l, e_m; 0\rangle, \\ |e_l, g_m; 0\rangle \rightarrow -i|g_l, g_m; 1\rangle, \\ |e_l, e_m; 0\rangle \rightarrow -i|g_l, e_m; 1\rangle. \end{cases} \quad (10.26)$$

As a result of this laser pulse, the quantum information of the control ion is mapped onto the vibrational mode.

- (2) A red detuned laser is applied to ion  $m$ . The corresponding unitary evolution, written in the  $\{|g_m, n=1\rangle, |a_m, n=0\rangle\}$  basis, is  $R_-(\theta = 2\pi, \phi = 0)$ , note that the auxiliary level  $|a_m\rangle$  is involved. Since  $\theta = 2\pi$ , the state  $|g_m, 1\rangle$  is mapped into  $-|g_m, 1\rangle$ . Therefore, the states obtained at the end of (10.26) are modified as follows:

$$\begin{cases} |g_l, g_m; 0\rangle \rightarrow |g_l, g_m; 0\rangle, \\ |g_l, e_m; 0\rangle \rightarrow |g_l, e_m; 0\rangle, \\ -i|e_l, g_m; 1\rangle \rightarrow i|e_l, g_m; 1\rangle, \\ -i|e_l, e_m; 1\rangle \rightarrow -i|e_l, e_m; 1\rangle. \end{cases} \quad (10.27)$$

- (3) A red detuned laser is applied to ion  $l$ , inducing again the unitary evolution  $R_-(\theta = \pi, \phi = 0)$ . This leads to

$$\begin{cases} |g_l, g_m; 0\rangle \rightarrow |g_l, g_m; 0\rangle, \\ |g_l, e_m; 0\rangle \rightarrow |g_l, e_m; 0\rangle, \\ i|e_l, g_m; 1\rangle \rightarrow |g_l, g_m; 0\rangle, \\ -i|e_l, e_m; 1\rangle \rightarrow -|g_l, e_m; 0\rangle. \end{cases} \quad (10.28)$$

The effect of this pulse is to map the state of the vibrational mode back onto the control qubit.

The global effect of the three laser pulses is to induce a controlled phase-shift gate CMINUS = CPHASE( $\pi$ ). The CNOT gate is then obtained from CMINUS after application of single-qubit (Hadamard) gates (see exercise 3.13).

It should be remarked that, although the vibrational mode could be regarded as an additional qubit (spanned by the phonon states  $|0\rangle$  and  $|1\rangle$ ), in practice it is only used as a bus to transfer quantum information between ions. Indeed, the “vibrational qubit” cannot be measured independently, as is the case for the internal electronic states of the ions.

### Quantum-jump detection

After a quantum computation, the state of each ion can be measured using *quantum-jump* detection: each ion is illuminated with laser light of polarization and frequency such that it absorbs and then re-emits photons only if it is in one particular qubit level (say, the state  $|e\rangle$ ). In contrast, if it is in the other ( $|g\rangle$ ) state, the laser frequency is out of resonance and does not induce any transition. Thus, the detection

of scattered fluorescence photons indicates that the ion was in the state  $|e\rangle$ . This is a projective measurement and a state discrimination efficiency above 99% can be reached. Moreover, it is possible to measure several ions in a trap individually. In order to uncover the average populations of the states  $|g\rangle$  and  $|e\rangle$  for each ion, one has to repeat the quantum computation and the final measurement a sufficient number of times.

### Optical and hyperfine qubits

In closing this section, we briefly discuss the choice of the states  $|g\rangle$  and  $|e\rangle$  in the experiments. One needs two “stable” levels; that is, two levels whose decay rates are much smaller than the Rabi frequencies associated with the laser-induced transition  $|g\rangle \leftrightarrow |e\rangle$ . Two different strategies have been followed:  $|g\rangle$  and  $|e\rangle$  are either the ground state and a metastable excited state connected by a forbidden optical transition (*optical qubit*, see e.g. Schmidt-Kaler *et al.*, 2003 for  $^{40}\text{Ca}^+$ ) or two hyperfine sub-levels of the ground state (*hyperfine qubit*, see e.g. Turchette *et al.*, 1998 for  $^9\text{Be}^+$ ). An optical transition is driven by a single laser, while for a hyperfine transition two lasers are used, far detuned from an intermediate level  $|c\rangle$ . Such a Raman configuration (see Fig. 10.4 and exercise 10.10) is used because the frequency  $\omega_0$  for a hyperfine transition is  $O(\text{GHz})$  and can be driven resonantly by a single electromagnetic wave with wavelength of order  $10^{-1}\text{ m}$ . It is clear that such a wavelength, much larger than the distance between two nearby ions in a trap (approximately  $10\text{ }\mu\text{m}$ ), would not allow single-ion addressing.

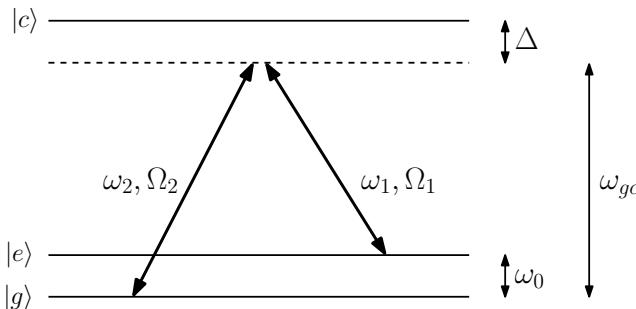


Fig. 10.4 A schematic diagram of a Raman transition. The two lasers have frequencies  $\omega_1$  and  $\omega_2$ , while  $\Omega_1$  and  $\Omega_2$  are the Rabi frequencies for the transitions  $|e\rangle \leftrightarrow |c\rangle$  and  $|g\rangle \leftrightarrow |c\rangle$ , respectively.

**Exercise 10.10** Show that in the Raman configuration drawn in Fig. 10.4, the transition  $|g\rangle \leftrightarrow |e\rangle$  takes place with Rabi frequency

$$\Omega_R \approx \frac{2\Omega_1\Omega_2}{\Delta}, \quad (10.29)$$

where  $\Omega_1$  and  $\Omega_2$  are the Rabi frequencies of the two applied laser fields and  $\Delta \gg \Omega_1, \Omega_2$  is the detuning with respect to the transitions  $|g\rangle \leftrightarrow |c\rangle$  and  $|e\rangle \leftrightarrow |c\rangle$  (note that we have set  $\hbar = 1$ ). We assume that the initial wave function  $|\psi(0)\rangle = |g\rangle$ .

Besides Eq. (10.29), the Raman approximation also predicts that level  $|c\rangle$  remains essentially unpopulated. Check the validity of the Raman approximation by direct numerical integration of the equations of motion for the overall three-level ( $|g\rangle$ ,  $|e\rangle$  and  $|c\rangle$ ) system.

A first implementation of the Cirac–Zoller CNOT quantum gate was reported in Schmidt-Kaler *et al.* (2003), while a different implementation of a two-ion gate was realized in Leibfried *et al.* (2003b). For the preparation and characterization of many-ion entangled states, see Leibfried *et al.* (2005), Häffner *et al.* (2005), and Monz *et al.* (2011). Many quantum protocols have been implemented using trapped ions including deterministic teleportation (Riebe *et al.*, 2004 and Barrett *et al.*, 2004), quantum error correction (Chiaverini *et al.*, 2004), and Grover’s, Deutsch–Jozsa, and Shor’s quantum algorithms with 2–7 qubits, see *e.g.* Monz *et al.* (2016).

### 10.3 Solid-state qubits

Qubits made out of solid-state devices may offer great advantages since fabrication by established lithographic methods allows for *scalability* (at least in principle). Moreover, another important feature of solid-state devices is their *flexibility* in design and manipulation schemes. Indeed, in contrast to “natural” atoms, “artificial” solid-state atoms can be lithographically designed to have specific characteristics such as a particular transition frequency. This tunability is an important advantage over natural atoms. Finally, solid-state qubits are easily embedded in electronic circuits and can take advantage of the rapid technological progress in solid-state devices as well as of continuous progress in the field of nanostructures. On the other hand, it should be remarked that there is a great variety of decoherence mechanisms, still not well understood, in solid-state devices.

Two main strategies have been followed for making solid-state qubits. In the first strategy, the qubits are single particles, such as nuclear spins in semiconductors or single electron spins in semiconductor quantum dots. In the second strategy, qubits are constructed from superconducting nanocircuits based on the Josephson effect.

#### 10.3.1 Spins in semiconductors

A proposal by Kane (1998) is sketched in Fig. 10.5. The qubits are the  $S = \frac{1}{2}$  nuclear spins of  $^{31}\text{P}$  impurities in silicon. Gate operations are performed by means of magnetic fields and static electric fields. Each qubit is controlled through the hyperfine interaction between the nucleus of  $^{31}\text{P}$  and the bound electron around it. Such an interaction is due to the coupling between the nuclear spin  $\mathbf{S}_n$  and the electronic spin  $\mathbf{S}_e$  and its strength is proportional to  $|\psi(\mathbf{0})|^2$ ; that is, to the probability density of the electron wave function at the nucleus position. The hyperfine coupling can be controlled by an applied electric field ( $A$ -gates in Fig. 10.5) that shifts the electron wave function from the phosphorus nucleus, thus reducing the hyperfine interaction. The transition frequency of each qubit ( $^{31}\text{P}$  nucleus)

is therefore determined by both the static magnetic field  $B$  applied to it and the hyperfine interaction. Thus, the  $A$ -gates can control the transition frequency of each single qubit and bring them into resonance with the oscillating magnetic field  $B_{AC}$ . In this manner, arbitrary single-qubit quantum gates can be realized with resonant pulses. Two-qubit quantum gates would be implemented using the  $J$ -gates of Fig. 10.5, which control the exchange interaction between two neighbouring bound electrons. Indeed, the exchange interaction depends on the overlap of the electron wave functions and can be controlled by the  $J$ -gates bringing the two electrons closer. Since the hyperfine interaction couple each qubit with its bound electron, the qubit–qubit interaction is mediated by the exchange interaction between the electrons. This proposal requires nanofabrication on the atomic scale, to place phosphorus impurities (and gates) in a silicon crystal in an ordered array with separation of order 10 nm. Kane’s proposal remains extremely challenging as a path to practical quantum computing. Nevertheless, one should consider the fact that silicon technology is a very rapidly developing field.

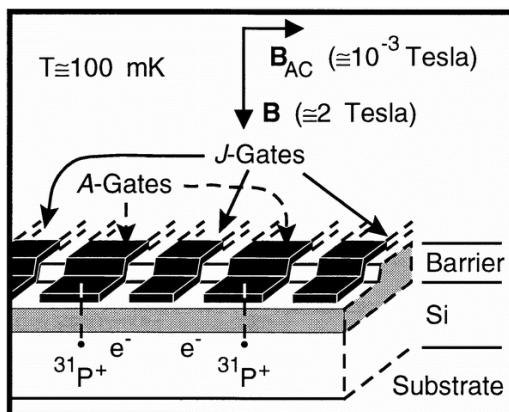


Fig. 10.5 A schematic drawing of Kane’s proposal. The figure is reprinted with permission from Kane (1998). ©(1998) Macmillan Publishers Ltd.

### 10.3.2 Quantum dots

Quantum dots are structures fabricated from semiconductor materials, in which electrostatic potentials confine electrons inside small “boxes”. When the size of the box is comparable to the wavelength of the electrons by which it is occupied, then the system exhibits a sequence of discrete energy levels, quite as in atoms. For this reason quantum dots are also known as artificial atoms. Typical binding energies and size of the orbits are 1 meV and 50 nm, to be compared with 10 eV and 0.05 nm (Bohr radius) for natural atoms. Quantum dots are fabricated starting with a semiconductor heterostructure, a sandwich of different layers of semiconducting materials (such as GaAs and AlGaAs), which are grown on top of each other

using molecular-beam epitaxy. By doping the AlGaAs layer with Si, free electrons are introduced, which accumulate at the interface between GaAs and AlGaAs, thus forming a two-dimensional gas of electrons that move along the interface. Metal gate electrodes applied on top of the heterostructure create an electric field that locally depletes the two-dimensional electron gas, creating one or more small islands (quantum dots) of confined electrons in an otherwise depleted region (see Fig. 10.6). Note that two sufficiently close quantum dots can be coupled through the overlapping of their electron wave functions, thus creating artificial molecules. At present, single and double quantum dots can be made, with a number of electrons controllable (by means of an applied voltage) down to just one electron.

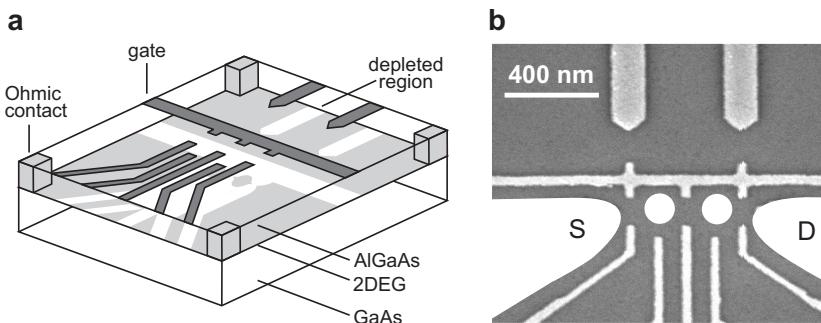


Fig. 10.6 A schematic drawing (a) and scanning electron micrograph (b) of a semiconductor heterostructure with two coupled quantum dots. In the left-hand figure, negative voltages applied to metal gate electrodes (dark gray) lead to depleted regions (white) in the two-dimensional electron gas (light gray). Electric contacts to reservoirs are obtained through ohmic contacts. In the right-hand figures, the gate electrodes (light gray) are shown on top of the surface of the heterostructure (dark gray). The source (S) and drain (D) reservoirs are connected to the two quantum dots (white circles) via tunnel barriers. The two upper electrodes can be used to measure changes in the number of electrons in the dots. The figure is taken from Elzerman *et al.* (2006), with kind permission of Società Italiana di Fisica, ©(2006) by the Italian Physical Society.

The qubit can be realized as the electronic spin of a single-electron quantum dot, the use of electron spins as qubits being attractive due to their long decoherence time. In the proposal of Loss and DiVincenzo (1998) (see Fig. 10.7) the dots that hold the electron spins (qubits) are placed in an array on top of a semiconductor heterostructure. A static magnetic field  $B$  induces an energy gap (Zeeman splitting) between the states  $|0\rangle$  (spin up) and  $|1\rangle$  (spin down) of each qubit. The Zeeman splitting is  $\Delta E = g\mu_B B$ , where  $g \approx -0.44$  is the Landé  $g$ -factor of GaAs and  $\mu_B \approx 9.27 \times 10^{-24}$  Joule/Tesla is the Bohr magneton. The spin state of single qubits can then be controlled by applying an oscillating magnetic field  $B_{ac}$  in resonance with the Zeeman splitting (that is, with angular frequency  $\Delta E/\hbar$ ). This technique is known as electron-spin resonance. A local difference in the Zeeman splittings could be obtained by means of gate potentials applied between the top and the bottom of the heterostructure. Each electron could then be shifted individually towards a layer of the heterostructure with a different  $g$ -factor. This would allow resonant addressing of individual qubits.

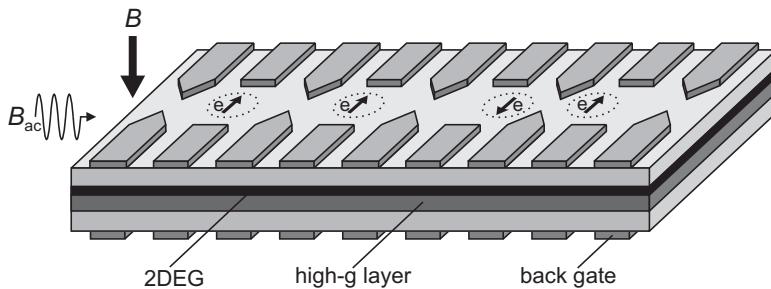


Fig. 10.7 A schematic picture of an array of quantum-dot spin qubits as proposed by Loss and DiVincenzo (1998). The quantum dots (circles) are created by metal electrodes on top of a semiconductor heterostructure containing a two-dimensional electron gas (2DEG). Each dot holds a single electron, whose spin state is pictorially represented by an arrow. The back gates can modify the Zeeman splitting by pulling the electron wave function into a layer with a large  $g$ -factor. The figure is taken from Elzerman *et al.* (2006), with kind permission of Società Italiana di Fisica, ©(2006) by the Italian Physical Society.

The interaction between two spins,  $\mathbf{S}_i$  and  $\mathbf{S}_j$ , can be modelled by the Heisenberg exchange interaction  $J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j$ , where  $J_{ij}$  depends on the overlap of the electronic wave functions. The coupling  $J_{ij}$  is relevant only between nearest neighbour qubits, provided each electron is well localized in a single quantum dot. Applying a gate voltage at the surface, the potential barrier between adjacent dots can be increased, thus reducing drastically the Heisenberg coupling. It is therefore possible to switch on and off the coupling between qubits and this provides a clear mechanism for the implementation of two-qubit gates. In particular, it can be checked that, if the interaction between two neighbouring qubits is switched on for a specific duration  $t_s$ , then the SWAP gate is realized. If on the other hand the duration is  $t_s/2$ , then by definition the  $\sqrt{\text{SWAP}}$  gate is implemented. The important point is that the  $\sqrt{\text{SWAP}}$  and single-qubit gates constitute a universal set of quantum gates (indeed, as shown in exercise 10.13, the CMINUS gate can be obtained from  $\sqrt{\text{SWAP}}$  and single-qubit gates).<sup>5</sup>

Readout is possible if the information contained in the spin is converted to information contained in the charge by a spin-dependent tunnelling process. First, the gate voltage is modified so that the electron stays in the dot if it has spin up, while it leaves the dot (tunnelling to a reservoir) if it has spin down. Detection of the charge of the dot is then possible using devices such as quantum point contacts (see Elzerman *et al.*, 2006). In this manner, the difficult problem of measuring the polarization of a single spin has been replaced by a much easier charge measurement.

Coherent control of two coupled electron spins in a double quantum dot was demonstrated by Petta *et al.* (2005); high-fidelity qubit initialization and read-out were demonstrated, see *e.g.* Kloeffel and Loss (2013) for an overview.

<sup>5</sup>Note that, by properly encoding each logical qubit into three spins instead of one, it is possible to perform universal quantum computation using only the Heisenberg exchange interaction (DiVincenzo *et al.*, 2000b). This possibility may be useful as it avoids the implementation of single-spin rotations, which is difficult in quantum-dot arrays.

Scalability is in principle possible since arrays of quantum dots can be produced with present technology. However, it should be taken into account that there are a great variety of possible decoherence processes in quantum dots and our knowledge of them is still very limited.

**Exercise 10.11** The simplest example to study the bound states of a particle is the infinitely deep one-dimensional square-well potential

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ +\infty, & x \leq 0, \quad a \leq x. \end{cases} \quad (10.30)$$

Find the stationary states and the energy levels for this model.

**Exercise 10.12** A more realistic case useful for the study of bound states is the well of finite depth:

$$V(x) = \begin{cases} -V_0, & -a < x < a, \\ 0, & x \leq -a, \quad a \leq x, \end{cases} \quad (10.31)$$

with  $V_0 > 0$ . Find the bound stationary states and energy levels for this model.

**Exercise 10.13** Show that the CMINUS gate can be obtained from the  $\sqrt{\text{SWAP}}$  and single-qubit gates as follows:

$$\begin{aligned} \text{CMINUS} = & (I \otimes R_z(\pi)) (\sqrt{\text{SWAP}})^{-1} (I \otimes R_z(\pi/2)) \\ & \times \text{SWAP} (I \otimes R_z(-\pi/2)) \sqrt{\text{SWAP}}. \end{aligned} \quad (10.32)$$

### 10.3.3 Superconducting qubit circuits

Superconductors have the ability to conduct electricity without loss of energy. In superconductors, pairs of electrons are bound together to form objects of twice the electron charge, known as Cooper pairs. A Josephson junction consists of two superconductors separated by a thin insulating barrier (see, e.g., Tinkham, 1996). Cooper pairs can tunnel through the barrier, this being a dissipationless process. Note that quantum tunnelling allows transport through regions that are classically forbidden owing to potential barriers (see exercise 10.14).

**Exercise 10.14** Study the transmission properties of a square barrier, described by the potential

$$V(x) = \begin{cases} V_0, & 0 < x < a, \\ 0, & x \leq 0, \quad a \leq x, \end{cases} \quad (10.33)$$

with  $V_0 > 0$ . Consider the case where the energy  $E < V_0$  (the *tunnel effect*).

Two energy scales determine the behaviour of a Josephson-junction circuit: the *Josephson energy*  $E_J$  and the electrostatic *charging energy*  $E_C$  for a single Cooper pair. The Josephson energy is related to the critical current  $I_J$  (the maximum

current that can flow through the junction without dissipation) by the relation  $E_J = I_J \hbar / 2e$ . Depending on the ratio  $E_J/E_C$ , one can distinguish between charge qubits ( $E_J \ll E_C$ , typically  $E_J/E_C \sim 0.1$ ), charge-flux qubits ( $E_J/E_C \sim 1$ ), flux qubits ( $E_J/E_C \sim 10$ ) and phase qubits ( $E_J \gg E_C$ , typically  $E_J/E_C \sim 10^6$ ). In what follows, we shall limit ourselves to the discussion of charge qubits, trying to give a flavour of the flexibility in the design and manipulation of superconducting qubits.

### Charge qubits

Electrostatic potentials can confine Cooper pairs in a “box” of micron size. In a Josephson junction a *Cooper-pair box*, known as the island, is connected by a thin insulator (tunnel junction) to a superconducting reservoir (see Fig. 10.8). Cooper pairs can move from the island to the reservoir and *vice versa* by quantum tunnelling effects. They enter the island one-by-one when a control-gate electrode (voltage  $U$ ), capacitively coupled to the island (capacitance  $C_g$ ), is varied. The island has discrete quantum states and, as we shall see below, under appropriate experimental conditions the two lowest energy states form a two-level system appropriate for a qubit. The charging energy is  $E_C = (2e)^2 / 2C$ , where  $C = C_J + C_g$  is the total capacitance of the island,  $C_J$  being the tunnel junction capacitance. If the capacitance  $C$  is in the range of a femtofarad or smaller, then  $E_C/k_B \geq 1\text{ K}$ . Typical values of  $E_J/k_B$  in the circuits considered here are instead 0.1 K, so that  $E_J/E_C \ll 1$ .

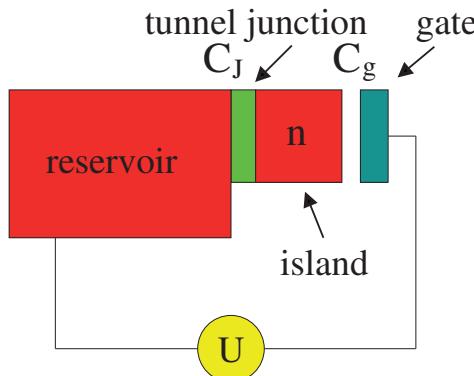


Fig. 10.8 A schematic drawing of a Josephson-junction qubit in its simplest design: a small superconducting island with  $n$  excess Cooper pairs (relative to some reference state) is connected by a tunnel junction with capacitance  $C_J$  and Josephson coupling energy  $E_J$  to a superconducting reservoir. The junction is biased by a gate voltage  $U$  with gate capacitance  $C_g$ .

The Cooper-pair box is described by the Hamiltonian

$$H = E_C(n - n_g)^2 - E_J \cos \phi, \quad (10.34)$$

where  $n$  is the number of extra Cooper pairs in the island and  $\phi$  the phase drop of the superconducting order parameter across the junction. The variables  $\phi$  and  $n$

are conjugate; that is,  $[\phi, n] = i$ . The dimensionless gate charge  $n_g = C_g U / 2e$  can be controlled by tuning the gate voltage  $U$ . The Josephson coupling  $E_J \cos \phi$  can be interpreted as a nonlinear potential energy term. Nonlinearity is crucial since it leads to energy levels that are not equally spaced. This permits the isolation of two energy levels to provide the two computational basis states for a qubit. In the regime  $E_J/E_C \ll 1$  a convenient basis is the basis of the eigenstates  $|n\rangle$  of the number operator  $n$ . The Josephson term is not diagonal in this basis since

$$\cos \phi |n\rangle = \frac{1}{2} (e^{i\phi} + e^{-i\phi}) |n\rangle = \frac{1}{2} (|n+1\rangle + |n-1\rangle). \quad (10.35)$$

Therefore, in the  $n$ -basis Hamiltonian (10.34) reads

$$H = E_C \sum_n (n - n_g)^2 |n\rangle \langle n| - \frac{1}{2} E_J \sum_n (|n+1\rangle \langle n| + |n\rangle \langle n+1|). \quad (10.36)$$

For  $E_J \ll E_C$  the charge states  $|n\rangle$  are weakly mixed by the Josephson term, except near the “optimal” operating points with  $n_g$  half-integer, where the electrostatic charging energy of the states  $|n_g - \frac{1}{2}\rangle$  and  $|n_g + \frac{1}{2}\rangle$  is the same and the Josephson coupling mixes them strongly (see Fig. 10.9). Therefore, the dynamics at low temperatures ( $k_B T \ll E_C$ ) is essentially limited to these two charge states. To simplify writing, we assume  $n_g$  around  $\frac{1}{2}$ , so that the two relevant charge states are  $|0\rangle$  and  $|1\rangle$ . Projection of the Hamiltonian (10.36) onto the subspace spanned by these two states leads (neglecting an irrelevant energy offset) to the single-qubit Hamiltonian

$$H_Q = \frac{1}{2} \epsilon \sigma_z - \frac{1}{2} \Delta \sigma_x, \quad (10.37)$$

where  $\epsilon = E_C(2n_g - 1)$  and  $\Delta = E_J$ . This Hamiltonian can be easily diagonalized. The energy splitting between its eigenvalues is  $\Omega = \sqrt{\epsilon^2 + \Delta^2}$ . The eigenvalues are  $\lambda_{\pm} = \pm \frac{\Omega}{2}$  and the corresponding eigenstates read

$$\begin{aligned} |+\rangle &= \cos \frac{\theta}{2} |0\rangle - \sin \frac{\theta}{2} |1\rangle, \\ |-\rangle &= \sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle, \end{aligned} \quad (10.38)$$

where we have introduced the mixing angle  $\theta$ , defined by  $\tan \theta = \Delta/\epsilon$ . At the degeneracy point,  $\theta = \frac{\pi}{2}$ , the eigenstates are equal superpositions of the states  $|0\rangle$  and  $|1\rangle$  and the energy splitting  $\Omega = \Delta = E_J$ . Far from the degeneracy point the eigenstates  $|\pm\rangle$  reduce to  $|0\rangle$  and  $|1\rangle$ , as the charging energy is the dominant term in the Hamiltonian  $H_Q$ . Typical frequencies are  $\Omega/2\pi \sim 10$  GHz.

It is now clear that generic single-qubit operations can be implemented by properly switching the gate voltage (thus tuning the Hamiltonian  $H_Q$ ) for a given time (see exercise 10.15). For instance, one can start far from the degeneracy point, move the system quickly to the degeneracy point (by means of a change in the gate voltage) for a time  $T$  and then back to the initial value of the gate voltage. This pulse implements a rotation about the  $x$ -axis of the Bloch sphere and was realized by Nakamura *et al.* (1999). To grasp the dramatic progress made by the field of quantum information processing with superconducting circuits, it is enough to note

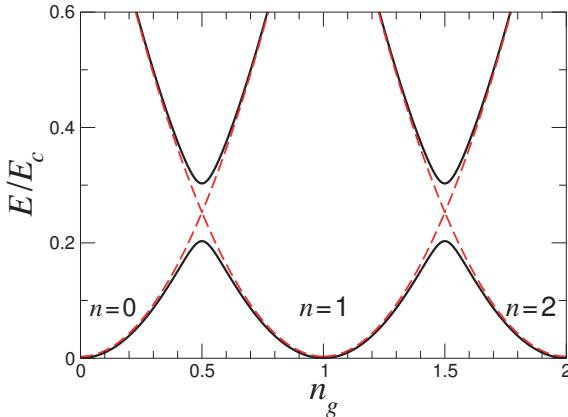


Fig. 10.9 The lowest energy levels of a Cooper-pair box, for  $E_J/E_C = 0.1$  (solid curves). The dashed curves show the energy levels for  $E_J = 0$ .

that the single-qubit coherence time has improved about five orders of magnitude from 1 ns in Nakamura *et al.* (1999) to 100  $\mu$ s, as reported for instance in Gambetta *et al.* (2017). Such progress has made it possible, among other things, the demonstration of two-qubit Grover and Deutsch quantum algorithms (DiCarlo *et al.*, 2009) and the implementation of quantum error-correction schemes (Kelly *et al.*, 2015).

**Exercise 10.15** For a two-level system, study the effect of a pulse of duration  $T$ , described by the Hamiltonian (10.37).

#### Circuit quantum electrodynamics

Circuit quantum electrodynamics (circuit QED) uses superconducting qubits as artificial atoms coupled to one or more transmission line resonators (see the schematic representation of Fig. 10.10). The qubits interact with the electromagnetic (microwave) field in the transmission lines (the cavity). Transmission lines with very high cavity quality factors  $Q$  have been demonstrated ( $Q \sim 10^6$  in Schoelkopf and Girvin, 2008, corresponding to photons making up to  $10^6$  bounces before being lost). This means that a photon of GHz frequency travels back and forth a total distance of 10 km before being dissipated. This allows multiple-qubit entanglement using transmission lines spanning distances of several millimetres.

Moreover, circuit QED allows to address the *ultrastrong-coupling regime*, where the matter-field coupling strength  $\lambda$  becomes comparable or even exceeds the resonator frequency  $\omega$  (Forn-Díaz *et al.*, 2017; Yoshihara *et al.*, 2017). Note that in cavity QED experiments the ratio  $\lambda/\omega \sim 10^{-6}$ . The coupling strength  $\lambda \propto 1/\sqrt{V}$ , with  $V$  the quantization volume (*i.e.*, the volume of the cavity). In contrast with cavity QED where  $V$  is usually larger than  $l^3$ , with  $l$  wave length of the cavity mode, in circuit QED by means of microstrips one can achieve quantization volumes much smaller than  $l^3$ . As a consequence,  $\lambda/\omega \sim 1$  can be achieved. This

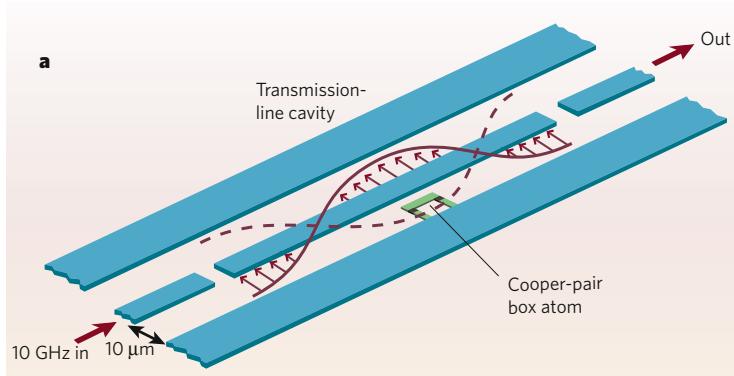


Fig. 10.10 Schematic drawing of circuit QED, with a superconducting qubit (Cooper-pair box atom) interacting with the electromagnetic field in a transmission line, consisting of a central conductor and two ground planes on either side. The cavity and qubit are measured by sending microwave signals on one side of the cavity and collecting the transmitted signals on the output side. The figure is reprinted with permission from Schoelkopf and Girvin (2008). ©(2008) Nature Publishing Group.

regime is interesting for high-speed manipulation of quantum systems, a problem of great relevance in quantum information processing: quantum gates should operate on a time scale much smaller than the decoherence time scale, to allow efficient error correction and fault-tolerant architectures. Furthermore, the ultrastrong coupling regime has interesting properties on its own, such as the emergence of a strongly correlated light-matter ground state (Forn-Díaz *et al.*, 2017; Yoshihara *et al.*, 2017). The RWA approximation is not justified and one cannot use the Jaynes-Cummings Hamiltonian (10.7) to describe the dynamics of the system, but rather the full Rabi Hamiltonian (7.111) (including the counter-rotating terms  $a^\dagger\sigma_+$  and  $a\sigma_-$ ) must be used.

A related problem is the detection of the dynamical Casimir effect (DCE), namely the generation of photons from the vacuum due to time-dependent boundary conditions or more generally to the nonadiabatic change of some parameters of a system. Indeed, a rapid variation of the matter-field coupling is needed to implement ultrafast quantum gates, and therefore the DCE appears as a fundamental limit to the implementation of high-speed quantum gates (see Benenti *et al.*, 2014b) and more generally to the development of ultrafast quantum technologies.

**Exercise 10.16** Consider the Rabi Hamiltonian with a time-dependent coupling:

$$\begin{aligned} H_{\text{Rabi}}(t) &= H_0 + H_I(t), \\ H_0 &= -\frac{1}{2}\hbar\omega_0\sigma_z + \hbar\omega(a^\dagger a + \frac{1}{2}), \\ H_I(t) &= \hbar f(t)(\lambda\sigma_+ + \lambda^*\sigma_-)(a^\dagger + a), \end{aligned} \quad (10.39)$$

and assume sudden switch on/off of the coupling:  $f(t) = 1$  for  $0 \leq t \leq \tau$ ,  $f(t) = 0$  otherwise. For simplicity's sake, consider the resonant case ( $\omega_0 = \omega$ ) and the coupling strength  $\lambda \in \mathbb{R}$ . Assume that initially ( $t = 0$ ) both the field and the qubit

are prepared in their ground state,  $|\psi(0)\rangle = |g, 0\rangle$ , so that within RWA there is no generation of photons,  $\langle n \rangle = 0$  at all times. On the other hand, in the *nonadiabatic regime* in which the coupling is switched on and off suddenly, the DCE leads to the generation of photons.

Compute numerically the mean number of photons,  $\langle n \rangle$ , as function of time, and compare the obtained results with those of time-dependent perturbation theory, derived as follows (see Benenti *et al.*, 2014a).

We first expand in the interaction picture the qubit-field state at time  $t$  as  $|\psi(t)\rangle = \sum_{l=g,e} \sum_{n=0}^{\infty} C_{l,n}(t) |l, n\rangle$ . The time-evolution of the coefficients  $C_{l,n}$  is governed by the equations

$$\begin{cases} i \dot{C}_{g,n}(t) = \Omega_n C_{e,n-1}(t) + \Omega_{n+1} e^{-2i\omega t} C_{e,n+1}(t), \\ i \dot{C}_{e,n}(t) = \Omega_{n+1} C_{g,n+1}(t) + \Omega_n e^{2i\omega t} C_{g,n-1}(t), \end{cases} \quad (10.40)$$

with the Rabi frequencies  $\Omega_n = \lambda \sqrt{n}$  ( $n = 0, 1, 2, \dots$ , where it is understood in the above equations that the terms  $C_{l,n}$  and  $\dot{C}_{l,n}$  must be set to zero when  $n < 0$ ).

The solution to the time-dependent Schrödinger equation (in the interaction picture)  $i\hbar\dot{\psi}(t) = H_I(t)\psi(t)$  can be approximated by the *Picard iterative process* (or *Picard series*). We start by writing the integral associated equation

$$\psi(t) = \psi(0) - \frac{i}{\hbar} \int_0^t H_I(t') \psi(t') dt'. \quad (10.41)$$

Iterating the process we obtain

$$\psi(t) = \psi(0) - \frac{i}{\hbar} \int_0^t H_I(t') \left[ \psi(0) - \frac{i}{\hbar} \int_0^{t'} H_I(t'') \psi(t'') dt'' \right] dt', \quad (10.42)$$

and so on. Hence we can write  $\psi(t) = \sum_{k=0}^{\infty} \psi^{(n)}(t)$ , with the zeroth-order approximation  $\psi^{(0)}(t) = \psi(0)$ , the first-order correction

$$\psi^{(1)}(t) = -\frac{i}{\hbar} \int_0^t H_I(t') \psi(0) dt', \quad (10.43)$$

the second-order correction

$$\psi^{(2)}(t) = -\frac{i}{\hbar} \int_0^t H_I(t') \psi^{(1)}(t') dt' = -\frac{1}{\hbar^2} \int_0^t H_I(t') \int_0^{t'} H_I(t'') \psi(0) dt' dt'', \quad (10.44)$$

and so on.

Intuitive diagrams can be used to represent the various terms in the Picard series. We consider as initial state  $|\psi(0)\rangle = |g, 0\rangle$ , so that the zeroth-order approximation  $|\psi^{(0)}(t)\rangle = |g, 0\rangle$ . Such state is diagrammatically represented as a vertical single line (see Fig. 10.11 (top left)), meaning that the qubit is in its ground state  $|g\rangle$ , while no photons are emitted. The two horizontal lines in Fig. 10.11 (top left) mean that interaction is switched on at time  $t = 0$  (lower line) and switched off at time  $t = \tau$  (upper line). That is, these lines outline the fact that we are dealing with finite-time QED. The diagrammatic representation of the first-order

contribution is shown in Fig. 10.11 (bottom left): the system starts from the state  $|g, 0\rangle$  and performs a transition to the state  $|e, 1\rangle$ , with the excited state  $|e\rangle$  of the qubit represented by two parallel vertical lines and the state with a single emitted (real) photon represented by a wavy line. The interaction vertex is represented by a full circle. Note that this diagram is beyond the RWA, since the energy is not conserved: both the qubit and the oscillator start from their ground states and are eventually excited. The diagrams representing the two second-order terms are shown in Fig. 10.11 (right). Notice that in the first case (top right diagram) the photon is virtual, while in the second (bottom right diagram) two real photons are emitted. The perturbative treatment here outlined can be easily iterated to higher orders.

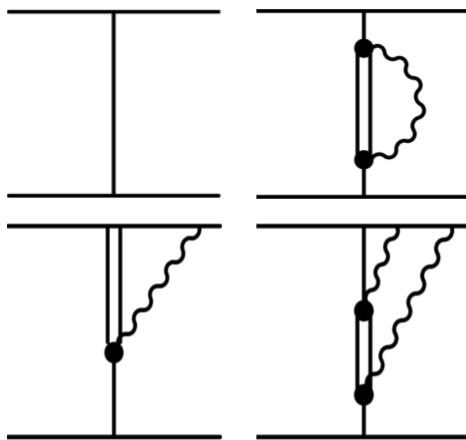


Fig. 10.11 Diagrammatic representation of the zeroth-order (top left), first order (bottom left) and second-order (right) contributions in the Picard series. Note that in the top right diagram the photon is virtual, in the bottom right diagram both photons are real. In both cases at the end the qubit is left in its ground state.

## 10.4 Quantum communication with photons

At present, the only appropriate physical system for long-distance communication of quantum states is the photon. Photons can travel long distances with low loss in optical fibres or even in free space. Furthermore, the state of a single photon can be manipulated using basic linear optical components; that is, phase shifters and beam splitters, which we shall discuss in Sec. 10.4.1. The purpose of this section is to present the basic principles of the experimental implementations of quantum cryptography with photons. Before doing so, a short introduction to linear optics is required.

### 10.4.1 Linear optics

An optical component is said to be linear if its output modes (with creation and annihilation operators  $b_j^\dagger$  and  $b_j$ ) are a linear combination of its input modes (with creation and annihilation operators  $a_k^\dagger$  and  $a_k$ ):

$$b_j^\dagger = \sum_k M_{jk} a_k^\dagger. \quad (10.45)$$

#### Phase shifter

This is defined by the transformation

$$U_P(\phi) = e^{i\phi m} = e^{i\phi a^\dagger a}. \quad (10.46)$$

Therefore, the Fock state  $|m\rangle$  is mapped into  $e^{i\phi m}|m\rangle$ . In practice, a phase shifter is a slab of transparent medium with refractive index  $n$  different from the free space refractive index  $n_0$ . Hence the wave vectors in the medium and in free space are  $k = n\omega/c$  and  $k_0 = n_0\omega/c$ , where  $\omega/2\pi$  is the photon frequency and  $c$  the speed of light in the vacuum. If the photon travels a distance  $L$  through the medium, its phase changes by  $e^{ikL}$ , which is different from the phase change  $e^{ik_0 L}$  for a photon travelling the same distance in free space. The phase shift  $\phi$  in (10.46) is then  $kL$  for the photon travelling in the medium and  $k_0 L$  for the photon travelling in free space.

#### Beam splitter

By definition, a beam splitter acts on two modes through the unitary transformation

$$U_B(\theta, \phi) = \begin{bmatrix} \cos \theta & -e^{i\phi} \sin \theta \\ e^{-i\phi} \sin \theta & \cos \theta \end{bmatrix}, \quad (10.47)$$

where the input and output modes are related through the linear mapping

$$a_l^\dagger |0\rangle \rightarrow \sum_m (U_B)_{ml} b_m^\dagger |0\rangle. \quad (10.48)$$

In particular, given the input state

$$|mn\rangle = \frac{(a_1^\dagger)^m}{\sqrt{m!}} \frac{(a_2^\dagger)^n}{\sqrt{n!}} |00\rangle, \quad (10.49)$$

we obtain the output state

$$\begin{aligned} U_B |mn\rangle &= \frac{1}{\sqrt{m!n!}} \left[ \sum_{i=1}^2 (U_B)_{i1} b_i^\dagger \right]^m \left[ \sum_{j=1}^2 (U_B)_{j2} b_j^\dagger \right]^n |00\rangle \\ &= \frac{1}{\sqrt{m!n!}} (\cos \theta b_1^\dagger + e^{-i\phi} \sin \theta b_2^\dagger)^m (-e^{i\phi} \sin \theta b_1^\dagger + \cos \theta b_2^\dagger)^n |00\rangle. \end{aligned} \quad (10.50)$$

For instance,

$$\begin{aligned}
 U_B|00\rangle &= |00\rangle, \\
 U_B|10\rangle &= \cos\theta|10\rangle + e^{-i\phi}\sin\theta|01\rangle, \\
 U_B|01\rangle &= -e^{i\phi}\sin\theta|10\rangle + \cos\theta|01\rangle, \\
 U_B|11\rangle &= -\sqrt{2}e^{i\phi}\sin\theta\cos\theta|20\rangle + \cos 2\theta|11\rangle + \sqrt{2}e^{-i\phi}\sin\theta\cos\theta|02\rangle, \\
 U_B|20\rangle &= \cos^2\theta|20\rangle + \sqrt{2}e^{-i\phi}\sin\theta\cos\theta|11\rangle + e^{-2i\phi}\sin^2\theta|02\rangle, \\
 U_B|02\rangle &= e^{2i\phi}\sin^2\theta|20\rangle - \sqrt{2}e^{i\phi}\sin\theta\cos\theta|11\rangle + \cos^2\theta|02\rangle. \tag{10.51}
 \end{aligned}$$

**Exercise 10.17** In the *dual-rail representation* a single photon can follow two different paths and the two states of the qubit ( $|0\rangle$  and  $|1\rangle$ ) correspond to the photon following one path or the other (see Fig. 10.12). The two logical states can be written as  $|0\rangle = a_0^\dagger|0\rangle_0|0\rangle_1 = |1\rangle_0|0\rangle_1$  and  $|1\rangle = a_1^\dagger|0\rangle_0|0\rangle_1 = |0\rangle_0|1\rangle_1$ , where the operators  $a_0^\dagger$  and  $a_1^\dagger$  create a photon in the input modes 0 and 1 and  $|0\rangle_0$ ,  $|0\rangle_1$  are the vacuum states corresponding to these modes. A beam splitter (see Eq. (10.47) with  $\theta = \frac{\pi}{4}$  and  $\phi = -\frac{\pi}{2}$ ) implements the transformation  $|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0'\rangle + i|1'\rangle)$  and  $|1\rangle \rightarrow \frac{1}{\sqrt{2}}(i|0'\rangle + |1'\rangle)$ , where  $|0'\rangle = b_{0'}^\dagger|0\rangle_{0'}|0\rangle_{1'} = |1\rangle_{0'}|0\rangle_{1'}$  and  $|1'\rangle = b_{1'}^\dagger|0\rangle_{0'}|0\rangle_{1'} = |0\rangle_{0'}|1\rangle_{1'}$ . Here  $b_{0'}^\dagger$  and  $b_{1'}^\dagger$  create a photon in the output modes  $0'$  and  $1'$ . Show that this beam splitter, together with two  $-\frac{\pi}{2}$  phase shifters, implements a Hadamard gate (see Fig. 10.12, left).

We can also introduce the *polarization qubit*: the two polarization states  $|h\rangle$  and  $|v\rangle$  stand for the states  $|0\rangle$  and  $|1\rangle$ . Show that the CNOT gate is implemented (up to a sign factor) by the circuit in Fig. 10.12 (right), provided the dual-rail qubit is the control and the polarization qubit the target and that a polarization rotator ( $|h\rangle \rightarrow |v\rangle$  and  $|v\rangle \rightarrow -|h\rangle$ ) is placed in the upper ( $1'$ ) path (see Cerf *et al.*, 1998).

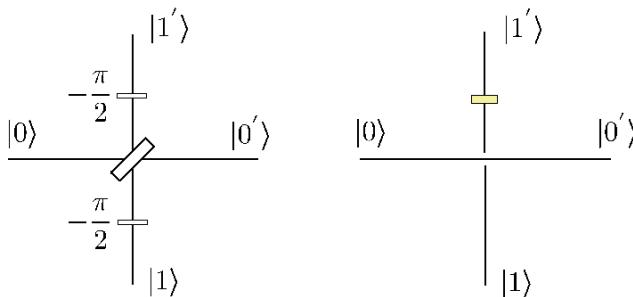


Fig. 10.12 Optical simulation of Hadamard (left) and CNOT (right) gates.

**Exercise 10.18** The two beams emerging from the beam splitter in Fig. 10.12 can be recombined using perfectly reflecting mirrors and another beam splitter. This is the principle of the *Mach-Zehnder interferometer* drawn in Fig. 10.13, an optical tool used to measure small phase shifts between the two paths connecting

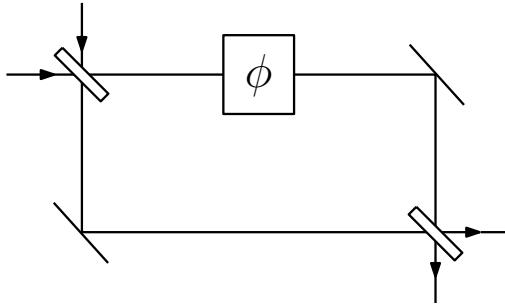


Fig. 10.13 The Mach–Zehnder interferometer. The two beam splitters stand for the circuit in Fig. 10.12 (left); that is, they implement Hadamard gates.

the two beam splitters. Show that, if a phase shifter is put into one arm of the interferometer, then the entire circuit is equivalent to a single beam splitter of arbitrary transmittivity (the transmittivity  $T$  and the reflectivity  $R$  in (10.47) are defined as  $T = \cos^2 \theta$ ,  $R = 1 - T = \sin^2 \theta$ ).

As discussed in Sec. 3.7, we can decompose any unitary operator  $U$  acting on a  $N = 2^n$ -dimensional Hilbert space into the product of  $O(N^2)$  operations, each only acting non-trivially on two-dimensional subspaces. More precisely, we can write (see Reck *et al.*, 1994)

$$U = DV_{2,1}V_{3,1}V_{3,2}V_{4,1} \cdots V_{4,3} \cdots V_{N,1}V_{N,2} \cdots V_{N,N-2}V_{N,N-1}, \quad (10.52)$$

where  $V_{p,q}$  differs from the  $N$ -dimensional identity matrix only in the matrix elements  $qq, qp, pq, pp$ , here given by the beam splitter matrix (10.47), and  $D$  is a  $N \times N$  diagonal matrix, with diagonal matrix elements of unit modulus. As shown in Fig. 10.14, transformation (10.52) can be implemented by means of a triangular array of  $\frac{N(N-1)}{2}$  beam splitters plus  $N$  phase shifters. The top left beam splitter in this figure realizes  $V_{N,N-1}$  and so on up to the top right beam splitter, realizing  $V_{2,1}$ . Finally, the phase shifters implement the diagonal matrix  $D$ . Therefore, any  $n$ -qubit quantum circuit can be simulated by a single-photon optical setup with  $N = 2^n$  optical paths. Note that the number of optical devices (beam splitters and phase shifters) grows exponentially with  $n$ . This is the price to pay because qubit–qubit interactions are not included in this model; that is, entanglement is not generated (see also Sec. 3.3 for a discussion of the importance of entanglement in quantifying the resources required for computation with waves).

### 10.4.2 Non-linear optics and probabilistic gates

In *non-linear optics* the two-qubit CMINUS gate could, in principle, be implemented by taking advantage of the indirect interaction between photons, mediated by atoms in a *Kerr medium*. As a result, the refractive index  $n$  is a linear function of the total intensity  $I$  of light crossing the medium ( $n(I) = n_0 + n_2 I$ ), so that an extra phase shift  $\phi \propto n_2 L$  is acquired if two photons propagate simultaneously through a Kerr

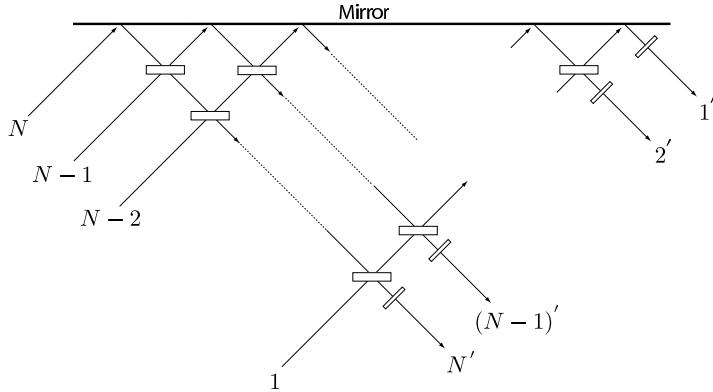


Fig. 10.14 A linear optics network implementing any  $N \times N$  unitary matrix.

medium of length  $L$ . If the medium is long enough to obtain  $\phi = \pi$  and the case in which both photons cross the medium corresponds, in the dual-rail representation, to the two-qubit state  $|11\rangle$ , then the CMINUS = CPHASE( $\pi$ ) gate is realized. The drawback is that in Kerr media it is difficult to obtain  $\phi = \pi$  before photon loss due to absorption becomes important.

#### Probabilistic quantum gates

As shown in Knill *et al.* (2001), see also Raussendorf and Briegel (2001), linear optics could be used in principle to implement an efficient quantum computation, provided we can detect photons and feed the results of measurements back to control future linear gates. This leads to *probabilistic quantum gates*, see for instance exercises 10.19 and 10.20. Even though these gates are not unitary, it is possible, using quantum teleportation and quantum error correction as basic ingredients, to approximate unitary operations efficiently.

**Exercise 10.19** *Non-linear sign shift.* Let us consider the quantum circuit drawn in Fig. 10.15, where the initial state is

$$|\psi\rangle|1\rangle|0\rangle = (\alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle)|1\rangle|0\rangle = \left(\alpha + \beta a_1^\dagger + \gamma \frac{(a_1^\dagger)^2}{\sqrt{2}}\right) a_2^\dagger |000\rangle \quad (10.53)$$

and the unitary transformation

$$U = \begin{bmatrix} 1-\sqrt{2} & \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{\sqrt{2}}-2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}-\frac{1}{\sqrt{2}} \\ \sqrt{\frac{3}{\sqrt{2}}-2} & \frac{1}{2}-\frac{1}{\sqrt{2}} & \sqrt{2}-\frac{1}{2} \end{bmatrix}. \quad (10.54)$$

Note that  $U$  can be realized using beam splitters and phase shifters as in Fig. 10.14. The circuit is *probabilistic*; that is, we accept the output  $|\psi'\rangle$  if and only if we

measure a single photon in mode 2 (second line in the circuit) and vacuum in mode 3 (lower line). Show that this measurement outcome is obtained with probability  $\frac{1}{4}$  and that

$$|\psi'\rangle = \alpha|0\rangle + \beta|1\rangle - \gamma|2\rangle. \quad (10.55)$$

The state  $|\psi'\rangle$  only differs from  $|\psi\rangle$  in the sign of the coefficient in front of the two-photon state  $|2\rangle$ . The transformation  $\alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle \rightarrow \alpha|0\rangle + \beta|1\rangle - \gamma|2\rangle$  is known as a non-linear sign-shift gate.

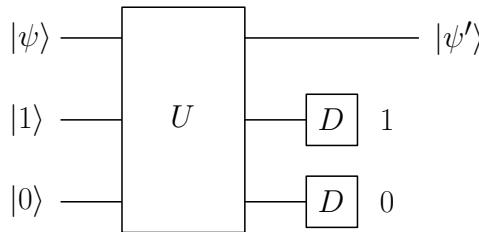


Fig. 10.15 A quantum circuit implementing the non-linear sign-shift gate.

**Exercise 10.20** Show that the circuit in Fig. 10.16 implements a probabilistic CMINUS gate, the probability of success being  $\frac{1}{16}$ .

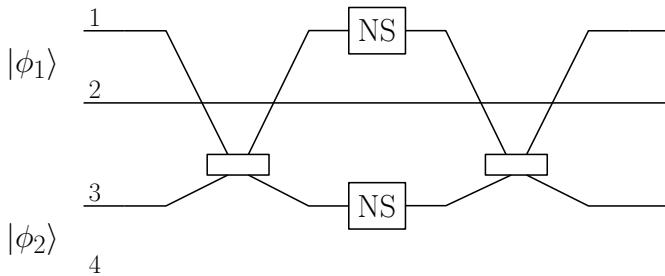


Fig. 10.16 A quantum circuit implementing a probabilistic CMINUS gate using two non-linear sign shift gates (NS) and two beam splitters with  $\phi = 0$  and  $\theta = \frac{\pi}{4}$  (left) or  $\theta = -\frac{\pi}{4}$  (right). The initial states of the two qubits are encoded in the dual-rail representation:  $|\phi_1\rangle = \alpha|0\rangle_1|1\rangle_2 + \beta|1\rangle_1|0\rangle_2$  and  $|\phi_2\rangle = \gamma|0\rangle + \delta|1\rangle = \gamma|0\rangle_3|1\rangle_4 + \delta|1\rangle_3|0\rangle_4$ .

### Parametric down-conversion

Entangled EPR states can be generated by *parametric down-conversion*. This phenomenon takes place when a laser beam passes through a non-linear crystal such as  $\beta$ -barium borate (BBO). Inside the crystal, an incoming pump photon can be converted into two photons of lower energy, one polarized vertically and the other polarized horizontally, conserving total energy and momentum.<sup>6</sup> In so-called type II

<sup>6</sup>If  $(\omega_p, \mathbf{k}_p)$ ,  $(\omega_1, \mathbf{k}_1)$ ,  $(\omega_2, \mathbf{k}_2)$  denote angular frequencies and wave vectors of the pump photon and of the two down-converted photons, then the relations  $\omega_p = \omega_1 + \omega_2$  and  $\mathbf{k}_p = \mathbf{k}_1 + \mathbf{k}_2$  hold.

down-conversion the photons are emitted along two cones (one photon per cone, see Fig. 10.17), corresponding to horizontally and vertically polarized photons. If the two photons travel along the cone intersections, neither photon has definite polarization. This corresponds to the entangled state

$$\frac{1}{\sqrt{2}}(|v\rangle_1|h\rangle_2 + e^{i\alpha}|h\rangle_1|v\rangle_2), \quad (10.56)$$

where  $|h\rangle_i$  and  $|v\rangle_i$  denote the horizontal and vertical polarization states of photon  $i$  ( $i = 1, 2$ ) and the relative phase  $\alpha$  arises from the crystal birefringence.

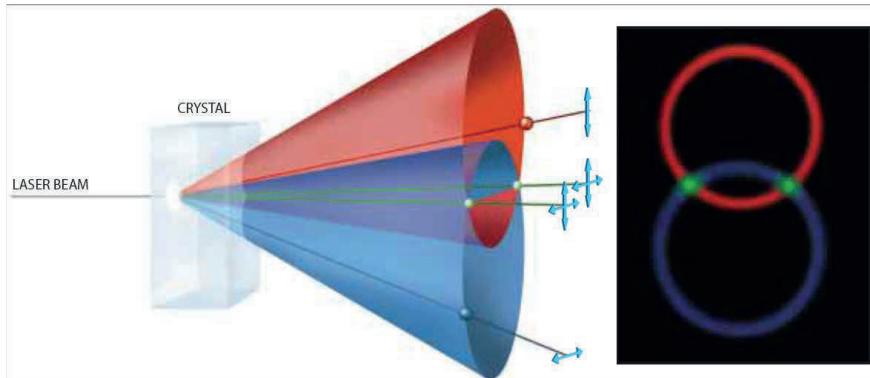


Fig. 10.17 Generation of entangled states by parametric down-conversion. Image courtesy of Anton Zeilinger, Wien.

Such entangled states are at the basis of teleportation experiments with photons (starting from Bouwmeester *et al.*, 1997 and Boschi *et al.*, 1998) and of quantum cryptographic experiments using EPR photon pairs (see below).

#### 10.4.3 Experimental quantum-key distribution

Quantum cryptography (or more precisely, quantum-key distribution) is the first quantum-information protocol with commercial applications, thanks to the enormous progress in the technology of optical-fibres and free-space optical communication.

Discrete-variable optical quantum cryptography is based on single-photon Fock states, emitted on demand. Such states are difficult to realize experimentally and are therefore approximated by attenuated laser pulses (see Sec. 5.7). Single photons are typically detected by means of semiconductor avalanche photodiodes (APD's).

##### Fibre-, ground-, and satellite-based systems

Photons can be transmitted from the sender (Bob) to the receiver (Alice) using *optical fibres* or *free space* (ground- or satellite-based quantum communication) as quantum channels (of course, such channels are only described as quantum because they are intended to transmit the quantum information encoded in single photons).

Let us briefly discuss the advantages and drawbacks of both approaches. Long-distance optical-fibre transmissions exploit the low loss of silica fibres in the 1.3 and 1.55  $\mu\text{m}$  wavelength bands. A further advantage is the possibility to employ standard fibres installed for classical communications. On the other hand, free-space applications are also possible. In this case, the emitter and the receiver are connected by telescopes pointing at each other (spectral filtering is used by the receiver to cut light outside the transmission bandwidth). A significant advantage of free-space quantum cryptography is that transmission over long distances is possible in a transmission window around 800 nm. In this window commercial photo-detectors (silicon avalanche photodiodes) have high detection efficiency. In contrast, at the wavelengths used in optical fibres silicon APD's are not efficient and one can take advantage of APD's made from germanium or indium gallium arsenide (with lower detection efficiency). A disadvantage of free-space quantum cryptography is that its performance strongly depends on weather conditions and air pollution. On the other hand, a major advantage of this approach is that it could offer the possibility to overcome the distance limitations of fibre-based quantum cryptography. Actually, the main drawback with quantum communication via optical fibres is that the probability for photon-absorption losses grows exponentially with the length of the fibre. On the basis of present technology, it appears difficult to employ optical fibres for quantum communication over distances of a few 100 kilometres.<sup>7</sup>

It is important to stress that, besides the practical interest of establishing a global-scale quantum-cryptographic network in the future, long-distance quantum communication is important also to investigate the validity of quantum laws. Indeed, it remains an open issue whether quantum laws, developed to describe phenomena in the microscopic domain, are applicable in the macroscopic domain, as for long distances. For instance, various proposals predict that entanglement is altered under specific gravitational fields (see Joshi *et al.*, 2018). Since quantum gravity is a theory still under development (see Rovelli, 2004, Thiemann, 2007, Rovelli and Vidotto, 2014, and Donoghue *et al.*, 2017), the issue is controversial, for a review see Bassi *et al.* (2017). Experiments using a 144 km free-space link between La Palma and Tenerife addressed this question (Ursin *et al.*, 2007, see Fig. 10.18). A source of entangled photons was installed in La Palma at an altitude of 2400 m and one of the photon of each entangled pair was sent to Tenerife and there received by a telescope, also at an altitude of 2400 m. The analysis of the polarization correlations between the photons showed that the photons remained entangled, even though they had been separated by a long distance. Moreover, quantum key distribution protocols using entangled and single photons were implemented. Long-distance free-space quantum teleportation experiments were also reported, over a 97 km link across the

---

<sup>7</sup>Quantum repeaters; that is, quantum purification schemes aimed at improving the fidelity of the transmitted photons (see Briegel *et al.*, 1998), would overcome this limitation. In principle, quantum repeaters could extend quantum communication to arbitrarily long distances. However, the development of quantum repeaters is yet in the early stages.

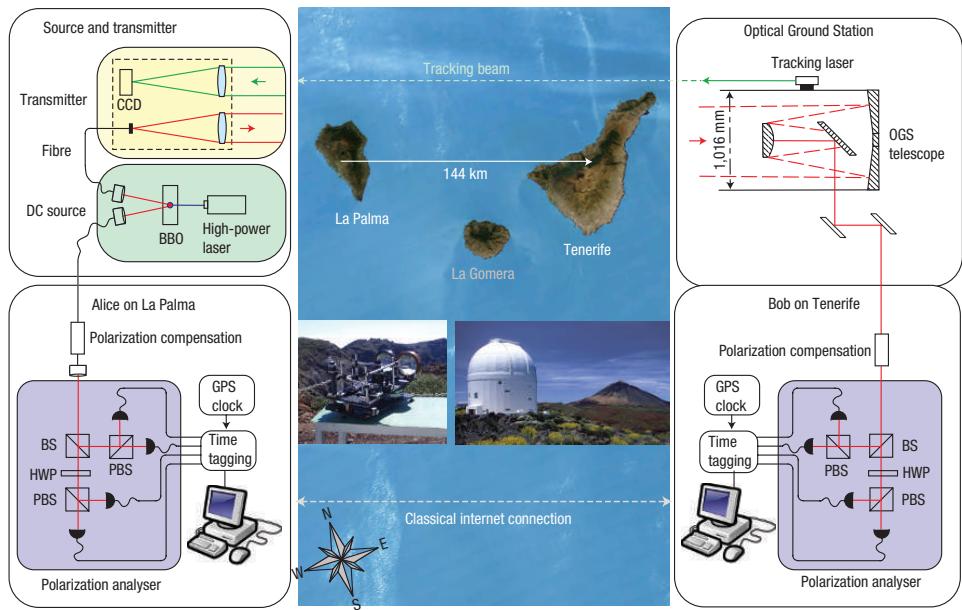


Fig. 10.18 Schematic illustration of the experimental setup in the distribution of entangled photons over 144 km between La Palma and Tenerife. The figure is reprinted with permission from Ursin *et al.* (2007). ©(2007) Macmillan Publishers Ltd.

Qinghai lake (Yin *et al.*, 2012) and over 143 km between La Palma and Tenerife (Ma *et al.*, 2012).

Direct links on significantly longer distances are not possible on ground, since they are prevented by the curvature of the earth. The development of quantum technologies into space promises to overcome such limitation. To set up a global scale quantum communication network, a satellite (nicknamed Micius) has been used to demonstrate satellite-based distribution of entangled photon pairs to two ground stations (Delingha and Lijiang in China) separated by 1203 Km (Yin *et al.*, 2017). More recently, another quantum secure communication backbone network has been built, from Beijing to Shanghai, with a fiber distance exceeding 2000 km (Zhang *et al.*, 2018). This opens the possibility to employ free-space photon transmission to distribute secret keys between parties located very far apart (say, in two different continents), using satellite-based links. Moreover, tests of the nonlocality of quantum mechanics can be extended to distances larger than 1000 km, with the distribution of entangled photons from the satellite to two distant ground stations (see Fig 10.19). Furthermore, a ground-space quantum link allows one to probe how gravity acts on the quantum properties of lights, studying for instance the propagation of an entangled state through a gravity gradient. At the ground to satellite distance the curvature of spacetime starts to become important and would affect the

polarisation of photons and hence quantum entanglement measurements.<sup>8</sup> Ground-to-satellite teleportation experiments based on entangled photons have been realized over distances of up to 1400 Km (Reu *et al.*, 2017). Following the first successful experiments, the launch of further Micius satellites is planned, with the purpose of establishing a European-Asian quantum-encrypted network by 2020, and a global network by 2030.

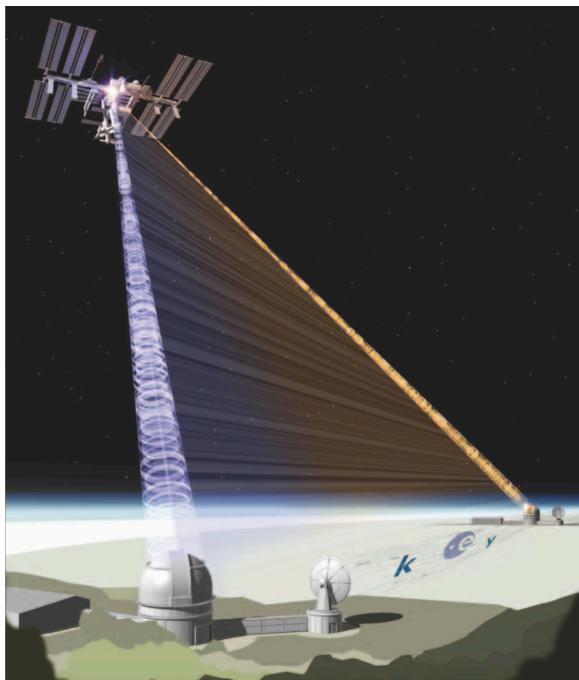


Fig. 10.19 Schematic illustration of entanglement distribution using satellites. Entangled photon pairs are simultaneously distributed to two separated locations on Earth. Drawing courtesy of ESA (European Space Agency).

#### Polarization and phase coding

A natural method to code the four states of the BB84 protocol is to employ photons polarized at  $-45^\circ$ ,  $0^\circ$ ,  $+45^\circ$  and  $90^\circ$ . For each pulse, Alice can rotate the polarization of one of these four states by means of electro-optic crystals (Pockels cells). Bob analyzes each photon in the vertical–horizontal basis or in the diagonal basis. If, for instance, a photon polarized at  $+45^\circ$  is sent and the measurement takes place in the diagonal basis, then the outcome is deterministic. On the other

<sup>8</sup>It might be useful to remark that in Global Positioning Systems (GPS) relativistic effects such as gravitational redshift and time dilation cannot be neglected. The frequency differences between clocks in orbit, and reference clocks on earth's surface, should be taken into account for navigation using GPS.

hand, if Bob chooses the horizontal-vertical basis, he randomly obtains one of the two possible outcomes. The main difficulty of this scheme is to maintain the photon polarization through the quantum channel connecting Alice and Bob. It is difficult to compensate for the polarization transformation induced by a long optical fibre since it is unstable over time, due, for instance, to temperature variations. Note that polarization coding is instead successful when used for free-space transmission over long distances.

Phase coding has proved to be more convenient for fibre-based implementations. The basic setup, drawn in Fig. 10.20, is an optical-fibre version of the Mach–Zehnder interferometer. It consists of two symmetric couplers (the equivalent of 50:50 beam-splitters) connected by two arms, each with a phase modulator (that is, a phase shifter). Alice and Bob can tune the phase shifts  $\phi_A$  and  $\phi_B$ , respectively. The “letters” used in the BB84 protocol (see Sec. 5.3.1) correspond to  $\phi = 0, \pi$  (first “alphabet”) and  $\phi = \frac{\pi}{2}, \frac{3\pi}{2}$  (second alphabet). Alice randomly applies one of the above four phase shifts to encode a bit value (she associates 0 and  $\frac{\pi}{2}$  with bit value 0 and  $\pi$  and  $\frac{3\pi}{2}$  with 1). On the other hand, Bob randomly chooses the measurement basis by applying a phase shift of either 0 or  $\frac{\pi}{2}$ . When  $|\phi_A - \phi_B| = 0, \pi$ , then Bob obtains with unit probability a deterministic output (see exercise 10.18). On the other hand, when the phase difference is equal to  $\frac{\pi}{2}$  or  $\frac{3\pi}{2}$ , then the photon is found with equal probability in one of Bob’s two detectors.

Note that the phase-coding scheme works inasmuch as the path mismatch  $k\Delta L$  ( $k$  is the wave number and  $\Delta L$  the difference between the lengths of the two arms) is much smaller than the photon wavelength (of order  $1\text{ }\mu\text{m}$ ). This condition cannot be fulfilled when Alice and Bob are separated by long distances. For this reason the configurations in Fig. 10.21 with two unbalanced Mach–Zehnder interferometers is

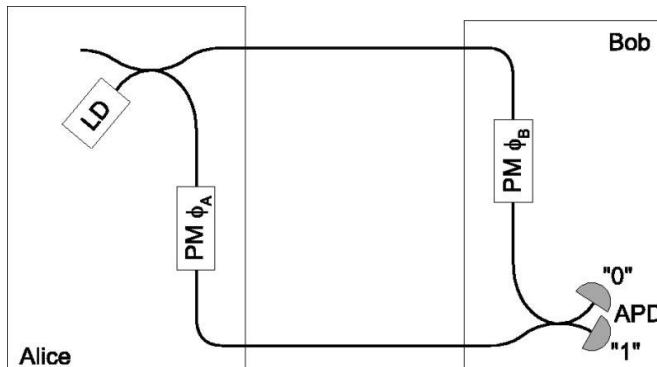


Fig. 10.20 A schematic drawing of an optical-fibre Mach–Zehnder interferometric setup for quantum cryptography. Photon pulses are emitted by a laser diode (LD) and then attenuated and sent from Alice to Bob by means of optical fibres; phase modulators (PM) of phase  $\phi_A$  and  $\phi_B$  are used in Alice’s and Bob’s laboratories; an avalanche photodiode (APD) is used to detect the photon in port 0 or 1. The figure is reprinted with permission from Gisin *et al.* (2002). ©(2002) by the American Physical Society.

used. In this case the two interferometers, one in Alice's laboratory and the other in Bob's, are connected by a *single* optical fibre. When monitoring counts as a function of time from photon emission, Bob observes three peaks (see the inset in Fig. 10.21). The left/right peak corresponds to photons that travel along the short/long path in both Alice's and Bob's interferometers. The central peak is instead associated with photons that choose the long path in Alice's interferometer and the short path in Bob's or *vice versa*. As these two processes are indistinguishable, they produce the interference required in the phase-coding scheme. The advantage of this system is that it is sufficient to keep stable within a small fraction of the photon wavelength the imbalances of Alice's and Bob's interferometers and not the path difference over a long distance as in the previous scheme of Fig. 10.20.

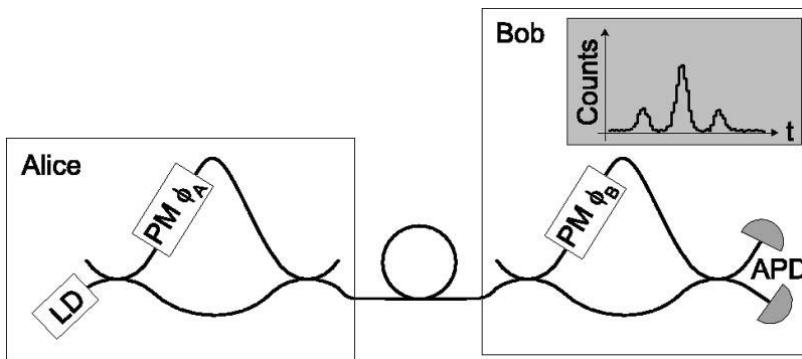


Fig. 10.21 A schematic drawing of a double Mach–Zehnder interferometer for quantum cryptography. The inset shows the temporal count distribution recorded as a function of the time passed since emission of a pulse by Alice (interference is observed in the central peak). The figure is reprinted with permission from Gisin *et al.* (2002). ©(2002) by the American Physical Society.

In closing this section, we mention that also with continuous variables it was possible to experimentally demonstrate long-distance quantum key distribution, see Jouguet *et al.* (2013). A major advantage of continuous-variable quantum cryptography is that only standard telecommunication technology is required. In particular, dedicated single-photon counters (necessary for discrete-variable quantum key distribution) are replaced by homodyne detectors (see Sec. 5.7), which are widely used in classical optical communications and have a much higher efficiency.

## 10.5 Problems and prospects

Quantum cryptography is the first quantum-information protocol to find commercial applications. Here the question is how extensive the market will be and this will largely depend on the transmission rates for long-distance quantum communication. The development of fast single-photon sources and high-efficiency detectors is required to improve significantly the transmission rates, thus broadening the prospects

of quantum cryptography. A significant step is also the development of satellite-based quantum communications, necessary to establish a global quantum-encrypted network.

With regard to quantum computation, the situation is much more difficult. It is not clear if and when we shall be able to build a useful quantum computer; that is, a quantum computer capable of outperforming existing classical computers in important computational tasks. When the problem of decoherence is taken into account for a complex many-qubit system, which we require to perform coherent controlled evolution, then large-scale quantum computers appear unrealistic with present technology. On the other hand, we should bear in mind that technical breakthroughs (such as the transistor was for the classical computer) are always possible and that no fundamental objections have been found against the possibility of building a quantum computer. Moreover, the simulation of many important and complex problems in condensed-matter physics, quantum chemistry, and high-energy physics could be implemented not only with a general purpose quantum computer, but also with simpler, analog devices, known as *quantum simulators*, whose principles and prospects shall be discussed in Chap. 11.

At any rate, even the first, few-qubit demonstrative experiments are remarkable, not only for quantum computation but also for addressing fundamental questions on quantum mechanics, such as the nature of the frontier between quantum and classical worlds or the nature of quantum entanglement in complex many-body systems.

It is also important to emphasize that basic research in the field of quantum information is strictly related to the emergence of *quantum technologies* such as quantum based sensors and clocks. For instance, entangled states could be used to improve the resolution of optical lithography and interferometric measurements.

## 10.6 A guide to the bibliography

The experimental effort in the field of quantum-information processing is huge and has produced beautiful experimental results. In the following, we shall limit our references to review papers that might be used by the reader as an entry point.

Various aspects of cavity quantum electrodynamics experiments manipulating Rydberg atoms and photons (including the generation of Fock states, atom-field entanglement, implementation of quantum gates, “Schrödinger’s cat” states of the field and emergence of classical behaviour due to decoherence effects) are reviewed in Raimond *et al.* (2001), Walther *et al.* (2006), and Haroche and Raimond (2006). Quantum computation with trapped ions is discussed in Häffner *et al.* (2008) and Blatt and Wineland (2008). Tutorial reviews on quantum information processing with atoms, ions and photons are Monroe (2002) and Cirac and Zoller (2004).

Spin qubits in semiconductor quantum dots are described in Hanson *et al.* (2007). A very readable introduction is Burkard and Loss (2002). Prospects for quantum dot-based quantum computing are discussed in Kloeffel and Loss (2013).

Superconducting quantum bits are reviewed in Clarke and Wilhelm (2008); Devoret and Schoelkopf (2013); for a simple introduction see You and Nori (2005); for a discussion of the state of the art and prospects of this implementation see Gambetta *et al.* (2017). On circuit QED, see Blais *et al.* (2004); Wallraff *et al.* (2004) and for a review Schoelkopf and Girvin (2008). For a discussion of the DCE in superconducting circuits, see Nation *et al.* (2012).

Linear optic quantum computation is discussed in Myers and Laflamme (2006) and Kok *et al.* (2007). Quantum optics implementations with continuous variables are reviewed in Braunstein and van Loock (2005). Quantum cryptography is reviewed in Gisin *et al.* (2002); for a discussion of experimental progress in ground- and space-based quantum communication see Krenn *et al.* (2016).

## Chapter 11

# Quantum information in many-body systems

This final chapter is devoted to an application of basic quantum information theory concepts in the context of strongly correlated systems and, more in general, of condensed matter physics. Recent groundbreaking experiments in the spirit of Feynman's idea of quantum simulators enabled to manipulate certain quantum many-body systems in a very accurate and controlled way, thus paving the way for a tremendously increasing theoretical activity in this field. We will see that a closer look to the intimate entanglement structure of typical many-body wave functions may reveal itself extremely helpful to understand the physics of such systems. Indeed, on this basis, it has been possible to conceive powerful analytical and numerical methods to uniquely characterize the thermodynamical properties of a wide class of complex quantum systems.

After glimpsing at how non-trivial quantum frustration mechanisms may emerge in such context, we shall consider in detail the spin-1/2 quantum Ising chain, and thoroughly present its analytic solution. This paradigmatic example of quantum many-body system provides a convenient playground where it is possible to recognize the emergence of a very peculiar behaviour for the ground-state entanglement, far from being specific to a single model. The latter observation can be formalized through the concept of the area-law scaling of bipartite quantum correlations for the low-energy states of generic locally interacting strongly correlated systems.

The second part of the chapter is devoted to an introduction of the so-called tensor-network formalism for quantum many-body lattice systems, a tool that is able to provide a reliable Ansatz for any wave function satisfying certain entanglement constraints in the Hilbert space. Emphasis is put on the simplest case of one-dimensional systems, where the so-called matrix product states (MPS) naturally emerge as a consequence of the area-law requirement for the bipartite entanglement. We shall discuss the density-matrix renormalization group algorithm, a variational technique for finding the ground state over the MPS class of wave functions. We also show how to manipulate such states to study real-time evolution, finite-temperature states, and systems coupled to an external environment. Finally we briefly focus on tensor-network schemes for higher dimensions, and on hierarchical structures which may violate the area-law constraint.

## 11.1 Quantum simulators

Understanding the thermodynamic behaviour of many-body quantum systems is a formidable task which tantalized the attention of many generations of physicists, starting from the dawn of quantum mechanics. This observation is particularly urgent in the context of non-equilibrium dynamics. Unfortunately up to a decade ago, even conceiving any experiment of controlled dynamics at the nanoscale was mostly seen as a chimera, due to the extreme difficulty in tailoring and controlling the physical setup. For this reason, the problem covered a purely academic interest.

The advent of the so-called *quantum simulators* profoundly changed this scenario: they enabled to realize a new concept of quantum devices originally devised by Feynman (1982), which could “mimic the physics of other quantum systems”. Such devices have unveiled the possibility to operate in the lab at the level of the single quantum object, with an exceptionally low degree of decoherence and high tunability. It is now possible to carefully probe the interplay between dimensionality, interactions and quantum coherence in a wealth of different physical implementations, from Josephson-junction arrays, to ultracold quantum gases, semiconductor nanostructures, and coupled quantum electrodynamical (QED) cavities. In some cases one may even engineer the dissipation in a controlled way, as a resource to probe the behaviour of driven-dissipative systems.

In synthesis, quantum simulators are fabricated devices that can experimentally simulate the model Hamiltonian underlying the non-trivial properties of the physical systems under consideration. The advantages of this approach are twofold. First of all, it is possible to explore the properties of strongly correlated model Hamiltonians also in the regions of the phase diagram which are elusive to numerical and analytical investigations. Secondly, it allows to test to which extent the model Hamiltonians are appropriate to treat the physical systems that they are supposed to describe, or whether additional ingredients are necessary.

### 11.1.1 *Ultracold atoms*

The first nano fabricated devices, designed with the specific purpose to simulate the physics of strongly correlated systems, were probably Josephson-junction arrays. The field however acquired full maturity only with the appearance of cold atoms trapped in magneto-optical potentials, which proved to be excellent simulators of a large variety of strongly interacting Fermi and Bose systems. The success of these latter systems, virtually immune to disorder, relies on their ultra-small coupling to the environment and to the extremely low operating temperatures, of the order of few nano Kelvin. This allows to access the equilibrium, as well as the non-equilibrium dynamics of closed many-body systems (a regime lying outside the experimental capabilities with more traditional solid-state systems).

As a matter of fact, cold atoms can be manipulated with an unprecedented accuracy. First of all, it is possible to trap them in periodic potentials that are

generated by shining two counter-propagating laser beams, which may interfere to form a modulated optical standing wave, the so-called *optical lattices*. Such artificial crystals, almost free from defects and vibrations, can be loaded with thousands of neutral atoms obeying either Bose or Fermi statistics (or even mixtures of them), such that each site may contain only few particles, of the order of one. Moreover, a good control over the collision properties of atoms can be achieved by tuning external parameters across the so-called Feshbach resonances,<sup>1</sup> making it possible to increase the two-body interaction strength and access a deep quantum regime of strongly correlated particles.

As shown by Jaksch *et al.* (1998), the prototype non-trivial Hamiltonian that can be naturally simulated by means of ultracold bosonic atoms loaded in an optical lattice is the *Bose–Hubbard model*. This is given by

$$H = -J \sum_{\langle i,j \rangle} (b_i^\dagger b_j + \text{H.c.}) + \frac{U}{2} \sum_j n_j(n_j - 1), \quad (11.1)$$

where the indices  $i$  and  $j$  label the various minima of the optical standing wave (i.e., the sites of the artificially constructed lattice), and the brackets  $\langle \cdot, \cdot \rangle$  limit the summation to nearest-neighbour pairs of sites. In the second quantization language, the operators  $b_j^{(\dagger)}$  annihilate (create) a boson on the  $j$ -th site of the lattice ( $n_j = b_j^\dagger b_j$  is the corresponding number operator, which counts how many bosons are located on the  $j$ -th site). We postpone a more rigorous theoretical analysis of such class of systems to Sec. 11.2.1, while providing here only a qualitative discussion in order to grasp its salient physical features.

The low-temperature physics of the Bose–Hubbard model is dominated by the competition between delocalization, induced by the hopping of bosons between adjacent cavities (the first term in Eq. (11.1), with strength  $J > 0$ ), and on-site repulsive interaction (the second term in Eq. (11.1), with strength  $U > 0$ ). Specifically, the ground state phase diagram consists of two phases. If the hopping term dominates over the interaction, that is  $J \gg U$ , the system enters a *superfluid* (SF) state where each atom tends to be spread over all the lattice, and the corresponding many-body wave function is a product of delocalized single-particle states:

$$|\psi\rangle_{\text{SF}} \propto \left( \sum_{j=1}^L b_j^\dagger \right)^N |0\rangle, \quad (11.2)$$

where  $N$  is the total number of bosons in a lattice of  $L$  sites, and  $|0\rangle$  is the vacuum of particles. The characteristic feature of this state is that of representing a macroscopic wave function, which possesses long-range phase coherence throughout the lattice. In the opposite limit where interactions dominate, that is  $U \gg J$ , the

---

<sup>1</sup>A Feshbach resonance describes the resonant scattering between two particles, when their incoming energies are very close to those of a two-particle bound state. It is possible to control their energy difference either via a magnetic field, or by optical methods. As a result of this process, the scattering length of the two atoms can be modified, thus providing a way to vary the interaction strength between atoms in the cloud by changing scattering length of elastic collisions. For a recent review, see Chin *et al.* (2010).

fluctuations in the atomic number on each site are very small, and the system stabilizes into a sequence of localized atomic wave functions with a fixed number of atoms per lattice site. In the case of a commensurate filling of  $n = N/L$  atoms per site, the global wave function can be written as a product of Fock states:

$$|\psi\rangle_{\text{MI}} \propto \prod_{j=1}^L (b_j^\dagger)^n |0\rangle, \quad (11.3)$$

whose resulting many-body features are that of a so-called *Mott insulator* (MI). This latter state does not entail phase coherence, and cannot be described by a macroscopic wave function.

In their seminal paper, Fisher *et al.* (1989) theoretically showed that, for integer values of  $n$ , a zero-temperature phase transition between the above mentioned SF and MI phases takes place at a finite value of the  $U/J$  ratio. On the other hand, for non-integer fillings the bosons are always superfluid. Thirteen years later, this transition was experimentally observed by Greiner *et al.* (2002): In response to a change in the relative strength of  $U$  vs.  $J$  (as obtained by tuning the optical lattice depth, through a control of the intensity of the trapping lasers), two clearly distinct types of matter wave interference pattern were spotlighted. Specifically, Fig. 11.1 displays experimentally observed absorption images of matter waves; the various panels from (a) to (h) stand for increasing values of  $U/J$ . As is clearly visible, in the SF phase interference fringes can be observed [panels (a)–(f)] while in the MI phase they completely disappear [panels (g)–(h)].

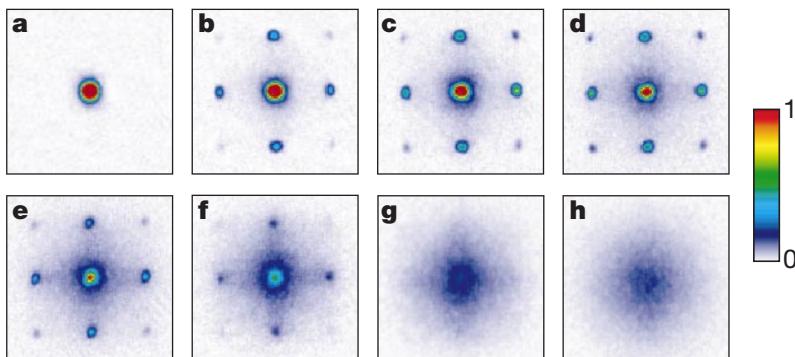


Fig. 11.1 Interference patterns of matter waves as obtained after releasing a confining three-dimensional optical lattice, and measuring the absorption of atoms on a screen at a fixed distance from the initial lattice. The various panels correspond to different optical potential depths, which tune the relative strength of  $U/J$  in Eq. (11.1). The figure is reprinted with permission from Greiner *et al.* (2002). ©(2002) Macmillan Magazines Ltd.

The extreme versatility of cold atoms also enables them to probe quantum magnetism through the direct measure of quantum correlations, thus opening up the possibility to study strongly correlated states of matter at very low temperatures. As an example, we quote the simulation of two-component fermionic models by Greif

*et al.* (2013), where nearest neighbour spin-spin correlations have been measured by means of cooling schemes based on the local redistribution of entropy within the lattice structure (see Fig. 11.2).

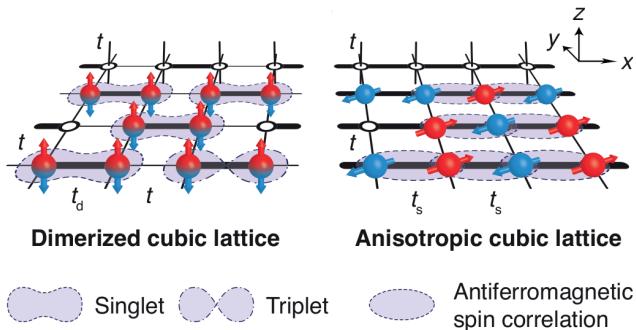


Fig. 11.2 Sketch of the magnetic spin-spin correlations that have been measured with a two-component ultracold Fermi gas with repulsive interactions, loaded in a three-dimensional cubic lattice. The system may be found in either a dimerized or an anisotropic configuration, depending on the relative strength of the tunnel terms along each spatial dimension. Drawing courtesy of Tilman Esslinger, Zurich.

However ultracold atoms are not the only available quantum simulators. As already discussed in Sec. 10.2.1, within the field of atomic and molecular optics, it is also possible to build up highly controllable systems of up to few tens of heavier ions, which can be trapped in electromagnetic potentials. In a completely different scenario, one could devise ensembles of QED cavities that are coupled together, such as to form artificial lattice structures.

### 11.1.2 Arrays of coupled QED cavities

A cavity array consists of a regular arrangement of QED cavities which, in the original conception, can be coupled through a photon-hopping mechanism. The emerging physical scenario is basically dominated by the interplay of two effects: on the one hand, light-matter interaction inside each cavity may lead to a strong effective Kerr nonlinearity between photons; on the other hand, photon hopping between neighbouring cavities favours delocalization, thus competing with the photon nonlinearity. The idea of implementing many-body states with light has been originally proposed in this context by Greentree *et al.* (2006), Hartmann *et al.* (2006) and Angelakis *et al.* (2007), showing that coupled cavities could be taken as another possibility towards the realization of quantum simulators.

Despite their young age, as compared to the more mature field of ultracold atoms in optical lattices or of ion-trap experiments, cavity arrays present a valid alternative to realize strongly correlated states of light. Indeed they can operate at high temperatures (with respect to Josephson arrays and optical lattices), and allow for single-site addressing, opening a clear way to experimentally access the

correlation functions. Most importantly, cavity arrays are specifically designed to be open-system quantum simulators, thus representing an opportunity to explore the physics of many-body quantum systems in contact with an environment. They naturally operate under non-equilibrium conditions: the population of photons that leak out because of unavoidable losses should be refilled by an external drive. Two classes of problems can be addressed: On the one hand, they offer a valid platform to understand how to realize and detect, under realistic non-equilibrium conditions, the dynamical counterpart of equilibrium phases for the underlying strongly-interacting models (Bose–Hubbard model, interacting spin models, . . . ), in situations where a steady state is formed. On the other hand, they pave the way to study the emerging non-equilibrium phase diagram in a controlled scenario, a field which is still, to a large extent, an unexplored territory. This may lead, as well, to the stabilization of exotic ordering, driven by completely new cooperative mechanisms.

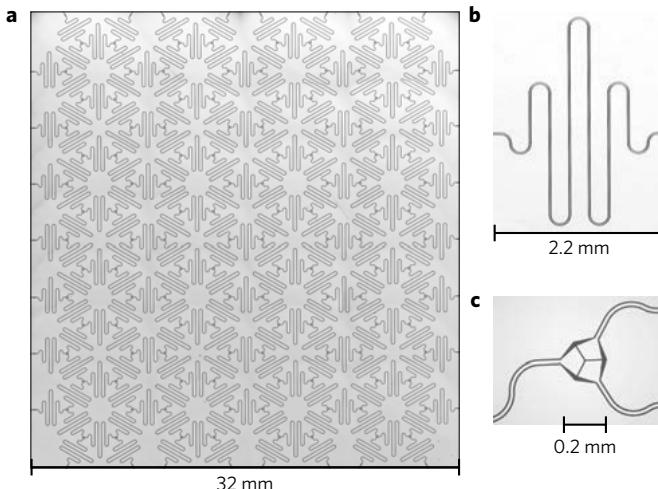


Fig. 11.3 QED cavities can be coupled in a lattice for quantum simulation. This figure shows a superconducting circuit of more than two hundred 7 GHz microwave cavities, coupled in a two-dimensional Kagome lattice structure. The junction between neighbouring cavities is realized by coupling capacitors, which enable photon hopping. Interactions between photons can be provided by adding superconducting qubits (an example is the so-called Cooper-pair box, constituted by two superconducting islands connected by two Josephson junctions, forming a superconducting quantum interference device) to the cavities, using an additional lithography layer. The figure is reprinted with permission from Houck *et al.* (2012). ©(2012) Macmillan Publishers Ltd.

There are several possible platforms that may lead to the implementation of a cavity array. One could engineer a photonic crystal, by coupling quantum dots to photonic band gap defect nanocavities. Another possibility are silicon structures of either a disc or a toroidal shape, where the light is trapped in whispering gallery modes that are localized close to the outer surface of the structure and have a small mode volume. We also mention Fabry-Pérot cavities where photons can be coupled

into and from each cavity through optical fibers, or using suitable on-chip arrays of cavity micro-mirrors. The most promising platform however seems to be based on present-day circuit-QED technology, through the use of superconducting resonators and Josephson junctions: scalable arrays of hundreds of cavities in large mesoscopic structures at the millimeter scale have been already realized (see Fig. 11.3).

### Theoretical model of a cavity array

As discussed above, the simplest Hamiltonian which describes an array of coupled cavities must include the light-matter interaction within each cavity, and the coupling between different (in many cases only neighbouring) cavities:

$$H = \sum_i H_i^{(\text{cavity})} + \sum_{\langle i,j \rangle} H_{i,j}^{(\text{coupl})}, \quad (11.4)$$

where the indices  $i$  and  $j$  label the position of a cavity on a generic  $d$ -dimensional lattice structure. The first term takes into account the dynamics of an isolated cavity, while the second one describes the interaction between cavities.

As explained in Sec 10.1.2, the minimal model of a cavity QED is the so-called Jaynes–Cummings Hamiltonian  $H_{\text{JC}}$  of Eq. (10.7). In the strong-coupling regime, light-matter interaction turns the cavity into a turnstile device, where only a photon can be present at the same time. Intuitively, this can be understood as the fact that one photon in the cavity strongly modifies the effective resonance frequency, inhibiting the injection of a second photon. This phenomenon has been termed *photon blockade*, after the Coulomb blockade effect of electrons in mesoscopic structures (see Imamōglu *et al.*, 1997). In several situations, as in circuit-QED, counter-rotating terms cannot be neglected, leading to the Rabi Hamiltonian  $H_{\text{Rabi}}$  of Eq. (7.111).

The coupling between cavities can have different origin, depending on the implementation. In the simplest scenario, if the cavities are sufficiently close to allow for photon hopping, an additional kinetic term

$$H_{i,j}^{(\text{coupl})} = -J(a_i^\dagger a_j + \text{H.c.}) \quad (11.5)$$

describes the tunnelling of a photon from the  $i$ -th to the  $j$ -th cavity, with an associated  $J$  rate. The operator  $a_j^{(\dagger)}$  annihilates (creates) a photon in the mode of the  $j$ -th cavity. In certain implementations it is even possible to couple two cavities through non-linear elements, such that  $H_{i,j}^{(\text{coupl})}$  would also contain cross-Kerr interaction terms, and/or correlated hopping terms.

The Hamiltonian in Eq. (11.4) bears strong similarities with the Bose–Hubbard model of Eq (11.1), and for this reason, it is sometimes referred to as the *Jaynes–Cummings–Hubbard model*. The similarity is evident: instead of the local on-site repulsion controlled by  $U$ , the non-linearity in the spectrum of Jaynes–Cummings–Hubbard model appears because of the light-matter interaction. Nonetheless, the competition between delocalization and local non-linearity does not crucially depend on these detailed differences.

Any realistic description of cavity arrays cannot exclude losses. In cavity systems there are numerous sources of dissipation and decoherence that need to be taken

into account, such as decoherence and relaxation of the atoms inside the cavities, and photon losses. In most relevant cases the latter are the dominant contribution, therefore here we only consider these ones. In the presence of losses, the state of the array can be faithfully described by its density matrix  $\rho$ , which obeys the GKLS master equation (7.169). Assuming that each cavity is coupled to an independent environment, the action of the dissipation can be modeled by the Lindblad operators

$$L_j = \sqrt{\kappa} a_j \quad (11.6)$$

on each cavity, with  $\kappa$  being the inverse of the photon lifetime. The fixed point of the resulting master equation is the vacuum state: eventually all the photons will have escaped from the cavities and the atoms will have decayed in their ground state. In order to refill the array with photons, an external drive is needed. If the pump is coherent it will contribute to an extra term in the Hamiltonian of the form

$$H_{\text{drive}} = \Omega \sum_j (a_j^\dagger + a_j). \quad (11.7)$$

The driving can be either continuous, or time-modulated through a laser pulse sequence. In the first case the array reaches a steady state, as a result of the interplay of the coherent and incoherent contributions to the dynamics of  $\rho$ ; in the latter case an initial population imbalance is formed, which subsequently relaxes to the vacuum state. More in general, a wealth of peculiar phenomena associated to non-equilibrium phases and phase transitions have been recently shown to emerge in such systems, as experimentally realized in Fitzpatrick *et al.* (2017). For a detailed discussion, we refer the interested reader to the literature in the bibliography.

### D-Wave device

In the last ten years, a scalable implementation of a circuit-QED quantum simulator has been realized by the private Canadian company D-Wave Systems Inc., with the intent to develop the first commercial “quantum computer” (see: <http://www.dwavesys.com>). This eventually led to the commercialization, starting from 2011, of prototype devices ranging from 128 (D-Wave One) to 2048 coupled qubits (D-Wave 2000Q), which have been purchased by big companies as, for example, Lockheed Martin, Google, and NASA.

The building block of the hardware is a so-called superconducting quantum interference device (SQUID), namely, a very sensitive magnetometer based on a superconducting loop containing Josephson junctions. The interference refers to the fact that different magnetic spin states (typically encoded in clockwise and anti-clockwise circulating current in the loop), which constitute the two levels of one qubit, can be put into a quantum mechanical superposition. In order to go from a single qubit to a multi-qubit processor, the qubits are then coupled by means of superconducting circuits. Here we will not detail the implementation of such device, but only briefly discuss its working principle.

As a matter of fact, D-Wave is not a device that implements quantum gates: it is an adiabatic quantum computer for solving optimization problems, in the spirit

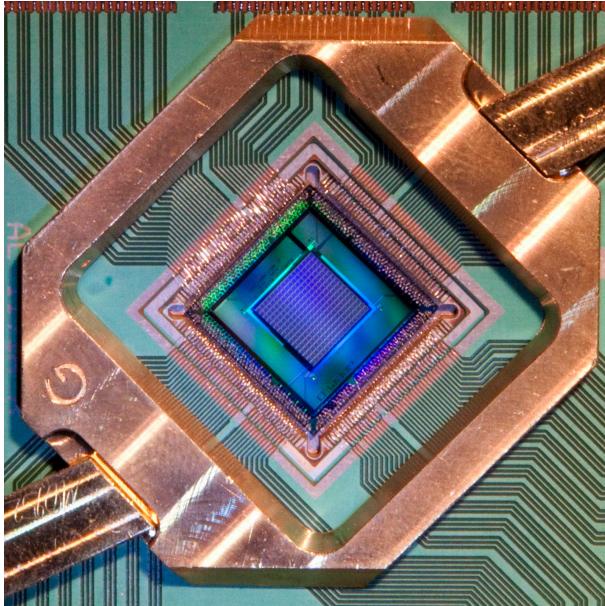


Fig. 11.4 Photograph of the D-Wave 2000Q chip, manufactured by D-Wave Systems Inc.

of what has been discussed in Sec. 3.13. Namely, the superconducting qubits are initialized into their lowest energy state, then suitable magnetic fields are gently modulated in time, in such a way as to drive this state (pseudo-)adiabatically toward a new one. This process is usually referred to as *quantum annealing*, and a machine like D-Wave is a *quantum annealer*. The possibility to realize this protocol with systems made of many qubits is particularly significant, since any quantum algorithm of arbitrary difficulty can be formulated in terms of identifying the global minimum (the ground state  $|\psi_F\rangle$ ) of a given function (the Hamiltonian  $H_F$ ) over a set of many local minima.<sup>2</sup> Provided the annealing protocol can be performed in a limited amount of time, in principle this would represent an exponential speedup with respect to any classical computation. It should be however stressed that, posing a problem in a form that D-Wave can handle, often requires several qubits to represent a single variable, thus limiting the size of the treatable problems.

Among the quantum annealing protocols, a particularly relevant one is the crossing of a quantum phase transition, an occurrence which typically emerges for any hard problems encoded in the annealer. If the control parameter is slowly changed on a timescale much larger than the typical inverse zero-temperature gap, the

---

<sup>2</sup>More precisely, Barahona (1982) has formally shown that finding the ground state of disordered spin-glass models of the Ising type in arbitrary geometry (similar to the classical version of the Ising model discussed in Sec. 11.3) belongs to the class of NP-complete classical problems. Unfortunately, it is presently not known whether there exists an efficient quantum algorithm, belonging to the BQP class, which would be able to solve such NP-complete problems in an efficient way (see Sec. 1.3.2 for a brief discussion on the complexity classes).

system stays in its instantaneous ground state, unless a critical point is crossed. In the latter case, the system is unable to follow the driving and to remain in its equilibrium/ground state: a finite density of defects will be produced, thus invalidating the outcome of the simulation. The problem of defect formation in the adiabatic dynamics of critical systems was examined well before quantum information, by Kibble (1976) and Zurek (1985), in the context of phase transitions in the early universe, and later extended to the quantum case, giving rise to an intense discussion (see, for example, Dziarmaga (2010)). The Kibble-Zurek mechanism roughly divides the dynamics in either adiabatic or impulsive, according to the distance from the critical point. The time at which the system switches from one regime to another depends on the annealing speed: the slower it is, the later the evolution will become impulsive. This allows to predict the scaling of such defects as a function of the rate of the annealing procedure, in terms of the critical exponents of the crossed phase transition.

The picture outlined above sheds light on the intrinsic limitations of the quantum annealing approach: the solution to “hard” problems, encoded into the ground state of  $H_F$ , cannot be adiabatically connected, through a time-dependent protocol of the type in Eq. (3.182), with the ground state  $|\psi_0\rangle$  an “easy” Hamiltonian  $H_I$ . More precisely, for NP problems the connection  $H(t)$  requires the passage through a critical point, where the gap closes exponentially with the system size, thus requiring an exponentially large amount of time for the annealing process to reach the target state  $|\psi_F\rangle$ . Moreover, as explained before, the presence of the coupling with the external environment for this kind of cavity-QED devices inevitably leads to decoherence. This represents a highly detrimental effect for the annealer. Nonetheless some preliminary studies, as the ones of Johnson *et al.* (2011), Boixo *et al.* (2013), and Lanting *et al.* (2014), have shown that genuine quantum effects, including the presence of entanglement, can survive in D-Wave machines already with few tens of qubits. Evidence of the power of quantum annealing was spotlighted even with more than one hundred coupled qubits (Boixo *et al.*, 2014). However, whether these machines already hold the so-called “quantum advantage” or not (that is, a clear exponential speedup over classical computers), is still under fervid debate (see, for example, Rønnow *et al.*, 2014, and Boixo *et al.*, 2018).

## 11.2 Emergence of quantum correlations

We now characterize more in detail some prototypical many-body systems on a lattice, whose physics can be naturally addressed by means of the quantum simulators described in the previous section. Specifically, here we would like to provide a flavour of how genuine quantum frustration effects may emerge, as a result of two or more competing mechanisms.

### 11.2.1 The Hubbard model

The Hubbard model is perhaps the simplest representative model of interacting quantum particles on a lattice, describing localized electronic orbitals, with on-site repulsive interactions between particles on the same site (Hubbard, 1963). To a good approximation it mimics what occurs in a crystal at low temperatures and in the Born-Oppenheimer approximation: few valence electrons per atom move in a periodic potential, in the lowest Bloch band, ignoring any long-range interactions between the electrons. The motion of electrons can be separated from that of nuclei, whose positions are supposed to be fixed in space, thus defining the underlying lattice structure. The Hamiltonian can be written as:

$$H = -t \sum_{\langle i,j \rangle, \sigma} (c_{i,\sigma}^\dagger c_{j,\sigma} + \text{H.c.}) + U \sum_j n_{j,\uparrow} n_{j,\downarrow}, \quad (11.8)$$

where  $c_{i,\sigma}^{(\dagger)}$  denote annihilation (creation) operators for electrons with a given spin  $\sigma = \uparrow, \downarrow$ , on the  $i$ -th site of the lattice. These satisfy the anti-commutation relations  $\{c_{i,\sigma}, c_{j,\tau}^\dagger\} = \delta_{ij} \delta_{\sigma\tau}$  and  $\{c_{i,\sigma}, c_{j,\tau}\} = \{c_{i,\sigma}^\dagger, c_{j,\tau}^\dagger\} = 0$ , where  $\delta_{ij}$  is the Kronecker symbol.<sup>3</sup> The first term in Eq. (11.8), with coupling constant  $t > 0$ , accounts for the possibility of an electron with spin  $\sigma$  to hop from the  $i$ -th site to an adjacent  $j$ -th site of the lattice. The interaction term, with strength  $U > 0$ , acts on states in which there are two particles on a given site (i.e., one particle of each spin state on the same site, since the Pauli principle prevents double occupancy of the same site with the same spin state), thus making them energetically unfavoured. The emerging scenario is similar to that discussed in Sec. (11.1.1), albeit with particles obeying a different quantum statistics.

In typical situations, the electrons are tightly bounded to the nuclei of the atoms, so that the interaction strength is much larger than the hopping amplitude:  $U \gg t$ . Moreover let us focus on the case of one electron per site, so that the number of electrons and the number of lattice sites will match. If interactions were absent, the Hubbard model would simply describe a conductor: the band is indeed half filled with one electron per site (thus there is no doping), but with two possible spin states per site. Due to translational invariance, for  $U = 0$  the model can be diagonalized straightforwardly in momentum space, with a sinusoidal dispersion relation in the momentum  $k$ . Conversely, in the presence of strong interactions the system becomes insulating, as it costs an energy  $U \gg t$  to move an electron onto an already singly occupied site.

#### Mapping to effective spin models

We now concentrate on the magnetic properties of the insulating phase, and perform a second-order perturbation theory in the hopping term of Eq. (11.8), that is, we

---

<sup>3</sup>The particles in Eq. (11.8) need not to be necessarily fermions, as in Hubbard's original work. One could also consider bosonic particles  $b_i^{(\dagger)}$  satisfying the commutation relations  $[b_i, b_j^\dagger] = \delta_{ij}$ ;  $[b_i, b_j] = [b_i^\dagger, b_j^\dagger] = 0$ . This gives rise to the Bose-Hubbard model of Eq. (11.1).

write  $H = H_0 + \delta H$ , where  $H_0$  is the interaction term ( $U$ ) and the perturbation  $\delta H$  is the hopping term ( $t$ ). We start by considering a pair of sites, hosting two electrons at half filling, and compare the energies of the triplet/singlet states:

$$|1, 1\rangle = |\uparrow, \uparrow\rangle, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle), \quad |1, -1\rangle = |\downarrow, \downarrow\rangle, \quad (11.9)$$

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle). \quad (11.10)$$

Since all these states have the same occupation on each site, and the hopping term will produce something orthogonal to each state, it is clear that at leading order:

$$\langle 1, 1|\delta H|1, 1\rangle = \langle 1, 0|\delta H|1, 0\rangle = \langle 1, -1|\delta H|1, -1\rangle = \langle 0, 0|\delta H|0, 0\rangle = 0. \quad (11.11)$$

However at second order in the hopping, the energy correction to the eigenstates  $|n\rangle$  of  $H_0$  is no longer zero, indeed:

$$\delta E_n^{(2)} = \sum_{m \neq n} \frac{\langle n|\delta H|m\rangle\langle m|\delta H|n\rangle}{E_n^{(0)} - E_m^{(0)}}, \quad (11.12)$$

where  $|m\rangle$  are intermediate virtual states which admit double occupancy on a single site, while  $E_n^{(0)}$  is the zero-th order energy of  $|n\rangle$ .

It is immediate to see that  $\delta H|1, 1\rangle = \delta H|1, -1\rangle = 0$ , since, due to Pauli principle, the hopping cannot move a spin up (down) particle to a site already occupied by a spin up (down) particle. For the other two states  $\delta H|1, 0\rangle$  and  $\delta H|0, 0\rangle$ , we need to fix a sign convention determined by the ordering of fermionic operators. We adopt the following choice: from left to right (i) operators appear in the same order as sites, and (ii) spin-up operators appear first. For example:  $|\uparrow\downarrow, \downarrow\rangle = c_{1,\uparrow}^\dagger c_{1,\downarrow}^\dagger c_{2,\downarrow}^\dagger |\Omega\rangle$ , where  $|\Omega\rangle$  is the vacuum state. Therefore we have:

$$\delta H|\uparrow, \downarrow\rangle = -t(c_{1,\uparrow}^\dagger c_{2,\uparrow} + c_{2,\uparrow}^\dagger c_{1,\uparrow} + c_{1,\downarrow}^\dagger c_{2,\downarrow} + c_{2,\downarrow}^\dagger c_{1,\downarrow})c_{1,\uparrow}^\dagger c_{2,\downarrow}^\dagger |\Omega\rangle \quad (11.13)$$

$$= -t(c_{2,\uparrow}^\dagger c_{1,\uparrow} c_{1,\uparrow}^\dagger c_{2,\downarrow} + c_{1,\downarrow}^\dagger c_{2,\downarrow} c_{1,\uparrow}^\dagger c_{2,\downarrow}^\dagger) |\Omega\rangle = -t(c_{2,\uparrow}^\dagger c_{2,\downarrow} + c_{1,\uparrow}^\dagger c_{1,\downarrow}) |\Omega\rangle$$

$$\delta H|\downarrow, \uparrow\rangle = -t(c_{1,\uparrow}^\dagger c_{2,\uparrow} + c_{2,\uparrow}^\dagger c_{1,\uparrow} + c_{1,\downarrow}^\dagger c_{2,\downarrow} + c_{2,\downarrow}^\dagger c_{1,\downarrow})c_{1,\downarrow}^\dagger c_{2,\uparrow}^\dagger |\Omega\rangle \quad (11.14)$$

$$= -t(c_{1,\uparrow}^\dagger c_{2,\uparrow} c_{1,\downarrow}^\dagger c_{2,\uparrow} + c_{2,\downarrow}^\dagger c_{1,\downarrow} c_{1,\uparrow}^\dagger c_{2,\uparrow}^\dagger) |\Omega\rangle = t(c_{1,\uparrow}^\dagger c_{1,\downarrow}^\dagger + c_{2,\uparrow}^\dagger c_{2,\downarrow}^\dagger) |\Omega\rangle,$$

where we used the fermionic ordering introduced above, and (in the last equalities of the two above expressions) the anti-commutation relations. It is thus clear that

$$\delta H|1, 0\rangle = 0, \quad \delta H|0, 0\rangle = -\frac{2t}{\sqrt{2}}(|\uparrow\downarrow, 0\rangle + |0, \uparrow\downarrow\rangle), \quad (11.15)$$

so that the singlet state is energetically favoured, with

$$\delta E_{0,0}^{(2)} = \frac{2t \times 2t}{0 - U} = -\frac{4t^2}{U}. \quad (11.16)$$

We have thus discovered that, at second order in perturbation theory for  $U \gg t$ , there is an antiferromagnetic interaction of strength  $J = 4t^2/U$ , which favours the projection onto the singlet state. Coming back to the pair of spin-1/2 particles, we can consider the operator

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2} \left[ (\mathbf{S}_1 + \mathbf{S}_2)^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2 \right] = \begin{cases} \frac{1}{2}(0 - 2 \times \frac{3}{4}) = -\frac{3}{4} & \text{for a singlet,} \\ \frac{1}{2}(2 - 2 \times \frac{3}{4}) = +\frac{1}{4} & \text{for a triplet,} \end{cases} \quad (11.17)$$

where  $\mathbf{S}_j = \boldsymbol{\sigma}_j/2$  and  $\boldsymbol{\sigma}_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z)$  are the usual Pauli matrices on the  $j$ -site. In this way, the effective hopping term can be written as an isotropic magnetic interaction of the form  $J(\mathbf{S}_1 \cdot \mathbf{S}_2)$ , with  $J = 4t^2/U$ . In the many-body scenario of several pairs of spin-1/2 particles, this leads to the following effective Hamiltonian:

$$H_{\text{eff}} \sim -J \sum_{\langle i,j \rangle} \mathcal{P}_{\text{singlet}}(i,j) = \frac{4t^2}{U} \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad (11.18)$$

where, according to Eq. (11.17),  $\mathcal{P}_{\text{singlet}}(i,j) = \frac{1}{4} - \mathbf{S}_i \cdot \mathbf{S}_j$  is the projector on the singlet state. This is named the antiferromagnetic spin-1/2 *Heisenberg model*.

At this point it is crucial to observe that, when considering a lattice of more than two sites (with a spin-1/2 particle per site),  $H_{\text{eff}}$  would prefer to place all the neighbouring spins in a singlet state. This is however not possible, and a problem of *quantum frustration* arises. Indeed, since the singlet state  $|0,0\rangle$  of Eq. (11.10) forms a maximally entangled Bell pair, the discussion in Sec. 6.8.1 tells us that, for any two spins  $A$  and  $B$  that are placed in such configuration, the monogamy of entanglement forbids to create any quantum state where the pair would be entangled with any other spin  $C$ . Specifically if  $A$  and  $B$  are in a singlet state, the concurrence  $C_{AB} = 1$ ; then Eq. (6.103) implies that  $C_{AC} = 0$ , and thus  $A$  cannot be entangled with  $C$ . This simple argument exemplifies the possibility to establish a highly non-trivial pattern of quantum correlations in complex many-body systems, clearly hinting at the fact that perturbative methods around a single state (as a pure state or Bloch waves) typically do not work in the presence of strong interactions.

In the remainder of this chapter, we will first discuss the prototypical example of an exactly solvable quantum many-body system. Later we will address a wider scenario, where general considerations on the scaling of bipartite entanglement for the low-lying energy states of a class of Hamiltonian models enable us to develop a theory that well captures the low-energy physics of typical strongly correlated quantum systems.

### 11.3 The spin-1/2 quantum Ising chain

In this section we present the one-dimensional (1D) quantum Ising chain in a transverse magnetic field, for spin-1/2 particles. This covers a pivotal importance for the understanding of the statistical mechanics of 1D quantum many-body systems. The possibility to find an easy analytical solution, firstly devised by Lieb *et al.* (1961) and by Pfeuty (1970), makes it the ideal playground where several emerging pivotal aspects of strongly correlated systems can be discussed. The Hamiltonian is

$$H = -J \sum_j \sigma_j^x \sigma_{j+1}^x - h \sum_j \sigma_j^z, \quad (11.19)$$

where the spin-1/2 Pauli matrices  $\sigma_j^\alpha$  ( $\alpha = x, y, z$ ) are defined on each site  $j$  of the chain ( $j = 1, \dots, L$  with  $L$  being the chain length),  $J$  denotes the nearest-neighbour coupling and  $h$  the magnetic field strength. For the sake of clarity and without loss of generality, we assume an antiferromagnetic coupling  $J > 0$ .

We are going to show how it is possible to fully diagonalize Eq. (11.19) using a Jordan–Wigner transformation into a quadratic fermionic Hamiltonian, followed by a Bogoliubov rotation in momentum space. We first concentrate on basic properties of the ground state, like its energy, magnetization and spin-spin correlation functions. These results, as well as a series of additional analytical findings on the whole excitation spectrum in out-of-equilibrium conditions which will not be touched in this presentation, can be found in a series of seminal papers by Barouch, McCoy and coworkers (Barouch *et al.*, 1970; Barouch and McCoy, 1971a,b; McCoy *et al.*, 1971). Later we will discuss more recent results where ground-state correlations have been reinterpreted in a quantum-information context, unveiling a peculiar behaviour for specific entanglement measures. Finally we will describe how the Ising chain can be read in the context of Majorana physics.

Let us first transform the spin-1/2 particles into hard-core bosons  $a_j$ , by identifying  $|\downarrow\downarrow\rangle \leftrightarrow |0\rangle$  and  $|\uparrow\uparrow\rangle \leftrightarrow |1\rangle$  at each site. The operators  $a_j^{(\dagger)}$  commute at different sites (as the original Pauli operators do), but are not ordinary bosonic operators, because they must satisfy the hard-core constraint  $(a_j^\dagger)^2 |0\rangle = 0$ , that is, at most one boson per site is allowed. Using the definition of the raising and lowering spin operators  $\sigma_\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y)$ , which act as  $\sigma_+ |\downarrow\rangle = |\uparrow\rangle$ ,  $\sigma_- |\uparrow\rangle = |\downarrow\rangle$ , we get:

$$\sigma_j^+ = a_j^\dagger, \quad \sigma_j^- = a_j, \quad \sigma_j^z = 2a_j^\dagger a_j - 1. \quad (11.20)$$

Notice that on the same site  $\{\sigma_j^-, \sigma_j^+\} = 1$ , thus implying that  $\{a_j, a_j^\dagger\} = 1$ , while ordinary bosons would have the commutator. In terms of hard-core bosons, the Ising Hamiltonian of Eq. (11.19) becomes:

$$H = -J \sum_j (a_j^\dagger a_{j+1} + a_j^\dagger a_{j+1}^\dagger + \text{H.c.}) - h \sum_j (2a_j^\dagger a_j - 1). \quad (11.21)$$

The next step is to switch to ordinary fermionic operators, with which it is possible to easily diagonalize the model.

### 11.3.1 Jordan–Wigner transformation

The hard-core constraint introduced above for the bosonic operators  $a_j$  seems to be ideally representable in terms of spinless fermions, where the absence of double occupancy is automatically enforced by the Pauli exclusion principle, and the anti-commutation on the same site comes for free. Unfortunately, since fermion operators on different sites anti-commute (while the bosons  $a_j$  commute), the resulting minus sign must be correctly handled by performing a non-local mapping, which is called the *Jordan–Wigner transformation* (JWT). Specifically, the JWT of hard-core  $a_j$  bosons into  $c_j$  fermions reads:

$$a_j = K_j c_j = e^{i\pi \sum_{i < j} n_i} c_j = \left[ \prod_{i=1}^{j-1} (1 - 2n_i) \right] c_j, \quad (11.22)$$

where  $n_j = c_j^\dagger c_j$  is the fermion number operator on site  $j$ , while  $K_j = e^{i\pi \sum_{i < j} n_i}$  is a string operator which simply accounts for the parity of the number of fermions

before site  $j$  (i.e., between site 1 and site  $j - 1$  of the chain), and thus multiplies the operator  $c_i$  by a phase-factor  $\pm 1$ . Notice that, in order to apply the JWT, one needs to define an ordering of the various sites of the system. This can be naturally done in one dimension, but becomes less meaningful in higher dimensions, thus essentially limiting the usefulness of the transformation (11.22) to 1D systems.

**Exercise 11.1** It is not difficult to verify that the string phase-factor  $K_j$  implies that the  $c_j$  fermions satisfy standard anti-commutation relations. Using the JWT defined in Eq. (11.22), show that indeed:

$$\{c_i, c_j^\dagger\} = \delta_{ij}, \quad \{c_i, c_j\} = \{c_i^\dagger, c_j^\dagger\} = 0. \quad (11.23)$$

To this purpose, you first need to employ the mapping of Eq. (11.20), and then use the commutation relations of the Pauli matrices. It can be useful to preliminarily prove that  $[\sigma_i^+, \sigma_j^-] = \delta_{ij} \sigma_j^z$  and  $[\sigma_i^z, \sigma_j^\pm] = \pm 2 \delta_{ij} \sigma_j^\pm$ .

We conclude with a summary of a few useful expressions where the string  $K_j$  cancels out exactly, in view of the 1D geometry of the problem:

$$\begin{aligned} a_j^\dagger a_j &= c_j^\dagger c_j, \\ a_j^\dagger a_{j+1}^\dagger &= c_j^\dagger (1 - 2n_j) c_{j+1}^\dagger = c_j^\dagger c_{j+1}^\dagger, \\ a_j^\dagger a_{j+1} &= c_j^\dagger (1 - 2n_j) c_{j+1} = c_j^\dagger c_{j+1}, \\ a_j a_{j+1} &= c_j (1 - 2n_j) c_{j+1} = c_j [1 - 2(1 - c_j c_j^\dagger)] c_{j+1} = -c_j c_{j+1}, \\ a_j a_{j+1}^\dagger &= c_j (1 - 2n_j) c_{j+1}^\dagger = c_j [1 - 2(1 - c_j c_j^\dagger)] c_{j+1}^\dagger = -c_j c_{j+1}^\dagger, \end{aligned}$$

which can be readily obtained by noting that  $K_j K_{j+1} = 1 - 2n_j$ , since  $(1 - 2n_i)(1 - 2n_i) = 1$ , and terms with different site index commute. Using such relations, it is immediate to show that both terms of the Ising model transform in a simple way into local fermionic operators:

$$\sigma_j^z = 2a_j^\dagger a_j - 1 = 2n_j - 1, \quad (11.24)$$

$$\sigma_j^x \sigma_{j+1}^x = (a_j^\dagger a_{j+1}^\dagger + a_j^\dagger a_{j+1} + \text{H.c.}) = (c_j^\dagger c_{j+1}^\dagger + c_j^\dagger c_{j+1} + \text{H.c.}). \quad (11.25)$$

We can thus write the Hamiltonian of Eq. (11.21) in a fermionic language as

$$H = -J \sum_j (c_j^\dagger c_{j+1} + c_j^\dagger c_{j+1}^\dagger + \text{H.c.}) - h \sum_j (2n_j - 1). \quad (11.26)$$

A final important remark concerns the boundary conditions. One often assumes periodic boundary conditions (PBC) for spin operators, which in turn immediately implies the same boundary conditions for the hard-core bosons, that is,  $a_L^\dagger a_{L+1} \equiv a_L^\dagger a_1$  (we identify site  $L+1$  with site 1 in the chain). However, when rewriting such term using spinless fermions, we have:

$$a_L^\dagger a_1 = e^{i\pi \sum_{i=1}^{L-1} n_i} c_L^\dagger c_1 = -e^{i\pi \sum_{i=1}^L n_i} c_L^\dagger c_1 = -(-1)^{N_F} c_L^\dagger c_1, \quad (11.27)$$

where  $N_F$  denotes the total number of  $c_j$ -fermions that are present in the lattice. The second equality follows since, to the left of  $c_L^\dagger$ , one certainly has  $n_L = 1$ , and

thus we have:  $-e^{i\pi n_L} = 1$ . Analogously we find that  $a_L^\dagger a_1^\dagger = -(-1)^{N_F} c_L^\dagger c_1^\dagger$ . This shows that boundary conditions are affected by the fermionic parity  $(-1)^{N_F}$ : for an odd number  $N_F$  of fermions one gets PBC in the fermionic model, while for  $N_F$  even, anti-periodic boundary condition are found. Notice that, although the number of fermions is not conserved by the Hamiltonian (11.26), its parity is conserved and  $(-1)^{N_F}$  is a constant of motion, with value  $\pm 1$ . Summarizing, in fermionic language, for an even number of fermions we have anti-periodic boundary conditions (that is, the  $L$ -th bond has an opposite sign with respect to the other ones:  $c_{L+1} = -c_1$ ). On the opposite, for an odd number of fermions periodic boundary conditions are required (all the bonds have the same sign:  $c_{L+1} = c_1$ ).

### 11.3.2 Diagonalization of the Ising chain

Since the Hamiltonian of Eq. (11.26) conserves the fermion parity, both the even (+) and the odd (-) sector of the fermionic Hilbert space have to be considered when diagonalizing the model. Let us denote with  $H^\pm$  the two even/odd Hamiltonian subspace restrictions, such that  $H = H^+ + H^-$ . Due to the translational invariance of the model (11.19), it is very helpful to switch to momentum space, and define a Fourier transform of the fermionic operators according to:

$$c_j = \frac{1}{\sqrt{L}} \sum_k d_k e^{i \frac{2\pi}{L} k j}, \quad j = 1, \dots, L. \quad (11.28)$$

In order to properly choose the possible values of the momenta  $k$ , we shall distinguish various cases, according to the parities of  $L$  and  $N_F$ . Recall that, according to the parity of  $N_F$ , the following fermionic boundary conditions are required:  $c_{L+1} = \pm c_1$ , where the plus sign is for  $N_F$  odd, while the minus sign is for  $N_F$  even. This is equivalent to asking that  $e^{i \frac{2\pi}{L} k L} = \pm 1$ . Without loss of generality, let us now redefine the labelling of sites using the following practical convention:

$$j = \begin{cases} -\frac{L-1}{2}, -\frac{L-3}{2}, -\frac{L-5}{2}, \dots, +\frac{L-1}{2} & (L \text{ odd}), \\ -\frac{L}{2} + 1, -\frac{L}{2} + 2, -\frac{L}{2} + 3, \dots, +\frac{L}{2} & (L \text{ even}). \end{cases} \quad (11.29)$$

Therefore with this labelling, we can use the following momenta (in units of  $2\pi/L$ ):

$$k = \begin{cases} -\frac{L-1}{2}, -\frac{L-3}{2}, -\frac{L-5}{2}, \dots, +\frac{L-1}{2} & (N_F + L \text{ even}), \\ -\frac{L}{2} + 1, -\frac{L}{2} + 2, -\frac{L}{2} + 3, \dots, +\frac{L}{2} & (N_F + L \text{ odd}). \end{cases} \quad (11.30)$$

**Exercise 11.2** Prove the consistency of our choice (11.30) with the boundary conditions for fermions.

Using the transformation (11.28), it is now possible to rewrite the two  $H^\pm$  Hamiltonian sectors in momentum space (with the appropriate choice of the

$k$ -vectors), according to:

$$\begin{aligned} H^\pm &= -\frac{J}{L} \sum_j \sum_{k,k'} \left\{ d_k^\dagger d_{k'} \left( e^{i\frac{2\pi}{L}[-kj+k'(j+1)]} + e^{i\frac{2\pi}{L}[-k(j+1)+k'j]} + \frac{2h}{J} e^{i\frac{2\pi}{L}[-kj+k'j]} \right) \right. \\ &\quad \left. + d_k^\dagger d_{k'}^\dagger e^{i\frac{2\pi}{L}[-kj-k'(j+1)]} + d_k d_{k'} e^{i\frac{2\pi}{L}[k(j+1)+k'j]} - \frac{h}{J} \delta_{k,k'} \right\} \\ &= -J \sum_k \left\{ d_k^\dagger d_k \left( e^{i\frac{2\pi}{L}k} + e^{-i\frac{2\pi}{L}k} + \frac{2h}{J} \right) + e^{i\frac{2\pi}{L}k} (d_k^\dagger d_{-k}^\dagger + d_k d_{-k}) - \frac{h}{J} \right\}, \end{aligned} \quad (11.31)$$

where we used the completeness relation  $\frac{1}{L} \sum_j e^{i\frac{2\pi}{L}kj} = \delta_{k,0}$ . Notice the coupling of  $-k$  with  $k$ , with the exceptions of momenta 0 and  $\pi$  for the cases where they appear (i.e.,  $k = 0$  and  $k = L/2$  in units of  $2\pi/L$ ), which do not have a separate  $-k$  partner. It is thus possible to group together terms with opposite momenta, such that the Hamiltonian is eventually decoupled into a sum of independent terms acting in the four-dimensional Hilbert spaces generated by  $k$  and  $-k$ :

$$H^+ = \sum_{k>0} H_k^+, \quad H^- = \sum_{k>0} H_k^- + H_0 + H_{L/2} \quad (L \text{ even}) \quad (11.32)$$

$$H^+ = \sum_{k>0} H_k^+ + H_{L/2}, \quad H^- = \sum_{k>0} H_k^- + H_0 \quad (L \text{ odd}). \quad (11.33)$$

Here we have singled out the unpaired contributions  $H_0 = -2(J+h)n_0 + h$  and  $H_{L/2} = 2(J-h)n_{L/2} + h$  (corresponding to  $k = 0$  and  $k = L/2$ ), and defined

$$H_k^\pm = -2 [J \cos(\frac{2\pi k}{L}) + h] (d_k^\dagger d_k - d_{-k} d_{-k}^\dagger) - 2iJ \sin(\frac{2\pi k}{L}) (d_k^\dagger d_{-k}^\dagger - d_{-k} d_k), \quad (11.34)$$

exploiting the anti-commutation relations for the fermions, and the fact that the cosine is an even function, while  $d_k^\dagger d_{-k}^\dagger - d_{-k} d_k$  is odd with respect to  $k \rightarrow -k$ . Now we introduce the two variables

$$f_k = -2 [J \cos(\frac{2\pi k}{L}) + h], \quad g_k = 2J \sin(\frac{2\pi k}{L}). \quad (11.35)$$

In this way, the expression for  $H_k^\pm$  can be recast in a compact form using the so-called Nambu formalism, with the fermionic two-component spinor  $\Psi_k^\dagger = [d_k^\dagger, d_{-k}]$ :

$$H_k = [d_k^\dagger \ d_{-k}] \begin{bmatrix} f_k & -ig_k \\ ig_k & -f_k \end{bmatrix} \begin{bmatrix} d_k \\ d_{-k}^\dagger \end{bmatrix} \equiv \Psi_k^\dagger \mathbb{H}_k \Psi_k, \quad (11.36)$$

where we omitted the apex  $\pm$ , and defined  $\mathbb{H}_k = f_k \sigma_z + g_k \sigma_y$ .

### Bogoliubov rotation

We will now focus on the diagonalization of the  $2 \times 2$  Hamiltonian in Eq. (11.36), since the unpaired terms  $H_0$  and  $H_{L/2}$  appearing in Eqs. (11.32)–(11.33) are already in a diagonal form. The matrix  $\mathbb{H}_k$  can be readily diagonalized by taking a rotation  $R_x(\theta_k)$  of an angle  $\theta_k = \arctan(g_k/f_k)$  around the  $x$  axis, that is, we introduce the unitary operator

$$U_k \equiv R_x(\theta_k) = \exp\left(i\frac{\theta_k}{2}\sigma^x\right) = \cos\left(\frac{\theta_k}{2}\right)\mathbb{I} + i\sin\left(\frac{\theta_k}{2}\right)\sigma^x. \quad (11.37)$$

The corresponding eigenvalues are given by:

$$\varepsilon_k^\pm = \pm \varepsilon_k, \quad \text{with } \varepsilon_k = \sqrt{f_k^2 + g_k^2} = 2J\sqrt{1 + \frac{h^2}{J^2} + \frac{2h}{J} \cos\left(\frac{2\pi k}{L}\right)}, \quad (11.38)$$

while the change-of-basis matrix  $U_k$  applied to the  $\Psi_k$  fermions defines the new operators  $\Phi_k = U_k^\dagger \Psi_k$ , which diagonalize the problem:

$$\Phi_k \equiv \begin{bmatrix} b_k \\ b_{-k}^\dagger \end{bmatrix} = \begin{bmatrix} \cos(\theta_k/2) & -i \sin(\theta_k/2) \\ -i \sin(\theta_k/2) & \cos(\theta_k/2) \end{bmatrix} \begin{bmatrix} d_k \\ d_{-k}^\dagger \end{bmatrix}. \quad (11.39)$$

If we now define  $u_k = \cos(\theta_k/2) = f_k/\varepsilon_k$  and  $v_k = \sin(\theta_k/2) = g_k/\varepsilon_k$ , such that  $u_{-k} = u_k$ ,  $v_{-k} = -v_k$ , and  $u_k^2 + v_k^2 = 1$ , Eq. (11.39) can be rewritten as:

$$\begin{cases} b_k = u_k d_k - i v_k d_{-k}^\dagger \\ b_{-k}^\dagger = -i v_k d_k + u_k d_{-k}^\dagger \end{cases}. \quad (11.40)$$

It is then immediate to verify that the  $b_k$  operators obey anti-commutation relations:

$$\{b_k, b_k^\dagger\} = \{u_k d_k - i v_k d_{-k}^\dagger, i v_k d_{-k} + u_k d_k^\dagger\} = u_k^2 \{d_k, d_k^\dagger\} + v_k^2 \{d_{-k}, d_{-k}^\dagger\} = 1, \quad (11.41)$$

and thus correspond to fermionic quasiparticles. Summarizing, we can write the Hamiltonian  $H_k$  of Eq. (11.36) in a diagonal form as

$$H_k = \Psi_k^\dagger U_k (U_k^\dagger \mathbb{H}_k U_k) U_k^\dagger \Psi_k = \Phi_k^\dagger \begin{bmatrix} \varepsilon_k & 0 \\ 0 & -\varepsilon_k \end{bmatrix} \Phi_k = \varepsilon_k (b_k^\dagger b_k + b_{-k}^\dagger b_{-k} - 1). \quad (11.42)$$

The procedure outlined here to diagonalize the fermionic Hamiltonian (11.36) in momentum space is called *Bogoliubov rotation*, while the operators  $b_k$  over which the Hamiltonian becomes diagonal are associated to the Bogoliubov *quasiparticles* of the Ising model.

**Exercise 11.3** Show that, using the same approach described above, it is possible to diagonalize a generic fermionic quadratic Hamiltonian of the type:

$$H = \sum_{i,j} c_i^\dagger A_{i,j} c_j + \frac{1}{2} \sum_{i,j} \left\{ c_i^\dagger B_{i,j} c_j^\dagger + \text{H.c.} \right\}, \quad (11.43)$$

where  $A$  needs to be a Hermitian matrix, due to the hermiticity of  $H$ , and  $B$  needs to be a skew-symmetric matrix, due to the anti-commutation rules among the  $c_j$  fermions. This construction generalizes the Bogoliubov rotation (11.39) for the  $\{k, -k\}$  momentum space, to the  $2L$  dimensional real space of the fermionic  $\{c_j\}_{j=1,\dots,L}$  and  $\{c_j^\dagger\}_{j=1,\dots,L}$  operators.

### Ground state

The expression (11.42) allows to conclude that, neglecting momenta 0 and  $\pi$  (when they appear), the ground state of the Hamiltonian must be the state  $|\psi_0\rangle$  which annihilates the  $b_k$  quasiparticles for all  $k$ , also called the *Bogoliubov vacuum*:

$$b_k |\psi_0\rangle = 0, \quad \forall k. \quad (11.44)$$

Moreover, the energy dispersion relation of Eq. (11.38) tells us that the low-energy excitations (i.e., those minimizing  $\varepsilon_k$ ) are quasiparticles with  $k \rightarrow L/2$  (corresponding to a momentum  $\pi$ , in units of  $2\pi/L$ ), which behave as:

$$\lim_{\tilde{k} \rightarrow 0} \varepsilon_{\tilde{k}} = \begin{cases} 2|h - J| + \frac{4\pi^2 J h}{|h - J| L^2} \tilde{k}^2 & \text{if } h \neq J, \\ \frac{4\pi}{L} \tilde{k} & \text{if } h = J, \end{cases} \quad (11.45)$$

with  $\tilde{k} = k - L/2$ . Therefore they exhibit a quadratic dispersion for  $h \neq J$ , and a linear dispersion for  $h = J$ . This also means that the energy gap between the ground state and the first excited state  $\Delta_E^{(0)} = 2|h - J|$  is always finite, except at  $h = J$  which is also called a gapless point. As we shall see later, the point  $h_c = J$  corresponds to a critical value for which the thermodynamic properties of the system exhibit a dramatic change.<sup>4</sup> This is a so-called *quantum critical point*.

All the conclusions drawn above hold for systems which are so large that the discretization of the momenta  $k$  makes the two uncoupled values 0 and  $L/2$ , as well as any detail on the choices of Eqs. (11.29)–(11.30), completely irrelevant. This is the case when the thermodynamic limit  $L \rightarrow \infty$  is approached. However, for finite systems, one should pay attention to the fact that, in principle, there are two competing ground states, coming from the  $N_F$ -even and from the  $N_F$ -odd parity sector. Moreover we recall that periodic boundary conditions have to be adopted with an odd number of fermions (thus resulting in the Hamiltonian  $H^-$ ), and anti-periodic boundary conditions hold with an even number of fermions (thus resulting in the Hamiltonian  $H^+$ ).

Specifically, in order to analyze the energies of the two ground states, we need to look at the specific form of the Hamiltonian, Eqs. (11.32) and (11.33), keeping in mind that, for a chain of length  $L$ , the vacuum of quasiparticles has always a parity  $(-1)^L$ . Let us also recall that the excitation energy for adding a  $k = 0$  quasiparticle is  $\varepsilon_0 = -2(J + h)$  and that for adding a  $k = L/2$  quasiparticle is  $\varepsilon_{L/2} = 2(J - h)$ . Therefore, for  $h > J$  both  $\varepsilon_0$  and  $\varepsilon_{L/2}$  are negative (and thus the creation of unpaired quasiparticles is favoured), for  $h < -J$  both  $\varepsilon_0$  and  $\varepsilon_{L/2}$  are positive (and thus the creation of unpaired quasiparticles is unfavoured), for  $|h| < J$  one of the two unpaired quasiparticles has positive energy, the other has negative energy. In conclusion we have various situations.

- For  $h < -J$  only the vacuum state with an even number  $N_F$  of fermions is permitted. The other vacuum does not match the required boundary conditions, therefore one particle with a positive and finite energy has to be added, thus lifting it to an excited state.
- For  $h > J$  again only one vacuum state is permitted: this corresponds to  $N_F$  even if  $L$  is even, and to  $N_F$  odd if  $L$  is odd.
- For  $|h| < J$  both two vacua, with  $N_F$  even and with  $N_F$  odd are permitted. This ends up into a quasi-degeneracy in the ground state, that becomes exact in the thermodynamic limit. We will come back to this point later in Sec. 11.3.5.

---

<sup>4</sup>An analogous behaviour occurs at  $h = -J$ . It is however important to stress that, for  $h < 0$ , the spectrum is reversed and thus one should look at excitations close to  $k = 0$ .

### Duality mapping

The peculiarity of the  $h = \pm J$  points is also signaled by the fact that it is possible to perform a canonical transformation of the spins  $\sigma_j^\alpha \rightarrow \tau_j^\alpha$  that preserves the structure of the Hamiltonian (11.19), and inverts the roles of the coupling constants  $J$  and  $h$ . Namely, this is defined by

$$\tau_j^x = \prod_{k < j} \sigma_k^z; \quad \tau_j^z = \sigma_j^x \sigma_{j+1}^x, \quad (11.46)$$

such that the Ising model in these new variables becomes

$$H_\tau = -J \sum_j \tau_j^z - h \sum_j \tau_j^x \tau_{j+1}^x. \quad (11.47)$$

Setting  $J = 1$  for simplicity, we find that  $H_\sigma(h) \longleftrightarrow h \times H_\tau(h^{-1})$ , where we only expressed the dependence of the model on the field, and  $H_\sigma$  denotes the original Hamiltonian of Eq. (11.19). The symbol  $\longleftrightarrow$  indicates the equivalence of two Hamiltonians, up to a canonical transformation. This self-duality property guarantees an identical spectral structure of the system in correspondence of fields with strength  $h$  and  $1/h$ , thus imposing the existence of a critical point at  $h = J$  (and a symmetric one at  $h = -J$ ), if we assume that the latter is unique.

### Correlation functions

We are now in the position to discuss how the integrability of the model can be exploited in order to extrapolate physical quantities, such as spin magnetizations or correlation functions. To this purpose, for a chain with  $L$  sites (we will take  $L$  odd) it is first convenient to introduce the following  $2L$  operators  $\gamma_j$ , with  $j = -L, -L+1, \dots, L-1$ :

$$\gamma_{2j-1} = \left( \prod_{m < j} \sigma_m^z \right) \sigma_j^x, \quad \gamma_{2j} = \left( \prod_{m < j} \sigma_m^z \right) \sigma_j^y. \quad (11.48)$$

Notice that the string  $\prod_{m < j} \sigma_m^z$  corresponds exactly to the one in front of Eq. (11.22), in the standard JWT. Now, using the fact that  $\sigma_j^x = \sigma_j^+ + \sigma_j^-$  and  $\sigma_j^y = -i(\sigma_j^+ - \sigma_j^-)$ , we obtain the following peculiar property:

$$\gamma_{2j-1} = c_j^\dagger + c_j, \quad \gamma_{2j} = -i(c_j^\dagger - c_j), \quad (11.49)$$

thus implying that the operators defined in (11.48) are Hermitian:  $\gamma_j^\dagger = \gamma_j$ . Any particle respecting this rule is called a *Majorana fermion*.

We are now going to show that any possible spin-correlation function of the Ising model can be obtained by knowing the  $2L \times 2L$  correlation matrix  $\Gamma_L$  of the above Majorana operators, whose matrix elements are defined by:

$$[\Gamma_L]_{m,n} = \langle \gamma_m \gamma_n \rangle, \quad m, n = -L, \dots, L-1. \quad (11.50)$$

Hereafter, unless specified, the averages of any observable  $\mathcal{O}$  have to be intended on the ground state:  $\langle \mathcal{O} \rangle = \langle \psi_0 | \mathcal{O} | \psi_0 \rangle$ . The ground-state expectation values  $\langle \gamma_m \gamma_n \rangle$ , entering the various matrix elements of  $\Gamma_L$ , are computed by applying a sequence of

recursive transformations which map  $\gamma_j$ -Majorana operators into  $c_j$ -fermions, then into  $d_k$ -fermions, and finally into the Bogoliubov  $b_k$ -quasiparticles that diagonalize the model. Putting them all together, we get

$$\begin{bmatrix} \gamma_{2j-1} \\ \gamma_{2j} \end{bmatrix} = \frac{1}{\sqrt{L}} \sum_k e^{i\frac{2\pi}{L}kj} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} u_k & iv_k \\ iv_k & u_k \end{bmatrix} \begin{bmatrix} b_k \\ b_{-k}^\dagger \end{bmatrix}. \quad (11.51)$$

As a matter of fact, from Eq. (11.44) we know that the ground state is defined as the one that is annihilated by all such  $b_k$  particles. Therefore we have:<sup>5</sup>

$$\langle b_k b_{k'}^\dagger \rangle = \delta_{kk'}; \quad \langle b_k b_{k'} \rangle = \langle b_k^\dagger b_{k'}^\dagger \rangle = \langle b_k^\dagger b_{k'} \rangle = 0. \quad (11.52)$$

Combining this with the relations (11.51), one can finally arrive to the following:

$$[\Gamma_L]_{mn} = \delta_{mn} + i\Lambda_{mn}, \quad \text{with } \Lambda_{mn} = \begin{bmatrix} \Pi_0 & \Pi_1 & \cdots & \Pi_{L-1} \\ -\Pi_1 & \Pi_0 & \cdots & \Pi_{L-2} \\ \vdots & \vdots & & \vdots \\ -\Pi_{L-1} & -\Pi_{L-2} & \cdots & \Pi_0 \end{bmatrix}, \quad (11.53)$$

where  $\Lambda$  is a skew-symmetric matrix with sub-blocks

$$\Pi_j = \begin{bmatrix} 0 & g(j) \\ -g(-j) & 0 \end{bmatrix}, \quad g(j) = \frac{1}{L} \sum_k e^{i\frac{2\pi}{L}kj} e^{i\theta_k/2}, \quad (11.54)$$

and  $\theta_k$  is the Bogoliubov angle defined above. In the thermodynamic limit we can write explicitly:

$$g(j) = \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{i\phi j} \frac{J \cos \phi + h - iJ \sin \phi}{|J \cos \phi + h - iJ \sin \phi|}. \quad (11.55)$$

Now suppose that we want to calculate a two-point correlation function of the form  $\langle \sigma_j^\alpha \sigma_l^\beta \rangle$ . First we have to express the Pauli matrices in terms of the Majorana fermions  $\gamma_j$ , using their definition (11.48):

$$\sigma_j^x = \left( \prod_{m < j} \sigma_m^z \right) \gamma_{2j-1}, \quad \sigma_j^y = \left( \prod_{m < j} \sigma_m^z \right) \gamma_{2j}, \quad i\sigma_j^z = \sigma_j^x \sigma_j^y = \gamma_{2j-1} \gamma_{2j}. \quad (11.56)$$

From this it is clear that  $\sigma^z$  correlators can be calculated straightforwardly, since they do not involve JW strings. On the other hand, correlators involving for example  $\langle \sigma_j^x \sigma_l^x \rangle$  are expressed as an expectation value of a string of Majorana fermions:

$$\langle \sigma_j^x \sigma_l^x \rangle = \langle \gamma_{2j} \gamma_{2j+1} \dots \gamma_{2l-1} \rangle, \quad (11.57)$$

where we used the fact that  $\gamma_{2j-1} \gamma_{2j-1} = (c_j^\dagger + c_j)(c_j^\dagger + c_j) = 1$ . These contractions can be drastically simplified by applying Wick's theorem to expectation values taken with respect to the ground (or the thermal) state of free-Fermi theory:

$$\langle \gamma_{2j} \gamma_{2j+1} \dots \gamma_{2l-1} \rangle = \sum_{\text{all pairings}} (-1)^P \prod_{\text{all pairs}} (\text{contraction of all pairs}), \quad (11.58)$$

---

<sup>5</sup> All the calculations presented here can be generalized to fermionic thermal states, observing that the relations (11.52) transform into:  $\langle b_k^\dagger b_{k'}^\dagger \rangle_\beta = \delta_{kk'}/[1 + \exp(\beta\varepsilon_k)] = \delta_{kk'} - \langle b_k b_{k'}^\dagger \rangle_\beta$ , and  $\langle b_k^\dagger b_{k'}^\dagger \rangle_\beta = \langle b_k b_{k'} \rangle_\beta = 0$ , where we used the Fermi-Dirac statistics and  $\beta = 1/(k_B T)$ .

where  $(-1)^P$  is the parity of the permutation that takes the indexes  $_{2j,2j+1,\dots,2l-1}$  into any sequence defining a string of operators to be contracted in pairs. The sum over all the pairings of the product of the pair contractions coincides with the square root of the determinant (i.e., the Pfaffian) of the  $(2l - 2j) \times (2l - 2j)$  reduced matrix  $\Gamma_{l-j}^{(R)}$ , obtained as a diagonal sub-block of the full correlation matrix  $\Gamma_L$ . This method is applicable to any spin-correlation function, which can be always taken as the expectation value of a string of Majorana fermions, as in Eq. (11.58).

A physically relevant result in this respect is the magnetization  $M^x$  along the coupling direction, obtainable as the long-distance limit of the two-point correlation:  $\lim_{n \rightarrow \infty} \langle \sigma_j^x \sigma_{j+n}^x \rangle = (M^x)^2$ , which can be calculated to give

$$M_x = \langle \sigma_x \rangle = \begin{cases} \left(1 - \frac{h^2}{J^2}\right)^{1/8} & \text{if } |h| < J, \\ 0 & \text{if } |h| \geq J, \end{cases} \quad (11.59)$$

where we omitted the irrelevant site index  $j$ , for a translationally invariant system. This shows once more that  $h_c = \pm 1$  is a special point, separating a region where the ground state of the system is unpolarized along the  $x$  axis ( $|h| > |h_c|$ ), from another one where the  $x$ -axis magnetization acquires a finite value ( $|h| < |h_c|$ ). More precisely,  $h_c$  denotes the onset of a zero-temperature *quantum phase transition*, where, as a consequence of a variation of the field strength  $h$ , the ground state undergoes drastic modifications: it changes from an  $x$ -paramagnet, where  $\langle \sigma_x \rangle = 0$ , to an  $x$ -ferromagnet, where  $\langle \sigma_x \rangle \neq 0$ . In passing we mention that the phenomenology outlined here is closely related to the finite-temperature phase transition in the classical Ising model above two dimensions (in the sense that the critical properties of the two phase transitions are the same), provided the parameter  $h$  is substituted by the temperature  $T$  and “quantum” fluctuations are replaced by “thermal” fluctuations. The interested reader can find a detailed discussion, for example, in Sachdev (2011).

### 11.3.3 Two-spin concurrence

A quantum phase transition may also reveal in a dramatic change of the ground-state entanglement pattern. Following the works by Osterloh *et al.* (2002) and Osborne and Nielsen (2002), we validate this statement by discussing the behaviour of the pairwise entanglement between any two spins  $i$  and  $j$  in the Ising model, when the transverse field  $h$  is varied. As shown in Sec. 6.7.1, it is possible to provide an analytic expression for the entanglement of formation of two qubits as a monotone of the so-called concurrence. Here we focus precisely on the concurrence  $C_d$ , where  $d = |i - j|$  denotes the distance between any two spin-1/2 particles.

It can be proven that, for any spin-chain model exhibiting the parity symmetry  $[H, \otimes_j \sigma_j^z] = 0$ , all the components of the wave function have an even (or odd) number of flipped spins. Under such restriction, the most general reduced density matrix of two arbitrary spins assumes a so-called  $X$ -shape, in which the only non-

zero elements are on the main diagonal and on the anti-diagonal. Using the Bloch-Fano representation of Eq. (7.84), this is given by

$$\rho_{ij} = \frac{1}{4} \left[ \sum_{\alpha_1, \alpha_2 = z, I} (r_{\alpha_1 \alpha_2} \sigma_i^{\alpha_1} \otimes \sigma_j^{\alpha_2}) + \sum_{\alpha_1, \alpha_2 = x, y} (r_{\alpha_1 \alpha_2} \sigma_i^{\alpha_1} \otimes \sigma_j^{\alpha_2}) \right]. \quad (11.60)$$

If we further assume translational invariance and reflection symmetry, as is the case for the Ising model of Eq. (11.19), we obtain  $r_{zI} = r_{Iz}$  and  $r_{xy} = r_{yx}$ , respectively. In this case the concurrence is given by the following compact formula:

$$C_d = 2 \max \{0, \tilde{C}_d^I, \tilde{C}_d^{II}\}, \quad (11.61)$$

where, recalling that  $r_{\alpha\beta} = \langle \sigma_i^\alpha \sigma_j^\beta \rangle$  and  $d = |i - j|$ , we defined

$$\tilde{C}_d^I = |r_{xx} + r_{yy}| - \sqrt{\left(\frac{1}{4} + r_{zz}\right)^2 - r_{zI}^2}, \quad \tilde{C}_d^{II} = |r_{xx} - r_{yy}| + r_{zz} - \frac{1}{4}. \quad (11.62)$$

After evaluating all the required correlation functions  $r_{\alpha\beta}$  with the machinery described in the previous section, it turns out that the two-spin concurrence for the Ising chain is always zero, unless the two sites are at most next-nearest-neighbours. Here we only focus on the nearest neighbour concurrence  $C_1$ . As is visible from the right inset of Fig. 11.5, this is a smooth function of the field, with a maximum located close to the critical point, but not exactly at  $h_c$  (note that, in the figure, the field is parametrized as  $h = 1/\lambda$ , with  $J = 1$ ). Moreover the concurrence clearly tends to zero in the limits  $h \ll 1$  and  $h \gg 1$ , since the ground state is fully polarized along the  $x$  or the  $z$  axis. The critical properties of the ground state are captured by the derivatives of  $C_1$  as a function of  $\lambda$ , as shown in the mail frame. In the thermodynamic limit, we have that

$$\partial_\lambda C_1 \sim \log |\lambda - \lambda_c|, \quad (11.63)$$

where  $\lambda_c = 1$  denotes the position of the critical point (in units of  $J$ ). For a finite system size  $L$ , the precursors of the critical behaviour can be analyzed by means of finite-size scaling, and the position of the minimum  $\lambda_m$  changes with  $L$  (left inset).

### 11.3.4 Entanglement block entropy

We now discuss the entanglement entropy of a block of  $\ell$  contiguous spins, for the ground state of the Ising model. Using again the Bloch-Fano representation, these are described by the reduced density matrix

$$\rho_\ell = \frac{1}{2^\ell} \sum_{\vec{\alpha} = \{x, y, z, I\}^\ell} r_{\alpha_1 \dots \alpha_\ell} \sigma_1^{\alpha_1} \otimes \dots \otimes \sigma_\ell^{\alpha_\ell}, \quad \text{with } r_{\alpha_1 \dots \alpha_\ell} = \langle \sigma_1^{\alpha_1} \otimes \dots \otimes \sigma_\ell^{\alpha_\ell} \rangle. \quad (11.64)$$

So in principle one would have to compute all the possible  $4^\ell$  spin correlation functions, using again the machinery described above. As shown by Vidal *et al.* (2003), it is however more practical to proceed directly with the reduced correlation matrix  $\Lambda_\ell^{(R)}$  of the first  $\ell$  spins, obtained by truncating the full matrix  $\Lambda$  of Eq. (11.53). Since the latter is skew-symmetric, it can be transformed into a block diagonal form:

$$\Lambda_\ell^F = V \Lambda_\ell^{(R)} V^T, \quad \text{with } \Lambda_\ell^F = \bigoplus_{i=1}^{\ell} \begin{bmatrix} 0 & \nu_\ell \\ -\nu_\ell & 0 \end{bmatrix}. \quad (11.65)$$

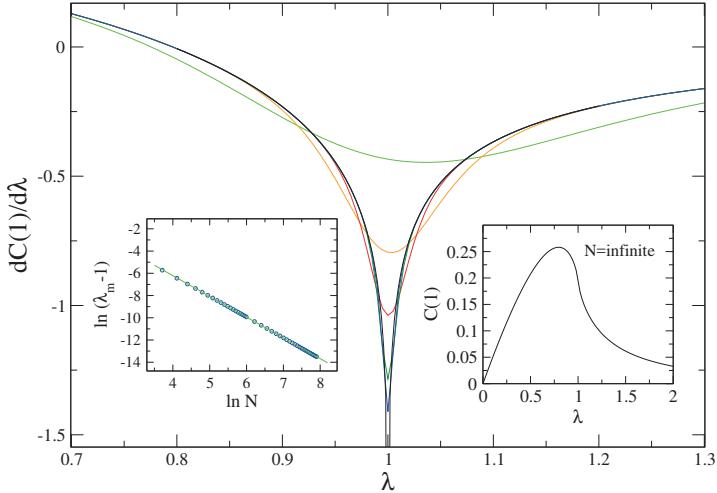


Fig. 11.5 Behaviour of the nearest-neighbour concurrence as a function of  $\lambda = J/h$  in the quantum Ising chain (here  $J = 1$ ). The main frame displays the first derivative in  $\lambda$ , and clearly shows a minimum that gets more pronounced as long as the size is increased (here referred to as  $N$ ), whose position tends to the critical value (left inset). The right inset shows that the maximum itself is not related to the critical properties of the model. The figure is reprinted with permission from Osterloh *et al.* (2002). ©(2002) Macmillan Magazines Ltd.

The transformation matrix  $V$  defines a new set of Majorana operators

$$\tilde{\gamma}_m = \sum_{n=1}^{2\ell} V_{mn} \gamma_n, \quad \text{such that} \quad \langle \tilde{\gamma}_m \tilde{\gamma}_n \rangle = \delta_{mn} + i[\Lambda_\ell^F]_{m,n}. \quad (11.66)$$

The structure of  $\Lambda_\ell^F$  implies that  $\tilde{\gamma}_{2j-1}$  is only correlated to  $\tilde{\gamma}_{2j}$ . For the sake of clarity, let us finally introduce the  $L$  spinless fermionic operators

$$\psi_j = \frac{1}{2}(\tilde{\gamma}_{2j-1} + i\tilde{\gamma}_{2j}), \quad (11.67)$$

satisfying canonical anti-commutation relations:  $\{\psi_m, \psi_n\} = 0$ ,  $\{\psi_m^\dagger, \psi_n\} = \delta_{mn}$ . By construction we have that

$$\begin{aligned} \langle \psi_m^\dagger \psi_n \rangle &= \frac{1}{4} \langle (\tilde{\gamma}_{2m-1} - i\tilde{\gamma}_{2m})(\tilde{\gamma}_{2n-1} + i\tilde{\gamma}_{2n}) \rangle \\ &= \frac{1}{4} \left( \langle \tilde{\gamma}_{2m-1} \tilde{\gamma}_{2n-1} \rangle - i \langle \tilde{\gamma}_{2m} \tilde{\gamma}_{2n-1} \rangle + i \langle \tilde{\gamma}_{2m-1} \tilde{\gamma}_{2n} \rangle + \langle \tilde{\gamma}_{2m} \tilde{\gamma}_{2n} \rangle \right) \\ &= \frac{1}{4} [\delta_{mn} - i\delta_{mn}(-i\nu_m) + i\delta_{mn}(i\nu_m) + \delta_{mn}] = \frac{1}{2} \delta_{mn} (1 - \nu_m) \end{aligned} \quad (11.68)$$

$$\begin{aligned} \langle \psi_m \psi_n \rangle &= \frac{1}{4} \langle (\tilde{\gamma}_{2m-1} + i\tilde{\gamma}_{2m})(\tilde{\gamma}_{2n-1} + i\tilde{\gamma}_{2n}) \rangle \\ &= \frac{1}{4} \left( \langle \tilde{\gamma}_{2m-1} \tilde{\gamma}_{2n-1} \rangle + i \langle \tilde{\gamma}_{2m} \tilde{\gamma}_{2n-1} \rangle + i \langle \tilde{\gamma}_{2m-1} \tilde{\gamma}_{2n} \rangle - \langle \tilde{\gamma}_{2m} \tilde{\gamma}_{2n} \rangle \right) \\ &= \frac{1}{4} [\delta_{mn} + i\delta_{mn}(-i\nu_m) + i\delta_{mn}(i\nu_m) - \delta_{mn}] = 0. \end{aligned} \quad (11.69)$$

Therefore the  $\ell$  modes  $\{\psi_j^{(\dagger)}\}_{j=1,\dots,\ell}$  are uncorrelated, and the reduced density matrix in Eq. (11.64) can be written as the tensor product of these modes:

$$\rho_\ell = \Upsilon_1 \otimes \Upsilon_2 \otimes \cdots \otimes \Upsilon_\ell, \quad (11.70)$$

where each density matrix  $\Upsilon_k$  is immediately cast in a diagonal form:

$$\Upsilon_k = \begin{bmatrix} \langle \psi_k \psi_k^\dagger \rangle & \langle \psi_k \psi_k \rangle \\ \langle \psi_k^\dagger \psi_k^\dagger \rangle & \langle \psi_k^\dagger \psi_k \rangle \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + \nu_k & 0 \\ 0 & 1 - \nu_k \end{bmatrix}. \quad (11.71)$$

Finally, using the decomposition (11.70), we can simply calculate the von Neumann entropy as the sum of the reduced entropies of the various independent modes:

$$S(\rho_\ell) = S(\Upsilon_1) + S(\Upsilon_2) + \dots + S(\Upsilon_\ell) = \sum_{j=1}^{\ell} H_{\text{bin}}\left(\frac{1 + \nu_j}{2}\right), \quad (11.72)$$

where  $H_{\text{bin}}(\cdot)$  is the Shannon binary entropy as defined in Eq. (6.26).

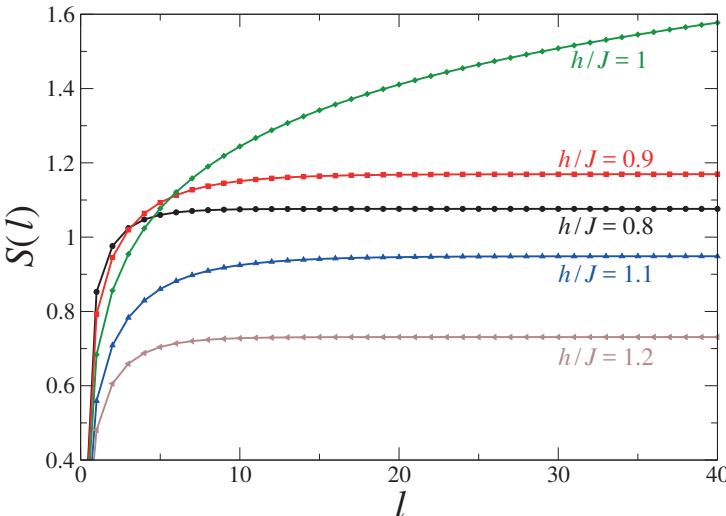


Fig. 11.6 Entanglement entropy of the reduced density matrix of  $\ell$  spins in the spin-1/2 Ising chain. The various curves are for different values of the external magnetic field  $h/J$ . While at the critical point ( $h/J = 1$ ) the entropy keeps increasing logarithmically with  $\ell$ , for all the other values of the field we can distinguish a clear saturating behaviour.

Using such formula to explicitly calculate the entropy  $S(\rho_\ell)$ , an intriguing picture emerges: a clear distinction between non-critical and critical behaviour hints at two forms of structurally inequivalent ground states, which will be thoroughly discussed in Sec. 11.4. As a matter of fact, the net result of this analysis, highlighted in Fig 11.6, shows that for  $h \neq h_c$  the entanglement entropy saturates with the size, while for  $h = h_c$  it keeps increasing logarithmically with  $\ell$ .

### 11.3.5 The Ising model revisited: Kitaev chain

We conclude our discussion on the quantum Ising chain by mentioning that this covers a fundamental importance in condensed matter, also because it represents the simplest toy model in which Majorana fermions appear, as originally devised

by Kitaev (2001). We have already introduced Majorana operators  $\gamma_j$  in Eq. (11.49), when showing that spin-spin correlation functions can be measured as the Pfaffian of a suitable correlation matrix of the  $\gamma_j$ 's. Quite curiously these are Hermitian particles:  $\gamma_j^\dagger = \gamma_j$  (in the high-energy language, we say that they are their own anti-particles). Moreover, using the anti-commutation relations for the  $c_j$ -fermions, it is easily verified that Majorana operators satisfy the anti-commutation relation

$$\{\gamma_i, \gamma_j\} = 2\delta_{ij}. \quad (11.73)$$

This implies that  $\gamma_i^2 = 1$ , and thus, acting twice with a Majorana operator, one comes back to the same initial state. Therefore there is no Pauli principle for them, and it is not even possible to speak about the occupancy of a Majorana mode. Indeed, by constructing a Majorana-like number operator, one would get  $n_j^{\text{MF}} \equiv \gamma_j^\dagger \gamma_j = \gamma_j \gamma_j = 1$ . Thus, a Majorana mode is in a sense always empty and always filled, and counting does not make any sense. As we shall see in a moment, their significance appears only in terms of normal fermions, from the interpretation that any fermion can be seen as a superposition of two Majorana operators, corresponding to its real and its imaginary part.

Let us come back to the formulation of the Ising chain in the fermionic language and consider open boundary conditions (OBC), that is, we remove the term for  $j = L$  in the first summation of Eq. (11.26). We shall now rewrite it in terms of the Majorana particles of Eq. (11.49). For the sake of simplicity in our notation, here we will change the site labelling by defining a double index according to:  $\gamma_{j,1} \equiv \gamma_{2j-1}$  and  $\gamma_{j,2} \equiv \gamma_{2j}$ . As mentioned above, Majoranas are readily obtained by splitting a  $c_j$ -fermion into its real and imaginary parts:

$$c_j^\dagger = \frac{1}{2}(\gamma_{j,1} + i\gamma_{j,2}), \quad c_j = \frac{1}{2}(\gamma_{j,1} - i\gamma_{j,2}), \quad (11.74)$$

where  $(\gamma_{j,1}, \gamma_{j,2})$  have to be intended as operators living on the  $j$ -th site of the chain (see Fig. 11.7). The physics that we are going to unveil here can be best understood in the simplest case where the external magnetic field is zero. Indeed, plugging the definitions (11.74) in Eq. (11.26) with  $h = 0$ , we arrive at the following simple expression for the fermionic Hamiltonian:

$$H = iJ \sum_{j=1}^{L-1} \gamma_{j,2} \gamma_{j+1,1}. \quad (11.75)$$

Notice that  $H$  is still Hermitian since Majorana operators are Hermitian but they anti-commute, according to Eq. (11.73). The expression (11.75) is nothing but an alternative way of writing the diagonalized Hamiltonian. To see this, we go back to a fermionic representation by defining new fermion operators with a similar construction as in Eq. (11.74), i.e., we combine Majorana operators from adjacent sites as shown in Fig. 11.7:

$$\tilde{c}_j^\dagger = \frac{1}{2}(\gamma_{j+1,1} + i\gamma_{j,2}). \quad (11.76)$$

In terms of these new fermions, we have  $i\gamma_{j,2}\gamma_{j+1,1} + 1 = 2\tilde{c}_j^\dagger\tilde{c}_j = 2\tilde{n}_j$ , and thus the Hamiltonian is automatically diagonal on this basis:

$$H = 2J \sum_{j=1}^{L-1} (\tilde{c}_j^\dagger\tilde{c}_j - \frac{1}{2}). \quad (11.77)$$

This shows that Majorana operators are nothing but a formal way of writing the fermionic Ising Hamiltonian; the physical excitations are fermionic states at finite energy, obtained by a superposition of nearest neighbour Majoranas.

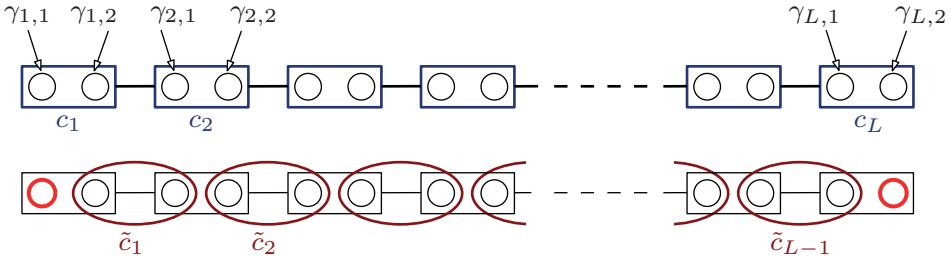


Fig. 11.7 Sketch of the Kitaev chain. Upper panel: the fermionic operators  $c_j$  on each site  $j$  of the chain can be split into two Majorana operators,  $\gamma_{j,1}$  and  $\gamma_{j,2}$ . Lower panel: for  $h = 0$ , the Hamiltonian is diagonalizable in the basis of the  $\tilde{c}_j$  fermions, which can be obtained by combining Majorana operators  $\gamma_{j,2}$  and  $\gamma_{j+1,1}$  on neighbouring sites. The two Majorana operators at the edges of the chain ( $\gamma_{1,1}$  and  $\gamma_{L,2}$ , red circles) are left unpaired, and can be combined to form a non-local zero energy fermionic mode  $\tilde{c}_M$ .

Now it turns out that, remarkably, the Majorana operators  $\gamma_{L,2}$  and  $\gamma_{1,1}$ , which are localized at the two ends of the wire, are completely missing from Eq. (11.75). Those operators can equivalently be described by a single fermionic state:

$$\tilde{c}_M^\dagger = \frac{1}{2}(\gamma_{L,2} + i\gamma_{1,1}), \quad (11.78)$$

which is highly non-local, since  $\gamma_{L,2}$  and  $\gamma_{1,1}$  are localized on opposite ends of the chain. Since such fermion is absent from the Hamiltonian, its state requires zero energy. The presence of this so-called *zero-energy Majorana mode* is intimately connected with the two-fold degeneracy of the Ising Hamiltonian in the ferromagnetic region  $|h| < J$ , corresponding to having in total an even or odd number  $N_F$  of fermions. The parity  $(-1)^{N_F}$  is associated to the eigenvalue of the number operator of the zero-energy fermion:  $n_M = \tilde{c}_M^\dagger\tilde{c}_M = 0/1$ , for even/odd parity.

The above argument has been made for  $h = 0$ , but a qualitatively analogous picture holds as long as  $|h| < h_c$ , with the emergence of zero-energy Majorana modes at the edges of the chain. In the general case  $h \neq 0$ , the Majorana fermions however are not completely localized at the two edge sites, but decay exponentially away from the edges. They remain at zero energy only if the chain is long enough that they do not overlap, thus reflecting in the quasi-degeneracy of the Ising ferromagnetic ground state that becomes exact only in the thermodynamic limit.

The importance of the framework described here resides in the possibility to create, without paying any cost in energy, a fermionic state that is made of a superposition of two Majorana particles, which are spatially separated and localized at the edges of the system. Such state can be manipulated, in principle, by physically exchanging the Majoranas, which obey a non-Abelian (non-commutative) statistics. Indeed, in virtue of the two-fold degeneracy of the ground state, which is separated from the excited states by a finite energy gap, there are certain adiabatic operations, such as the slow exchange of Majorana positions, that may bring the system from one ground state to another. Moreover, in virtue of the high degree of delocalization of the zero-energy fermionic state (part is localized on one edge, part on the other edge of the chain), this turns out to be protected from most types of decoherence, since it cannot be modified by any local perturbation, which would only affect one of its Majorana constituents. This principle has lead to the formulation of *topological quantum computation*, where the information is stored in suitably delocalized fermionic states that are robust against most sources of decoherence, which do not couple simultaneously to more than one Majorana mode.

#### 11.4 Area-law scaling of the entanglement

Let us now come back to the entanglement properties of the Ising model: a remarkable observation that we spotlighted in Sec. 11.3.4 is the different scaling of the reduced von Neumann entropy for a system lying at the critical point  $h_c$ , with respect to any other non-critical point (see Fig. 11.6 for a visual image). Quite remarkably, this behaviour is not specific to the Ising model, but entails a fundamental structural property of the ground-state many-body wave function that brings together a wide class of quantum systems. The essence of such property relies on general considerations based on the entanglement theory.

To fix the ideas, we consider a generic quantum system living on a lattice, such that each site  $i$  is associated to a complex Hilbert space  $\mathcal{H}_i = \mathbb{C}^d$  of a given finite dimension  $d$ . Considering a system made up of  $L$  sites, the global Hilbert space  $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d$  will be exponentially large in  $L$ , having dimension  $d^L$ . This eventually represents the fundamental obstacle in the description of quantum many-body systems that are sufficiently large, in the thermodynamic sense. Indeed, a generic wave function  $|\psi\rangle \in \mathcal{H}$  on that space can be parametrized as

$$|\psi\rangle = \sum_{i_1, \dots, i_L} c_{i_1, \dots, i_L} |i_1, \dots, i_L\rangle, \quad i_j = 1, \dots, d, \quad (11.79)$$

where  $c_{i_1, \dots, i_L}$  are  $d^L$  complex parameters containing the amplitudes on the basis elements  $|i_1, i_2, \dots, i_L\rangle = |i_1\rangle_1 \otimes |i_2\rangle_2 \otimes \cdots \otimes |i_L\rangle_L$  of the Hilbert space. Unfortunately, the number of such parameters grows so fast that it becomes unpractical to keep track of the full details of a generic  $|\psi\rangle$ . For example, considering a system of spin-1/2 particles, already for  $L = 20$  one would require to store and manipulate  $2^{20} = 1\,048\,576$  complex coefficients, making it impossible, for a standard classical computer, even to write a generic operator in the huge space  $\mathcal{H} = \mathbb{C}^{4^L}$ .

There are however specific situations where the required computational resources are tremendously reduced. This is the case, e.g., for describing product states  $|\psi\rangle_{\text{prod}}$ . For spin-1/2 systems, these can be indeed written as:

$$|\psi\rangle_{\text{prod}} = (a_1|\downarrow\rangle_1 + b_1|\uparrow\rangle_1) \otimes (a_2|\downarrow\rangle_2 + b_2|\uparrow\rangle_2) \otimes \cdots \otimes (a_L|\downarrow\rangle_L + b_L|\uparrow\rangle_L), \quad (11.80)$$

so that, out of all possible  $2^L$  complex parameters, here we have only  $2L$ . For translationally invariant systems, these further reduce to only two parameters  $a$  and  $b$ . Approximating a generic many-body wave function (11.79) as a product state, like in Eq. (11.80), corresponds to performing a *mean-field decoupling* approximation, where correlations between different sites in the chain are completely neglected.

Now it would be highly desirable to find a way for an efficient description of strongly correlated systems, in such a way to capture the relevant physics beyond the highly simplified mean-field picture. The hope for this possibility comes from the fact that in most situations, as for the Hubbard model in Eq. (11.8) or the Ising chain in Eq. (11.19), the Hamiltonian can be written as a sum of few-body terms (even without restricting to lattice systems). Namely, we have:

$$H = \sum_{\langle\langle i,j \rangle\rangle} h_{ij}, \quad (11.81)$$

where  $i$  and  $j$  are, in general, two neighbouring sites in the lattice.<sup>6</sup> Indeed, while the Hilbert space for a lattice of  $L$  sites is  $d^L$ -dimensional, the most general Hamiltonian  $H$  of this form is specified by not more than  $O(L^2 \times d^4)$  parameters, being  $L^2$  all the possible couples  $(i,j)$  on the lattice, and  $d^4$  the total number of parameters in a  $d^2$ -dimensional Hilbert space of two sites.

The answer to the above question is positive: while a generic  $L$ -body quantum state can occupy an exponentially large Hilbert space, many physically relevant states live in a tiny “corner” of this space. The difficulty is to find an efficient and clever parametrization which captures the states in this corner of Hilbert space, while at the same time allowing for efficient simulation methods. As we will show later in detail, there is a precise formalism for constructing such kind of Ansatz. Here we formulate the underlying entanglement property on which this is based.

To this end, we consider a state  $|\psi\rangle$  on a lattice, and make a contiguous bipartition. Namely, we consider a continuous region  $A$  of either  $\ell$  sites (a segment in 1D),  $\ell \times \ell$  sites (an area in 2D), or  $\ell \times \ell \times \ell$  sites (a volume in 3D). The entanglement between such region and the rest of the system is quantified by the von Neumann entropy of the reduced density matrix  $\rho_A = \text{Tr}_{L-A} [|\psi\rangle\langle\psi|]$ :

$$S(\rho_A) = -\text{Tr}[\rho_A \log \rho_A]. \quad (11.82)$$

As shown in Sec. 6.5.1, it is not difficult to prove that, for a random state, this entanglement will be close to its maximal value  $|A| \log d$ , where  $|A|$  is the number of

---

<sup>6</sup>We stress that all the observations that we will discuss hereafter do not strictly depend on the finite range of the Hamiltonian interactions  $h_{ij}$ : it is sufficient that interactions decay at most exponentially with the distance  $|i-j|$ . The double brackets notation  $\langle\langle \cdot, \cdot \rangle\rangle$  in Eq. (11.81) indeed extends the summation to distant sites, provided that such exponential decay of interactions holds.

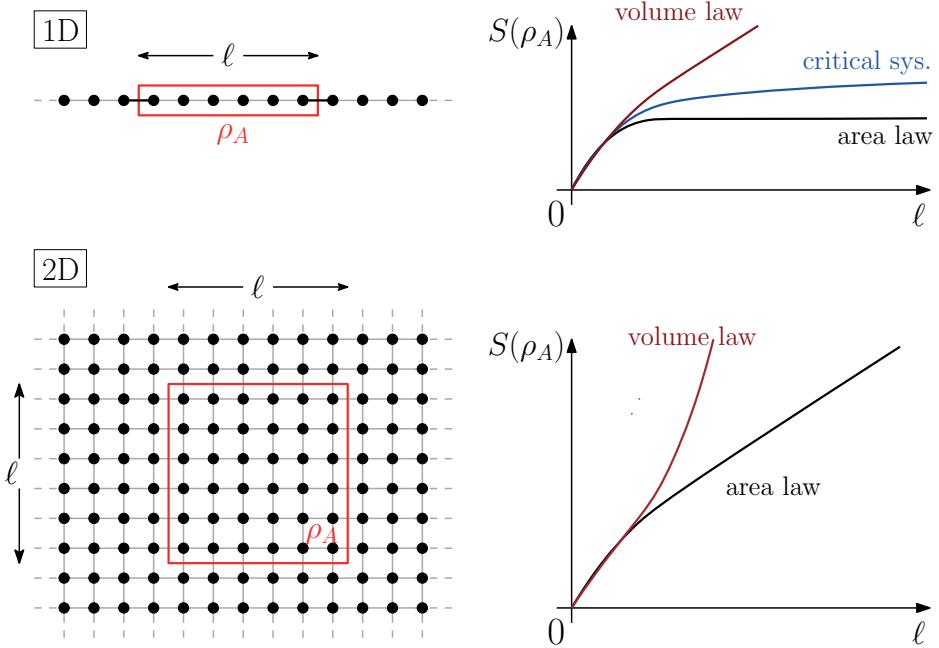


Fig. 11.8 Sketch of the area-law behaviour: the bipartite entanglement entropy of the reduced state  $\rho_A$  of block  $A$  scales as the length of its boundary  $\partial A$ . In 1D (upper panels), the entropy  $S(\rho_A)$  is bounded by a constant, irrespective of the number  $\ell$  of sites in  $A$ . For critical systems one finds a logarithmic increase  $\sim \log \ell$ , while for generic states it exhibits a volume-law linear increase with  $\ell$ . In 2D (lower panels), the entropy  $S(\rho_A)$  grows linearly with  $\ell$ , where  $\ell^2$  is the number of sites in  $A$ . Conversely, for generic states one would find a volume-law quadratic increase  $\sim \ell^2$ .

spins in the region  $A$ . Indeed, if we consider a generic quantum state of the type in Eq. (11.79), where each of the coefficients  $c_{i_1, \dots, i_L}$  has a constant amplitude  $\sim d^{L/2}$  and random phase, to the leading order in  $L$ , the reduced density matrix of  $\ell$  sites is a multiple of the identity. Thus the reduced von Neumann entropy is  $S_\ell \sim \ell \log d$ . However it turns out that, for most of the many-body states one can be interested in, such entanglement will be much less.

**Area law for the entanglement entropy:** *For any ground state of local non-critical Hamiltonians of the type in Eq. (11.81), the entanglement entropy of a generic bipartition  $A|B$  scales as the boundary of one of the two regions:*

$$S(\rho_A) \propto |\partial A|. \quad (11.83)$$

We stress that the area-law behaviour is very peculiar to the low-energy physics of Hamiltonian models as those in Eq. (11.81), and has been verified in a variety of situations. For 1D gapped Hamiltonians, the area law has been rigorously proved by Hastings (2007). For gapless Hamiltonians, as it happens at criticality, subleading corrections, which are at most logarithmic in the volume of  $A$ , may arise. It is

however worth mentioning that there are situations where, despite the locality of the Hamiltonian, the gap closes faster than  $1/L^2$  and the area law is “supercritically” violated by a square root factor  $\sim \sqrt{\ell}$ , that is, exponentially more than the logarithm (Movassagh and Shor, 2016).

The wording “area law” comes from the three-dimensional situation, where the boundary of a region  $A$  with  $\ell^3$  sites is just its surface, and scales as  $\ell^2$ . This has to be contrasted with the so-called *volume law* for the entanglement entropy, which typically occurs for random states or for thermal states. In two-dimensional systems, an area-law behaviour means that the entanglement entropy of the region  $A$  with  $\ell^2$  sites (which is an area itself) grows linearly with  $\ell$ . The most remarkable situation occurs in 1D, where the area-law condition means that the entropy does not scale at all with the size of the region  $A$ . See Fig. 11.8 for a sketch of typical situations in 1D and 2D. Physically speaking, area-law means that the entanglement is concentrated around the boundary of the bipartition. For excited states, or in general in non-equilibrium conditions, this highly restrictive law is not at all guaranteed and a typical volume-law behaviour has to be expected. We mention, however, that a significant counterexample is represented by the eigenstates of many-body localized systems, typically occurring in the presence of both disorder and interactions. These systems, when driven out of equilibrium, usually fail to thermalize, in the sense that their physics is not captured by the standard ensembles of quantum statistical mechanics and the volume-law scaling of the entanglement entropy is not observed.

## 11.5 Matrix product states

We now show how to build up an Ansatz for a quantum state that satisfies an area law in 1D. It is first useful to decompose each site of a chain of length  $L$  into two ancillary subsystems, each of them living on a Hilbert space  $\mathbb{C}^\chi$  of dimension  $\chi$ . In the language of the auxiliary  $2L$  subsystems, any state can be written as

$$|\psi_{\text{aux}}\rangle = \sum_{\vec{\alpha}} \sum_{\vec{\beta}} c_{\vec{\alpha}, \vec{\beta}} |\vec{\alpha}\rangle_A \otimes |\vec{\beta}\rangle_B, \quad (11.84)$$

where  $\vec{\alpha}$  and  $\vec{\beta}$  are two strings of indexes  $\{\alpha_j\}_{j=1,\dots,L}$  and  $\{\beta_j\}_{j=1,\dots,L}$ , each of them running from 1 to  $\chi$ , and we use the notation  $|\vec{\alpha}\rangle_A = |\alpha_1, \alpha_2, \dots, \alpha_L\rangle_A$  and  $|\vec{\beta}\rangle_B = |\beta_1, \beta_2, \dots, \beta_L\rangle_B$ . These represent a basis respectively for the first ( $A$ ) and the second ( $B$ ) subsystem corresponding to each physical site. On neighbouring sites  $(s, s+1)$ , adjacent pairs of such ancillary systems are then linked in a maximally entangled state (for the sake of clarity, the situation is visualized in Fig. 11.9). Mathematically, we can write each pair using its Schmidt decomposition

$$|w_\chi\rangle = \frac{1}{\sqrt{\chi}} \sum_{k=1}^{\chi} |k\rangle_{B_s} \otimes |k'\rangle_{A_{s+1}}, \quad (11.85)$$

so that  $\{|k\rangle_{B_s}\}_{k=1,\dots,\chi}$  and  $\{|k'\rangle_{A_{s+1}}\}_{k'=1,\dots,\chi}$  denote the elements of a given basis, for each of the two ancilla sites. We can thus write the global auxiliary state as

$$|\psi_{\text{aux}}\rangle = \bigotimes_{s=1}^L |w_\chi\rangle = \sum_{\vec{i}} \sum_{\vec{j}} (W_{j_1,i_2} W_{j_2,i_3} \dots W_{j_L,i_1}) |\vec{i}\rangle_A \otimes |\vec{j}\rangle_B, \quad (11.86)$$

where  $W$  is any  $\chi \times \chi$  matrix with maximum rank. Using the same basis defined in Eq. (11.84), this can be safely taken as a multiple of the identity. The Ansatz (11.86) clearly satisfies an area-law behaviour, since for any bipartition one cuts two bonds, and thus  $S(\rho_L) = 2 \log \chi$ . Notice that here we have considered PBC. In the case of OBC, one simply has to omit the term  $W_{j_L,i_1}$ , so that the first and last auxiliary subsystems remain unpaired.

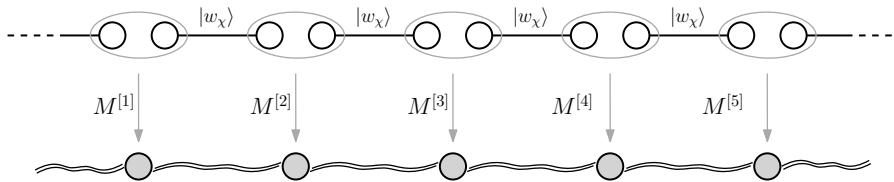


Fig. 11.9 Schematic representation of how to derive a matrix-product-state representation for a 1D system with  $L$  sites (grey circles), from the mapping of a pairwise entangled state of an ancillary chain with  $2L$  sites (white circles).

After having constructed the auxiliary space, we eventually need to map it back to the original space. This operation can be performed by defining a set of  $L$  generic maps on each physical site  $s$ , that is,  $M^{[s]} : \mathbb{C}^\chi \otimes \mathbb{C}^\chi \rightarrow \mathbb{C}^d$ , each of them taking the states of the two auxiliary sites  $A_s$  and  $B_s$  to the states of site  $s$ . A map of this type can be formally written as

$$M^{[s]} = \sum_{i=1}^d \sum_{\alpha,\beta=1}^\chi A_{\alpha,\beta}^{[s]i} |i\rangle \langle \alpha, \beta|, \quad (11.87)$$

where  $A_{\alpha,\beta}^{[s]i}$  denotes a three-rank tensor in which the Roman index  $i$  runs over the  $d$  states of the local Hilbert basis of site  $s$ , while each of the two Greek indexes runs over the  $\chi$  states of the Hilbert basis of one ancillary system. Hereafter we will always adopt the following convention: Greek indexes label states of the auxiliary subsystems ( $\alpha, \beta, \gamma, \dots = 1, \dots, \chi$ ), while Roman indexes label states of the real lattice sites ( $i, j, k, \dots = 1, \dots, d$ ).

Using this approach, it is easy to construct the general form of the resulting wave function in the physical space, after the mapping operation applied to the state  $|\psi_{\text{aux}}\rangle$  of Eq. (11.86). Consider for example the simple case of only two physical sites, with PBC in the ancillary space. Apart from a normalization constant, we

have:

$$\begin{aligned}
M^{[1]} M^{[2]} |\psi_{\text{aux}}\rangle &= \left( M_{A_1 B_1}^{[1]} \otimes M_{A_2 B_2}^{[2]} \right) |\omega_\chi\rangle_{B_1 A_2} |\omega_\chi\rangle_{B_2 A_1} \\
&= \left\{ \sum_{i,\alpha,\beta} A_{\alpha,\beta}^{[1]i} |i\rangle_{A_1 B_1} \langle \alpha, \beta| \right\} \left\{ \sum_{j,\gamma,\delta} A_{\gamma,\delta}^{[2]j} |j\rangle_{A_2 B_2} \langle \gamma, \delta| \right\} \left( \sum_k |k, k\rangle_{B_1 A_2} \sum_l |l, l\rangle_{B_2 A_1} \right) \\
&= \sum_{i,j} \sum_{\alpha,\beta,\gamma,\delta} A_{\alpha,\beta}^{[1]i} A_{\gamma,\delta}^{[2]j} |i, j\rangle_{12} \langle \beta, \gamma| \sum_k |k, k\rangle_{B_1 A_2} \langle \alpha, \delta| \sum_l |l, l\rangle_{B_2 A_1} \\
&= \sum_{i,j} \sum_{\alpha,\beta} A_{\alpha,\beta}^{[1]i} A_{\beta,\alpha}^{[2]j} |i, j\rangle_{12}, \tag{11.88}
\end{aligned}$$

where all the subscripts corresponding to the various subsystems have been specified for clarity. This tells us that, in the resulting state, the two Greek indexes have been contracted; we can thus write it in a compact form as:

$$M^{[1]} M^{[2]} |\psi_{\text{aux}}\rangle = \sum_{i,j} \text{Tr}[A^{[1]i} A^{[2]j}] |i, j\rangle, \tag{11.89}$$

where the trace represents the contraction operation of the two tensors, and Greek indexes have been omitted (for the ease of compactness in the notation, hereafter in most of the formulas we will also omit the tensor product symbol between the various subsystems). The wave function in Eq. (11.89) is an example of a so-called *matrix product state* (MPS).<sup>7</sup> The name comes from the fact that, interpreting each of the three-rank tensors  $A_{\alpha,\beta}^{[s]i}$  as a matrix, once the index  $i$  is fixed, the trace operation is equivalent to a row-by-column matrix multiplication. Every time the Greek indexes are omitted, such as for  $A^{[s]i}$ , we will always treat the corresponding tensor as a matrix, supposing that the Roman index  $i$  is fixed.

The construction presented above can be easily generalized to the case of many sites, in such a way that the resulting  $L$ -site MPS is given by a concatenation of such row-by-column matrix multiplications:

$$|\psi\rangle_{\text{MPS}} = \sum_{\vec{i}} \text{Tr}[A^{[1]i_1} A^{[2]i_2} \cdots A^{[L]i_L}] |i_1, i_2, \dots, i_L\rangle, \tag{11.90}$$

where  $\vec{i} = \{i_j\}_{j=1, \dots, L}$ , as introduced for the first time by Fannes *et al.* (1992). Similarly, for OBC, the resulting MPS is given by:

$$|\psi\rangle_{\text{MPS}} = \sum_{\vec{i}} \sum_{\alpha_1, \dots, \alpha_{L-1}} A_{\alpha_1}^{[1]i_1} A_{\alpha_1, \alpha_2}^{[2]i_2} \cdots A_{\alpha_{L-1}}^{[L]i_L} |i_1, i_2, \dots, i_L\rangle, \tag{11.91}$$

<sup>7</sup>We note that, while Eq. (11.86) in the auxiliary space automatically satisfies an area-law behaviour for the entanglement entropy, this also holds for the MPS form of Eq. (11.89): as a matter of fact, the maps  $M^{[s]}$  onto the physical degrees of freedom are local with respect to the sites (i.e., to the bipartitions), and thus they cannot increase the entanglement entropy. We will however come back later to this point in Sec. 11.6.4, giving some technical details on how to approximate generic many-body states as MPS wave functions. We also mention that, while a MPS can be shown to obey an area law, the converse is not necessarily true (despite this holds, for all the physically relevant ground states of local gapped Hamiltonians). For a formal thorough discussion on the approximability of area-law states with MPS, we refer the interested reader to Verstraete and Cirac (2006), Hastings (2007) and Schuch *et al.* (2008).

where the first and the last tensors have rank two, each of them having only one Greek index. Equivalently one could use the same MPS representation of Eq. (11.90) for periodic systems, and define the two border tensors as

$$\langle l | A^{[1]i_1} \equiv \tilde{A}^{[1]i_1}, \quad A^{[L]i_L} | r \rangle \equiv \tilde{A}^{[L]i_L}, \quad (11.92)$$

$|l\rangle$  and  $|r\rangle$  being the left and right vectors.

### 11.5.1 Examples of MPS wave functions

To fix the ideas, we now provide few simple examples of MPS representations of certain known quantum states. In order to simplify our notation, unless specified, we will always omit all the subscripts corresponding to the associated sites.

#### The GHZ-state

Let us consider the following GHZ-state with  $n$  qubits:

$$|\text{GHZ}\rangle_n = \frac{1}{\sqrt{2}}(|00\cdots 0\rangle + |11\cdots 1\rangle). \quad (11.93)$$

It is easy to see that there is an exact representation of  $|\text{GHZ}\rangle_n$  in terms of a MPS as in Eq. (11.90), with  $\chi = 2$ . We can consider the following matrices:

$$A^{i=0} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A^{i=1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad (11.94)$$

where the superscript  $[s]$  can be omitted, since the GHZ-state (11.93) is translationally invariant and thus all the three-rank tensors  $A^{[s]}$  can be taken equal. Indeed it is easy to verify that  $A^0 A^0 = A^0$ ,  $A^0 A^1 = 0$ ,  $A^1 A^1 = A^1$ , therefore

$$\text{Tr}[A^{i_1} \cdots A^{i_L}] = \begin{cases} \text{Tr}[A^0] = 1 & \text{if } i_1 = \dots = i_L = 0, \\ \text{Tr}[A^1] = 1 & \text{if } i_1 = \dots = i_L = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (11.95)$$

#### The W-state

The  $n$ -qubit Werner state is given by:

$$|\text{W}\rangle_n = \frac{1}{\sqrt{L}}(|10\cdots 0\rangle + |01\cdots 0\rangle + \dots + |00\cdots 1\rangle). \quad (11.96)$$

This example shows that the natural setting to write a MPS representation of the W-state is with OBC. Indeed if we force, for a moment, PBC, a natural choice would be that of taking

$$A^{i=0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A^{i=1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad (11.97)$$

such that  $A^0 A^0 = A^0$ ,  $A^0 A^1 = A^1$ ,  $A^1 A^1 = 0$ , and thus for a given product of  $A$  matrices, one should have only one  $A^1$  otherwise the net result would be the null matrix. However unfortunately, when evaluating the trace  $\text{Tr}[A^{i_1} A^{i_2} \cdots A^{i_L}]$ , one would get a non-zero contribution also for the case of  $\sum_k i_k = 0$ , corresponding to a state  $|00\cdots 0\rangle$  which does not enter the definition (11.96).

A natural way to solve this issue is to consider OBC with the border vectors

$$\langle l| = [1 \ 0], \quad |r\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (11.98)$$

such that

$$A^{i_1} \cdots A^{i_L} = \begin{cases} A^0 & \text{if } \sum_k i_k = 0, \\ A^1 & \text{if } \sum_k i_k = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (11.99)$$

Therefore  $\langle l|A^{i_1} \cdots A^{i_L}|r\rangle = 1$  if and only if  $\sum_k i_k = 1$ , while for all the other choices of  $i_k$  indexes the matrix product state gives zero.

### The Affleck-Kennedy-Lieb-Tasaki state

We now provide the MPS representation of a non-trivial quantum state, which corresponds to the ground state of the Heisenberg spin-1 Hamiltonian

$$H = \sum_i \vec{S}_i \cdot \vec{S}_{i+1} + \frac{1}{3} (\vec{S}_i \cdot \vec{S}_{i+1})^2, \quad (11.100)$$

where  $\vec{S}_i = (S_i^x, S_i^y, S_i^z)$  are the corresponding spin matrices for a spin-1 particle, that is

$$S_i^x = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad S_i^y = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad S_i^z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad (11.101)$$

Such state has been dubbed AKLT, after the name of its four discoverers, Affleck, Kennedy, Lieb, and Tasaki (Affleck *et al.*, 1987).

It can be shown that the AKLT state is easily represented by a MPS with the lowest non-trivial dimension  $\chi = 2$ . In order to explicitly construct it, one needs to follow the procedure detailed in the previous section, and sketched in Fig. 11.9. Each individual spin-1 is replaced by a pair of spin-1/2 particles which are completely symmetrized, that is, of the four possible states one only considers the triplet, naturally identified as the  $S = 1$  states:

$$|+\rangle = |\uparrow\uparrow\rangle, \quad |0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\uparrow\rangle), \quad |- \rangle = |\downarrow\downarrow\rangle. \quad (11.102)$$

On neighbouring sites, adjacent pairs of spin-1/2 particles are placed in the maximally entangled singlet state  $|w_2\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ . In the standard computational basis for the qubits, the  $i$ -th singlet bond  $|w_2\rangle$  is encoded in the following way:

$$|w_2\rangle = \sum_{\beta_i=0}^1 \sum_{\alpha_{i+1}=0}^1 W_{\beta_i, \alpha_{i+1}} |\beta_i\rangle |\alpha_{i+1}\rangle, \quad \text{with } W = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (11.103)$$

Therefore the state  $|\psi_{\text{aux}}\rangle$  with singlets on all bonds factorizes upon splitting any site  $i$  into its two constituents. The identification (11.102) of the symmetrized states of the auxiliary spins with the physical spin can be performed by introducing a mapping from the states of the two auxiliary spin-1/2 particles,  $|\alpha_i \beta_i\rangle \in \{|\uparrow\rangle, |\downarrow\rangle\}^{\otimes 2}$

to the states of the physical spin-1 particles,  $|s_i\rangle \in \{|+\rangle, |0\rangle, |-\rangle\}$ . This is formally done after defining the  $2 \times 2$  matrices

$$M^+ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M^0 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad M^- = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (11.104)$$

The state  $|\psi_{\text{aux}}\rangle = |w_2\rangle^{\otimes L}$  is then mapped into:

$$\begin{aligned} |w_2\rangle^{\otimes L} &\rightarrow \sum_{\vec{i}} \sum_{\vec{\alpha}, \vec{\beta}} (M_{\alpha_1, \beta_1}^{i_1} W_{\beta_1, \alpha_2}) (M_{\alpha_2, \beta_2}^{i_2} W_{\beta_2, \alpha_3}) \cdots (M_{\alpha_L, \beta_L}^{i_L} W_{\beta_L, \alpha_1}) |\vec{i}\rangle \\ &= \sum_{\vec{i}} \text{Tr}[M^{i_1} W M^{i_2} W \cdots M^{i_L} W] |\vec{i}\rangle, \end{aligned} \quad (11.105)$$

which is exactly a MPS of the form (11.90), where the matrices  $\tilde{A}^i = M^i W$  are given by:

$$\tilde{A}^+ = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix}, \quad \tilde{A}^0 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad \tilde{A}^- = \begin{bmatrix} 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}. \quad (11.106)$$

The properly normalized AKLT state is finally obtained by imposing the normalization constraint  $\langle \psi | \psi \rangle$ , which can be satisfied by rescaling the  $\tilde{A}$  matrices by a factor  $2/\sqrt{3}$ . Therefore the normalized matrices  $A$ 's are:

$$A^+ = \begin{bmatrix} 0 & \sqrt{\frac{2}{3}} \\ 0 & 0 \end{bmatrix}, \quad A^0 = \begin{bmatrix} -\frac{1}{\sqrt{3}} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{bmatrix}, \quad A^- = \begin{bmatrix} 0 & 0 \\ -\sqrt{\frac{2}{3}} & 0 \end{bmatrix}. \quad (11.107)$$

## 11.6 Graphical representation of matrix product states

There is an elegant diagrammatic way to represent  $n$ -rank tensors in a graphical form. This turns out to be particularly useful when contractions among several indexes have to be performed. A given tensor can be depicted as a box, having a number of outgoing legs which is equal to its rank  $n$ . For example, for each of the 3-rank tensors  $A^{[s]}$  that enters the MPS decomposition, we will draw a box with one vertical leg, corresponding to its Roman index, and two horizontal legs, corresponding to its Greek indexes (Fig. 11.10, left drawing). Summing up over a given index is denoted by the contraction of one leg going out from a box, with another leg going into another box. For example, the right drawing of Fig. 11.10 represents the following summation among the tensors  $A^{[1]}$  and  $A^{[2]}$ :

$$\sum_{\beta} A_{\alpha, \beta}^{[1]i_1} A_{\beta, \gamma}^{[2]i_2} = (A^{[1]i_1} A^{[2]i_2})_{\alpha, \gamma} \equiv T_{\alpha, \gamma}^{[1, 2]i_1 i_2}. \quad (11.108)$$

The outcome of such operation is a four-rank tensor  $T^{[1, 2]}$  having two (vertical) Roman indexes  $i_1, i_2$  and two (horizontal) Greek indexes  $\alpha, \gamma$ .

In the remainder of the chapter we will mostly deal with MPS wave functions, which represent the simplest example of *tensor networks*, postponing the discussion

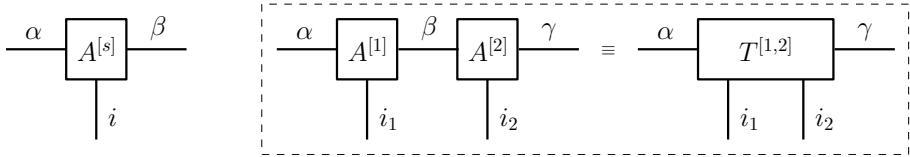


Fig. 11.10 Left: graphical representation of a rank-three tensor  $A_{\alpha,\beta}^{[s]i}$ , with one vertical leg corresponding to the Roman index  $i$ , and two horizontal legs corresponding to the Greek indexes  $\alpha$  and  $\beta$ . Right: row-by-column matrix multiplication among the two tensors  $A^{[1]i_1}$  and  $A^{[2]i_2}$ .

of more complicated tensor-network structures to Sec. 11.9. All the MPS technicalities presented hereafter can be found in the review paper by Schollwöck (2011).

The above contraction of two tensors can be generalized to the trace operation of a generic MPS in Eq. (11.90), which represents a concatenated row-by-column matrix multiplication. Using the graphical notation, the net result is the upper drawing of Fig. 11.11. All the  $L$  horizontal legs are contracted in such a way that the various boxes have been linked together. On the other hand, the vertical legs are left open. In the following we will refer to the horizontal legs (representing Greek indexes) as the *bond links* of the MPS, while the vertical legs (representing Roman indexes) will be called *physical links*. In conclusion, we have found that a MPS denotes a specific form of many-body wave function  $|\psi\rangle \in \mathbb{C}^{2^L}$ , that is

$$|\psi\rangle = \sum_{\vec{i}} c_{i_1, \dots, i_L} |i_1 \dots i_L\rangle \quad \text{with } c_{i_1, \dots, i_L} = \text{Tr}[A^{[1]i_1} \dots A^{[L]i_L}]. \quad (11.109)$$

This has to be contrasted with a generic many-body wave function in the same space, where the  $L$ -rank tensor  $c_{i_1, \dots, i_L}$  cannot be further decomposed. Its graphical representation would be a box with  $L$  vertical legs, corresponding to the Roman indexes  $i_1, \dots, i_L$  (Fig. 11.11 bottom).

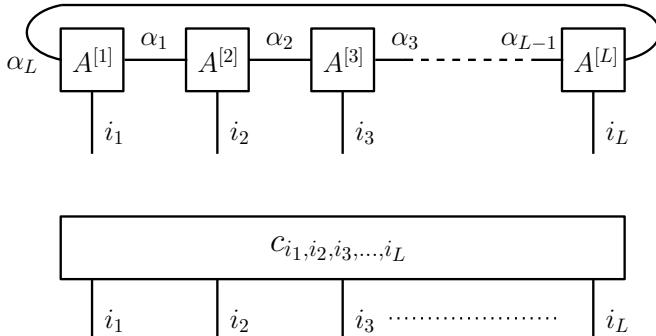


Fig. 11.11 Top: graphical representation of a MPS for  $L$  physical sites, with periodic boundary conditions. Bottom: comparison with the representation of a generic many-body wave function in a  $d^L$ -dimensional Hilbert space.

### Normalization

We now explicitly work out the contractions required to calculate the norm  $N = \langle\psi|\psi\rangle$  of a given quantum state. For a given wave function  $|\psi\rangle = \sum_{\vec{i}} c_{\vec{i}} |\vec{i}\rangle$  we have:

$$N = \sum_{\vec{i}} \sum_{\vec{j}} \langle j_1, \dots, j_L | c_{j_1, \dots, j_L}^* c_{i_1, \dots, i_L} | i_1, \dots, i_L \rangle = \sum_{\vec{i}} |c_{i_1, \dots, i_L}|^2, \quad (11.110)$$

where  $*$  denotes the complex conjugation. In order to do a similar calculation for the MPS state in Eq. (11.90), we first need to clarify all the indexes. Specifically, let us write the “ket state”  $|\psi\rangle_{\text{MPS}}$  and the “bra state”  $\text{MPS}\langle\psi|$  according to

$$|\psi\rangle_{\text{MPS}} = \sum_{\vec{i}} \sum_{\vec{\alpha}} A_{\alpha_1, \alpha_2}^{[1]i_1} A_{\alpha_2, \alpha_3}^{[2]i_2} \cdots A_{\alpha_L, \alpha_1}^{[L]i_L} |i_1, i_2, \dots, i_L\rangle, \quad (11.111)$$

$$\text{MPS}\langle\psi| = \sum_{\vec{j}} \sum_{\vec{\beta}} (A_{\beta_1, \beta_2}^{[1]j_1})^* (A_{\beta_2, \beta_3}^{[2]j_2})^* \cdots (A_{\beta_L, \beta_1}^{[L]j_L})^* \langle j_1, j_2, \dots, j_L|, \quad (11.112)$$

so that a contraction of them gives the norm  $N = \text{MPS}\langle\psi|\psi\rangle_{\text{MPS}}$ :

$$N = \sum_{\vec{i}} \sum_{\vec{\alpha}, \vec{\beta}} \left\{ (A_{\beta_1, \beta_2}^{[1]i_1})^* A_{\alpha_1, \alpha_2}^{[1]i_1} \right\} \left\{ (A_{\beta_2, \beta_3}^{[2]i_2})^* A_{\alpha_2, \alpha_3}^{[2]i_2} \right\} \cdots \left\{ (A_{\beta_L, \beta_1}^{[L]i_L})^* A_{\alpha_L, \alpha_1}^{[L]i_L} \right\}. \quad (11.113)$$

In order to be consistent with our graphical notation, where MPS “ket states” have downward legs, we will represent MPS “bra states” with upward legs, adopting the convention to take the complex conjugate of the corresponding tensors. In this way, for example, the contractions in Eq. (11.113) over the group indexes  $\vec{i}$ ,  $\vec{\alpha}$  and  $\vec{\beta}$  can be visualized as in Fig. 11.12.

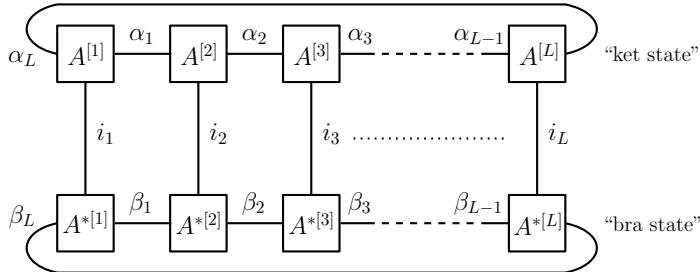


Fig. 11.12 Graphical representation of the contraction of a MPS with itself, which is needed to calculate its norm  $N = \text{MPS}\langle\psi|\psi\rangle_{\text{MPS}}$ . Contracting the MPS “bra state” with the MPS “ket state” equals to summing up over the vertical legs.

### 11.6.1 *Expectation values of observables*

When talking about the statistical mechanics of 1D systems, there is a very powerful tool that enables to work out a complete theoretical characterization of the scaling properties of typical correlators with the distance. This is the so-called *transfer matrix* which, for a MPS, is a linear map on a given site  $s$  of the chain, defined as

$$\mathbb{E}^{[s]} = \sum_{i_s=1}^d (A^{[s]i_s}) \otimes (A^{[s]i_s})^*. \quad (11.114)$$

Namely, within the tensor-network formalism, it corresponds to a four-rank tensor characterized by matrix elements

$$\mathbb{E}_{(\alpha_s, \beta_s), (\alpha_{s+1}, \beta_{s+1})}^{[s]} = \sum_{i_s} (A_{\alpha_s, \alpha_{s+1}}^{[s] i_s})(A_{\beta_s, \beta_{s+1}}^{[s] i_s})^*, \quad (11.115)$$

which can be equivalently seen as a  $\chi^2 \times \chi^2$  matrix, after grouping together the indexes  $\alpha_s$  with  $\beta_s$ , and  $\alpha_{s+1}$  with  $\beta_{s+1}$ . Its graphical representation is depicted in Fig. 11.13, left.

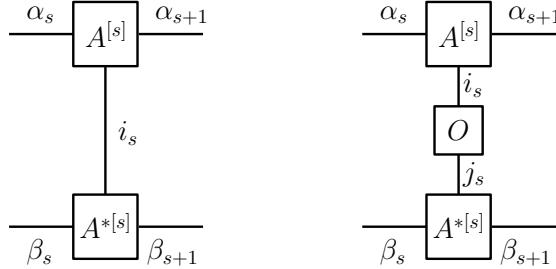


Fig. 11.13 Graphical representation of the transfer matrices  $\mathbb{E}^{[s]}$  (left) and  $\mathbb{E}_O^{[s]}$  (right).

The notion of transfer matrix can be straightforwardly extended to include local operators, according to the following:

$$\mathbb{E}_O^{[s]} = \sum_{i_s=1}^d \sum_{j_s=1}^d (A^{[s] i_s}) \otimes (A^{[s] j_s})^* \langle j_s | O | i_s \rangle, \quad (11.116)$$

where  $O$  is a generic on-site operator living on the local Hilbert space of site  $s$ . It is thus immediate to see that  $\mathbb{E}_I^{[s]} \equiv \mathbb{E}^{[s]}$ . Hereafter we will omit the operator if this coincides with the identity. The graphical representation of  $\mathbb{E}_O^{[s]}$  is provided in Fig. 11.13, right.

Equipped with this tool, we can immediately realize that the norm (11.113) can be cast in a compact form as a product of the  $L$  transfer matrices for all the sites:

$$N = \text{Tr}[\mathbb{E}^{[1]} \mathbb{E}^{[2]} \dots \mathbb{E}^{[L]}]. \quad (11.117)$$

It is also possible to write a similar expression for any expectation values of observables that can be written as tensor products of operators on different sites. For example, for 2-point correlators of the type  $\langle P_k Q_\ell \rangle$ , we have that

$$\text{MPS} \langle \psi | P_k \otimes Q_\ell | \psi \rangle_{\text{MPS}} = \text{Tr} \left[ \mathbb{E}^{[1]} \mathbb{E}^{[2]} \dots \mathbb{E}_{P_k}^{[k]} \mathbb{E}^{[k+1]} \dots \mathbb{E}_{Q_\ell}^{[\ell]} \dots \mathbb{E}^{[L]} \right]. \quad (11.118)$$

Once more, the corresponding graphical representation is shown in Fig. 11.14.

At this stage, it is useful to count the number of operations that are required in order to calculate a sequence of contractions over  $L$  transfer matrices, each of size  $\chi^2 \times \chi^2$ . Using a direct scheme, it is easy to see that for PBC one needs  $O(L \chi^6)$  operations. Indeed a row-by-column matrix multiplication of two matrices of size

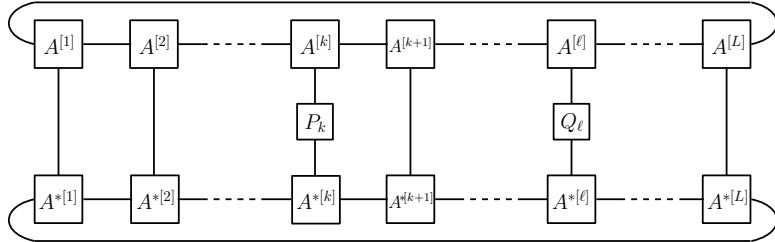


Fig. 11.14 Graphical representation of the contraction scheme that is needed in order to calculate the 2-point correlator  $\text{MPS}\langle\psi|P_kQ_\ell|\psi\rangle_{\text{MPS}}$  in Eq. (11.118).

$m \times m$  costs  $m^3$  operations. On the other hand, for OBC it is more convenient to start contracting transfer matrices from one border: the transfer matrices of site 1 and site  $L$  are indeed of size  $1 \times \chi^2$  and  $\chi^2 \times 1$ , respectively. In this way it is not difficult to see that the total number of required operations in order to contract  $L$  of them scales as  $O(L\chi^4)$ . This shows that it is generally better to work with open boundaries. Unless specified, in the following we will thus specialize to OBC.

**Exercise 11.4** Show that the scaling of the number of operations discussed above is not optimal, since it is possible to perform the same contraction scheme by using  $O(Ld\chi^5)$  and  $O(Ld\chi^3)$  operations, respectively for PBC and for OBC.

### 11.6.2 \* Scaling of correlation functions with the distance

The specific geometry of 1D systems enables us to derive the behaviour of generic two-point correlation functions as a function of the distance, for a given MPS wave function. Namely, we would like to see how the observable

$$C(k, \ell) = \frac{\text{MPS}\langle\psi|P_k \otimes Q_\ell|\psi\rangle_{\text{MPS}}}{\text{MPS}\langle\psi|\psi\rangle_{\text{MPS}}} \quad (11.119)$$

behaves with  $r = |k - \ell|$ . In this respect, the transfer matrix formalism turns out to be particularly useful, since we can rewrite Eq. (11.119) using Eq. (11.118):

$$C(k, \ell) = \frac{\text{Tr}[\mathbb{E}^{[1]} \dots \mathbb{E}_{P_k}^{[k]} \dots \mathbb{E}_{Q_\ell}^{[\ell]} \dots \mathbb{E}^{[L]}]}{\text{Tr}[\mathbb{E}^{[1]} \dots \mathbb{E}^{[L]}]} = \frac{\text{Tr}[\mathbb{E}_{P_k} \mathbb{E}^{\ell-k-1} \mathbb{E}_{Q_\ell} \mathbb{E}^{L-\ell+k-1}]}{\text{Tr}[\mathbb{E}^L]}, \quad (11.120)$$

where the second equality comes from the cyclicity of the trace, and holds for translational invariant systems: in that case, the transfer matrix  $\mathbb{E}^{[s]}$  is independent of the physical site  $s$ , and the superscript  $^{[s]}$  can be safely omitted. The correlation function can thus be evaluated after suitably decomposing  $\mathbb{E}$ .

Unfortunately the expression (11.114) is in general not Hermitian, and may not be diagonalizable. Let us however assume for a moment that this is the case, and write its spectral decomposition:  $\mathbb{E} = \sum_j \lambda_j |\lambda_j\rangle\langle\lambda_j|$ , with  $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$ . If the largest eigenvalue  $\lambda_1$  is non-degenerate, then we know that

$$\lim_{M \rightarrow \infty} \mathbb{E}^M = \lambda_1^M |\lambda_1\rangle\langle\lambda_1|, \quad (11.121)$$

and thus we can write that

$$C(k, \ell) \xrightarrow{L \rightarrow \infty} \frac{\langle \lambda_1 | \mathbb{E}_{P_k} \mathbb{E}^{[\ell-k-1]} \mathbb{E}_{Q_\ell} | \lambda_1 \rangle}{\lambda_1^{\ell-k+1}} = \sum_k \left( \frac{\lambda_k}{\lambda_1} \right)^{\ell-k-1} \frac{\langle \lambda_1 | \mathbb{E}_{P_k} | \lambda_k \rangle \langle \lambda_k | \mathbb{E}_{Q_\ell} | \lambda_1 \rangle}{\lambda_1^2}.$$

Therefore all the correlators can either decay exponentially, being a superposition of exponentials with decay length  $\xi_k = -1/\log \lambda_k$ , or have an infinite range, if  $\langle \lambda_1 | \mathbb{E} | \lambda_1 \rangle$  is finite. A similar situation holds if  $\lambda_1$  is a degenerate eigenvalue.

In the generic case of a non-diagonalizable  $\mathbb{E}$ , it is always possible to transform it in a canonical Jordan form, through a similarity transformation:  $\mathbb{E} = T \mathbb{J} T^{-1}$ . Here  $T$  is an invertible matrix and  $\mathbb{J}$  is a block-diagonal matrix in the Jordan form:

$$\mathbb{J} = \bigoplus_j \mathbb{J}_{d_j}(\eta_j) = \bigoplus_j (\eta_j \mathbb{I}_{d_j} + \mathbb{N}_{d_j}), \quad (11.122)$$

where  $\mathbb{J}_{d_j}(\eta_j)$  is a Jordan block of dimension  $d_j$ , corresponding to the generalized eigenvalue  $\eta_j$ ,  $\mathbb{I}_{d_j}$  denotes a  $d_j \times d_j$  identity matrix, and  $\mathbb{N}_{d_j}$  a nilpotent matrix such that  $[\mathbb{N}_{d_j}]^p = 0$ ,  $\forall p \geq d_j$ . It can be proven that

$$[\mathbb{J}_{d_j}(\eta_j)]^L = \eta_j^L \cdot \mathbb{Q}^{(L)}(\eta_j), \quad \text{with } \mathbb{Q}^{(L)}(\eta_j) = \sum_{q=0}^{d_j-1} \binom{L}{q} \eta_j^{-q} (\mathbb{N}_{d_j})^q, \quad (11.123)$$

so that

$$\lim_{L \rightarrow \infty} [\mathbb{J}_{d_j}(\eta_j)]^L = 0_{d_j}, \quad \text{for } |\eta_j| < 1. \quad (11.124)$$

It is eventually possible to prove the final result:

$$\text{MPS} \langle \psi | P_k Q_\ell | \psi \rangle_{\text{MPS}} = c_1 + \sum_{k>1} c_k \eta_k^{r-1}, \quad (11.125)$$

thus showing that the decay terms are of the type  $e^{-r \xi_k}$ , with  $\xi_k = -1/\log \eta_k$ .

In conclusion, any finite-dimensional MPS is only able to approximate a given correlation functions as a superposition of exponentials. For practical purposes, this approximation might be very good on short distances, even for power laws.

### Examples

- The transfer matrix of the AKLT state, using the normalized MPS representation of Eq. (11.107), is given by

$$\mathbb{E} = A^+ \otimes (A^+)^* + A^0 \otimes (A^0)^* + A^- \otimes (A^-)^* = \frac{1}{3} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}. \quad (11.126)$$

It is easy to show that  $\mathbb{E}$  has eigenvalues  $\lambda_1 = 1$ ,  $\lambda_2 = \lambda_3 = \lambda_4 = -1/3$ . Therefore correlation functions drop exponentially with decay length  $\xi = -1/\log 1/3 \approx 0.91$ .

- For the GHZ state, we use the MPS representation of Eq. (11.94) and find

$$\mathbb{E} = A^0 \otimes (A^0)^* + A^1 \otimes (A^1)^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (11.127)$$

Therefore, since  $\lambda_1 = \lambda_2 = 1$ , infinite-range correlations set in.

### 11.6.3 Gauge freedom

We now discuss the fact that the representation of a wave function in terms of a MPS is not unique, thus exhibiting a gauge freedom. Indeed, in the wave function representation of Eq. (11.90) we can replace the tensors in such a way that

$$A^{[k]i_k} \rightarrow X_k A^{[k]i_k} X_{k+1}^{-1}, \quad \forall k = 1, \dots, \ell, \quad (11.128)$$

where  $X_k$  are generic non-singular (invertible)  $\chi \times \chi$  matrices. Indeed, working for simplicity with PBC, one has

$$\text{Tr}[A^{[1]i_1} A^{[2]i_2} \dots] = \text{Tr}[(X_1 A^{[1]i_1} X_2^{-1})(X_2 A^{[2]i_2} X_3^{-1}) \dots (X_L A^{[L]i_L} X_1^{-1})], \quad (11.129)$$

where the equality holds since  $X_k X_k^{-1} = \mathbb{I}$ . Let us adopt OBC, since these are the most convenient boundary conditions.

#### 11.6.3.1 Isometric gauge

There is a particularly useful gauge named *isometric gauge*, which turns out to drastically simplify all the calculations of observables, as we shall show below.<sup>8</sup> We consider a generic MPS representation of the many-body amplitude

$$c_{i_1, \dots, i_L} = \underbrace{A^{[1]i_1}}_{1 \times \chi_1} \underbrace{A^{[2]i_2}}_{\chi_1 \times \chi_2} \underbrace{A^{[3]i_3}}_{\chi_2 \times \chi_3} \dots \underbrace{A^{[L]i_L}}_{\chi_{L-1} \times 1}, \quad (11.130)$$

where the various tensors  $A^{[s]i_s}$  can have different dimensions as indicated here, such to preserve the row-by-column construction. In order to put such MPS in the isometric gauge, a number of sequential singular value decompositions (SVD) have to be performed.

Let us start from the first site, being represented by the tensor  $A_\alpha^{[1]i_1}$  (the subscript  $\alpha$  accounts for the only non-trivial Greek index). We perform a SVD with respect to the indexes  $i_1$  (row index) and  $\alpha$  (column index):

$$A_\alpha^{[1]i_1} = \sum_{\lambda=1}^{\chi_1} V_\lambda^{[1]i_1} X_{\lambda\alpha}^{[1]}, \quad \text{where } X_{\lambda\alpha}^{[1]} = \Sigma_{\lambda\lambda}^{[1]} W_{\lambda\alpha}^{[1]}, \quad (11.131)$$

with  $\sum_{i_1} V_\lambda^{[1]i_1} (V_{\lambda'}^{[1]i_1})^* = \delta_{\lambda\lambda'}$ , that is, in matrix form  $(V^{[1]})^\dagger V^{[1]} = \mathbb{I}$ . We will call such condition for  $V^{[1]}$  the *left isometric gauge*. A similar relation holds for  $W^{[1]}$ , such that  $W^{[1]} (W^{[1]})^\dagger = \mathbb{I}$ . The diagonal matrix  $\Sigma^{[1]}$  contains the singular values of the SVD. Notice that, in order to apply the SVD, we require that  $d \geq \chi_1$ . Substituting Eq. (11.131) into the expression (11.130) for  $c_{i_1, \dots, i_L}$ , we find

$$c_{i_1, \dots, i_L} = V^{[1]i_1} X^{[1]} A^{[2]i_2} A^{[3]i_3} \dots A^{[L]i_L}. \quad (11.132)$$

Defining now  $B^{[2]} = X^{[1]} A^{[2]}$ , we can repeat the above strategy and perform a SVD on it. Here we operate on a three-rank tensor, of the form  $B_{\alpha\beta}^{[2]i_2}$ . The SVD

---

<sup>8</sup>This condition however does not exhaust the gauge freedom on the tensors, since there is still a unitary left. Fixing it would then give the canonical form, which is basically unique.

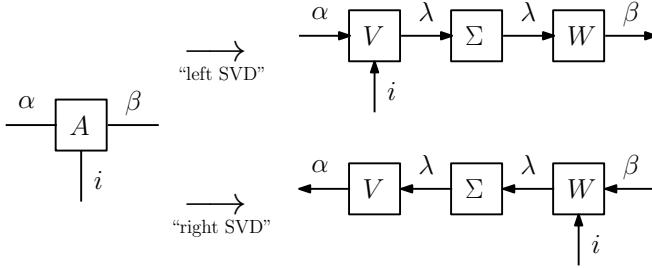


Fig. 11.15 Graphical representation of the “left SVD” and of the “right SVD”. The arrows encode the shaping of the matrix indexes: ingoing arrows stand for row indexes, while outgoing arrows stand for column indexes.

has to be performed with respect to the composite index  $(i_2, \alpha)$  (row index) and  $\beta$  (column index):

$$B_{\alpha\beta}^{[2]i_2} = \sum_{\lambda=1}^{\chi_2} V_{\alpha\lambda}^{[2]i_2} X_{\lambda\beta}^{[2]}, \quad \text{where } X_{\lambda\beta}^{[2]} = \Sigma_{\lambda\lambda}^{[2]} W_{\lambda\beta}^{[2]}, \quad (11.133)$$

with  $V^{[2]}$  respecting the left isometric gauge, that is  $\sum_{i_2, \alpha} V_{\alpha\lambda}^{[2]i_2} (V_{\alpha\lambda}^{[2]i_2})^* = \delta_{\lambda\lambda'}$ , or equivalently  $(V^{[2]})^\dagger V^{[2]} = \mathbb{I}$ . This kind of SVD following the above index compression will be referred to as the “left SVD”, as depicted in Fig. 11.15, upper panel. Similarly to before, we require  $d\chi_1 \geq \chi_2$ . Substituting Eq. (11.133) into Eq. (11.132) we get

$$c_{i_1, \dots, i_L} = V^{[1]i_1} V^{[2]i_2} X^{[2]} A^{[3]i_3} \dots A^{[L]i_L}. \quad (11.134)$$

It is now easy to recognize that we can define  $B^{[3]i_3} = X^{[2]} A^{[3]i_3}$  and iterate this procedure, until we reach the right boundary of the chain. So far, we will obtain a representation of the type:

$$c_{i_1, \dots, i_L} = V^{[1]i_1} \dots V^{[L-1]i_{L-1}} B^{[L]i_L}, \quad \text{with } (V^{[k]})^\dagger V^{[k]} = \mathbb{I} (\forall k). \quad (11.135)$$

It is even possible to perform an analogous construction starting from the right edge, and retaining the  $W$  matrices of the various SVD, keeping attention to the index compression leading to the “right SVD”, as depicted in Fig. 11.15, lower panel. When reaching the left end of the chain, we would have obtained a complementary representation of the type:

$$c_{i_1, \dots, i_L} = B^{[1]i_1} W^{[2]i_2} \dots W^{[L]i_L}, \quad \text{with } W^{[k]} (W^{[k]})^\dagger = \mathbb{I} (\forall k), \quad (11.136)$$

where the condition for the  $W^{[k]}$  is now called *right isometric gauge*. In general, one could even employ a mixed left/right isometric gauge such that, for a given site  $s$  in the middle of the chain, all the tensors on the left of it are in the left isometric gauge, while all the tensors on the right of it are in the right isometric gauge. Using the graphical notation, these conditions can be elegantly explained through the diagrams of Fig. 11.16.

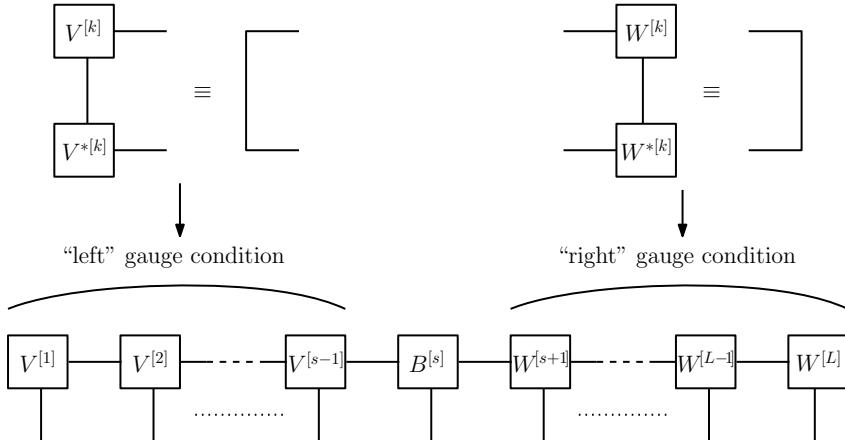


Fig. 11.16 Graphical representation of the unitary gauge for a generic MPS wave function.

#### MPS contraction for observables

We shall now see the isometric gauge at work, by realizing that it enables to calculate expectation values of observables a much more efficiently. Suppose, for example, that we want to measure a one-point observable  $P$  on the site  $k$ . It is then convenient to employ the isometric gauge with respect to point  $k$ . In this way it is easy to see that the whole contraction of the MPS that is required to calculate it, is reduced to a one-point contraction:

$$\text{MPS} \langle \psi | P_k | \psi \rangle_{\text{MPS}} = \sum_{\alpha, \beta} \left[ \mathbb{E}_P^{[k]} \right]_{(\alpha\alpha), (\beta\beta)}. \quad (11.137)$$

Similarly, for two-point observables living on sites  $k$  and  $\ell$ , we have:

$$\text{MPS} \langle \psi | P_k Q_\ell | \psi \rangle_{\text{MPS}} = \sum_{\alpha, \beta} \left[ \mathbb{E}_P^{[k]} \mathbb{E}^{[k+1]} \dots \mathbb{E}^{[\ell-1]} \mathbb{E}_Q^{[\ell]} \right]_{(\alpha\alpha), (\beta\beta)}. \quad (11.138)$$

As a matter of fact, provided one has chosen the proper isometric gauge conditions, the number of operations required in these two situations, respectively  $d^2 \chi^2$  for one-point and  $O((\ell - k + 1) d \chi^3)$  for two-point observables, is drastically reduced as compared to the one needed to calculate a generic full diagram of the type in Fig. 11.14, that is,  $O(L d \chi^3)$ .

#### 11.6.4 Schmidt decomposition of a MPS

Apart from providing a computational speedup, the isometric gauge has an important conceptual significance as well, since it gives easy access to the Schmidt decomposition of the MPS wave function with respect to the intermediate point. Indeed, suppose that we are in the isometric gauge of Fig. 11.16. By performing a

final “left SVD” on the central tensor:  $B_{\alpha\beta}^{[s]i_s} = \sum_{\lambda} V_{\alpha\lambda}^{[s]i_s} \Sigma_{\lambda\lambda}^{[s]} \widetilde{W}_{\lambda\beta}^{[s]}$ , and defining<sup>9</sup>

$$|l_{\lambda}\rangle = \sum_{i_1, \dots, i_s} [V^{[1]i_1} \dots V^{[s]i_s}]_{\lambda} |i_1, \dots, i_s\rangle, \quad (11.139)$$

$$|r_{\lambda}\rangle = \sum_{i_{s+1}, \dots, i_L} [\widetilde{W}^{[s]} W^{[s+1]i_{s+1}} \dots W^{[L]i_L}]_{\lambda} |i_{s+1}, \dots, i_L\rangle, \quad (11.140)$$

it is immediate to see that, in virtue of the isometric condition for the  $V$  and  $W$  matrices, we have  $\langle l_{\alpha}|l_{\beta}\rangle = \delta_{\alpha\beta}$  and similarly  $\langle r_{\alpha}|r_{\beta}\rangle = \delta_{\alpha\beta}$ . We are thus in the position to rewrite the MPS wave function in a Schmidt form:

$$|\psi\rangle_{\text{MPS}} = \sum_{\lambda=1}^{\chi} \Sigma_{\lambda\lambda}^{[k]} |l_{\lambda}\rangle \otimes |r_{\lambda}\rangle, \quad (11.141)$$

where the squares of the singular values of  $B^{[s]}$ , namely  $(\Sigma_{\lambda\lambda}^{[k]})^2$ , are the eigenvalues of the reduced density matrix of the subsystem with sites from 1 to  $s$ .

Since the total number of non-zero singular values cannot exceed  $\chi$ , the maximum amount of bipartite entanglement entropy is obtained when the  $\chi \times \chi$  reduced density matrix  $\rho_{[1\dots s]}$  of the first  $s$  sites is a multiple of the identity. In that case

$$S(\rho_{[1\dots s]}) = - \sum_{\lambda=1}^{\chi} (\Sigma_{\lambda\lambda}^{[k]})^2 \log \left[ (\Sigma_{\lambda\lambda}^{[k]})^2 \right] \sim \log \chi. \quad (11.142)$$

We have thus discovered that the bipartite entanglement of a MPS wave function cannot exceed a threshold which is equal to the logarithm of the bond link  $\chi$ .

The isometric gauge thus provides a simple way to keep the bond link of the MPS controlled, according to the amount of entanglement entropy at each bipartition. In practice, one can iterate the SVD sequence described above, *sweeping* from left to right in the chain (or vice versa), and retaining only the largest resulting  $\chi$  singular values (with the corresponding left and right singular vectors). By construction, after sweeping back and forth until convergence, this procedure eventually produces the most accurate approximation of a given MPS wave function in terms of a new MPS with a bond link  $\chi' \leq \chi$ .

### MPS representation of a generic quantum state

The sweeping procedure outlined above also enables to write a generic many-body wave function of the type in Eq. (11.79) as a MPS, without any approximation. The idea is to start from the  $L$ -rank tensor  $c_{i_1, \dots, i_L}$  and sequentially separate the various site indexes through a sequence of “left SVDs” of properly compressed indexes, following the scheme of Fig. 11.17. It is not difficult to realize that, at each step from the left to the right, the bond dimension of the generated tensors  $V^{[k]}$  keeps increasing exponentially: indeed  $V^{[1]i_1}$  is a  $1 \times d$  matrix,  $V^{[2]i_2}$  is a  $d \times d^2$  matrix, and so on (at the  $k$ -th step,  $V^{[k]i_k}$  is a  $d^{k-1} \times d^k$  matrix,  $d$  being the on-site Hilbert

<sup>9</sup>The tilde in  $\widetilde{W}^{[s]}$  denotes the fact that such tensor has only two Greek indexes and does not have a vertical leg, contrary to the other three-index tensors  $W^{[k]}$  for  $s+1 \leq k \leq L$ .

space dimension<sup>10</sup>). Analogously, one can proceed with a sequence of “right SVDs” from the right to the left, thus generating the tensors  $W^{[k]}$  with an exponentially increasing bond:  $W^{[L]i_L}$  is a  $d \times 1$  matrix,  $W^{[L-1]i_{L-1}}$  is a  $d^2 \times d$  matrix, and so on. The resulting bond-link size pattern is the following:

$$\{\chi_j\}_{j=1,\dots,L} = \{1, d, d^2, \dots, d^{L/2-1}, d^{L/2}, d^{L/2-1}, \dots, d^2, d, 1\}, \quad (11.143)$$

where the largest value is achieved for the central one, connecting site  $\frac{L}{2}$  with  $\frac{L}{2}+1$ , having a dimension  $\chi_{L/2} = d^{L/2}$ . Since this procedure is exact, the exponential growth of  $\chi$  in the middle is necessary to keep track of all the  $c_{i_1,\dots,i_L}$  complex amplitudes of the wave function in its original representation.

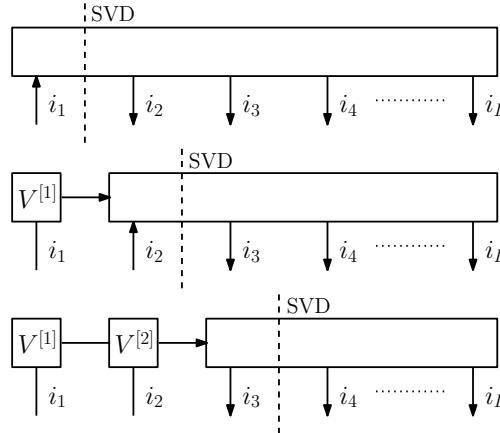


Fig. 11.17 Sequential SVD operations that are needed to decompose a generic  $L$ -rank tensor into a MPS with  $L$  three-rank tensors.

Therefore, every time we want to approximate a given state  $|\psi\rangle$  with a MPS, in such a way to reduce the number of parameters to a tractable size, we shall perform the above sweeping procedure. Typically during the SVD sequence, it is convenient to keep the bond-link dimension in the middle of the chain below a fixed threshold  $\chi_{\max}$ . At the  $k$ -th step, the wave function (11.141) is thus approximated by

$$|\psi'\rangle = \sum_{\lambda=1}^{\chi_{\max}} \Sigma_{\lambda\lambda}^{[k]} |l_\lambda\rangle \otimes |r_\lambda\rangle. \quad (11.144)$$

The committed error is

$$\varepsilon_{\text{trunc}}^{[k]} = \| |\psi\rangle - |\psi'\rangle \|^2 = 1 - \langle \psi | \psi' \rangle = 1 - \sum_{\lambda=1}^{\chi_{\max}} (\Sigma_{\lambda\lambda}^{[k]})^2 = \sum_{\lambda=\chi_{\max}+1}^{\chi} (\Sigma_{\lambda\lambda}^{[k]})^2. \quad (11.145)$$

When performing many cuts along the whole chain, the errors are roughly independent, so that the total committed error is approximately

$$\varepsilon_{\text{trunc}}^{\text{TOT}} = \| |\psi\rangle - |\psi_{\chi_{\max}}\rangle \|^2 \approx \sum_{j=1}^{L-1} \varepsilon_{\text{trunc}}^{[j]}. \quad (11.146)$$

<sup>10</sup>To be precise, the left and right bond dimensions of  $V^{[k]i_k}$  will be  $\min(d^{k-1}, d^{L-k+1})$  and  $\min(d^k, d^{L-k})$ , respectively.

From what we learned before, we can conclude that for typical many-body ground states of local Hamiltonians, the total truncation error can be always kept under control with a relatively small value of  $\chi_{\max}$ , which does not scale with  $L$ .

### 11.7 Ground-state search in the Hilbert space corner

Having at our disposal the required technical skills that are needed to manipulate matrix product states, we are now in the position to describe the working mechanism of variational optimization protocols over such Ansatzes, that are routinely employed to investigate certain quantum many-body systems. A crucially important task in this context is understanding the ground-state properties of low-dimensional (typically 1D or quasi-1D) quantum lattice Hamiltonians with local interactions. In view of the considerations done in Sec. 11.4, the search of such state in the MPS class is optimal, since (almost) every quantum state in an arbitrarily large 1D system satisfying an area law can be always accurately approximated by a MPS with a finite bond link.

To fix the ideas, let us concentrate on local Hamiltonians that can be written as

$$H = \sum_j h_{j,j+1}, \quad (11.147)$$

where  $h_{j,j+1}$  is an Hermitian operator having support on sites  $j$  and  $j + 1$ . The variational protocol for the ground-state search works as follows.

- (1) Start from a given initial state, which can be chosen randomly over the set of MPS with finite bond link  $\chi$ :

$$|\psi\rangle_{\text{MPS}} = |\psi[A^{[1]}, A^{[2]}, \dots, A^{[L]}]\rangle. \quad (11.148)$$

To shorten the notation, from now on we will omit the subscript  $\text{MPS}$ .

- (2) Pick one site  $s$  and optimize the energy cost function  $E = \langle H \rangle$  over the tensor  $A^{[s]}$ . This equals to finding the minimum of the functional

$$E(\mathbf{X}) = \frac{\langle \psi(\mathbf{X}) | H | \psi(\mathbf{X}) \rangle}{\langle \psi(\mathbf{X}) | \psi(\mathbf{X}) \rangle}, \quad (11.149)$$

as a function of the  $(d\chi^2)$  entries of  $A_{\alpha_{s-1}\alpha_s}^{[s]i_s}$ , identified as the unknown  $\mathbf{X}$ . All the other tensors  $A^{[k \neq s]}$  composing the MPS (11.148) are kept fixed.

- (3) Iterate the previous step over all the lattice sites, going back and forth until the energy  $E$  has reached convergence.

Let us now explain in more detail how to perform the minimization of the functional (11.149), on a generic site  $s$  of the chain. It is very convenient to adopt the isometric gauge centered on  $s$ , in such a way that

$$\langle \psi(\mathbf{X}) | \psi(\mathbf{X}) \rangle = \sum_i \sum_{\alpha,\beta} (A_{\alpha\beta}^{[s]i_s})^* A_{\alpha\beta}^{[s]i_s}. \quad (11.150)$$

If we interpret  $\mathbf{X}$  as a vector, it is immediate to see that our task is equivalent to minimizing  $\langle \psi(\mathbf{X}) | H | \psi(\mathbf{X}) \rangle$  under the constraint  $\langle \mathbf{X} | \mathbf{X} \rangle = 1$ . Now, since  $|\psi(\mathbf{X})\rangle$  is linear in the unknown  $\mathbf{X}$ , we can equivalently see our problem as a quadratic form:

$$\langle \psi(\mathbf{X}) | H | \psi(\mathbf{X}) \rangle = \mathbf{X}^\dagger \mathbb{H} \mathbf{X}, \quad (11.151)$$

where  $\mathbb{H}$  is an Hermitian matrix which denotes the effective Hamiltonian on the site  $s$ . We can thus interpret the minimization of the functional (11.149) as an eigenvalue problem, such that the unknown satisfies

$$\mathbb{H} \mathbf{X} = E_{\min} \mathbf{X}, \quad (11.152)$$

being  $E_{\min}$  the smallest eigenvalue of  $\mathbb{H}$ . Therefore once  $\mathbb{H}$  is known, we just need to find its smallest eigenvector and plug it into the entries of the target tensor  $A^{[s]}$ .

In order to compute  $\mathbb{H}$ , we first note that

$$\langle \psi(\mathbf{X}) | \psi(\mathbf{X}) \rangle = \sum_j \langle \psi(\mathbf{X}) | h_{j,j+1} | \psi(\mathbf{X}) \rangle = \sum_j \mathbf{X}^\dagger \mathbb{H}_j \mathbf{X}. \quad (11.153)$$

The various matrices  $\mathbb{H}_j$  can be easily computed by suitably contracting the various indexes, according to the position of the optimization site  $k$  with respect to the sites  $j$  and  $j+1$ . As an illustrative example, for  $j+1 < k$  we obtain  $\mathbb{H}_j = \mathbb{L}_j[k] \otimes \mathbb{I} \otimes \mathbb{I}$ ; for  $j+1 = k$ , we obtain  $\mathbb{H}_j = \mathbb{A}[k] \otimes \mathbb{B} \otimes \mathbb{I}$ . The corresponding graphical representations are given in Fig. 11.18.

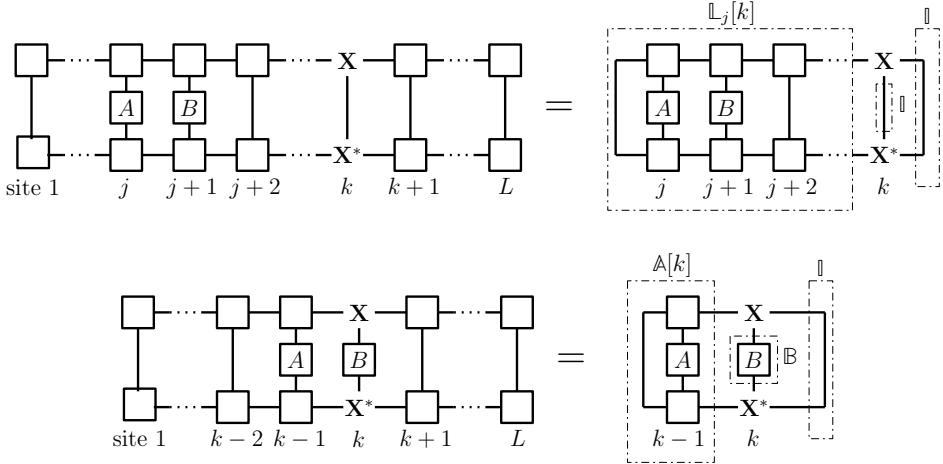


Fig. 11.18 Graphical representation of the contractions that are needed to compute the effective Hamiltonian  $\mathbb{H}_j$ , during the optimization on site  $k$ . The upper diagram is for  $j+1 < k$ , while the lower diagram is for  $j+1 = k$ .

It is possible to speed up the algorithm by co-moving the isometric gauge, so that it is always centered on the site where the minimization is being performed. In practice, sweeping throughout the lattice site-by-site from left to right (or from

right to left), amounts to performing a “left SVD” (resp. “right SVD”) each time. Thus it is computationally cheap to build up the various compound matrices from the previous steps (e.g.  $\mathbb{L}_j[k+1]$  is easily constructed from  $\mathbb{L}_j[k]$ ).

### Periodic boundary conditions

Unfortunately, with PBC the isometric gauge cannot help in the minimization procedure, since there is no edge where to start contracting the various isometries. In any case, the MPS wave function (11.148) remains linear in the  $A^{[s]}$  tensors, therefore the minimization problem (11.149) can be written as

$$E(\mathbf{X}) = \frac{\mathbf{X}^\dagger \mathbb{H} \mathbf{X}}{\mathbf{X}^\dagger \mathbb{N} \mathbf{X}}, \quad (11.154)$$

where the effective matrices  $\mathbb{H}$  and  $\mathbb{N}$  are computed similarly as before. The computational cost of the various contractions now scales as  $\chi^5$ , instead than  $\chi^3$ . The most delicate issue is that we are now facing a generalized eigenvalue problem:

$$\mathbb{H} \mathbf{X} = \lambda \mathbb{N} \mathbf{X}. \quad (11.155)$$

The difference with respect to the OBC case is the presence of the matrix  $\mathbb{N}$  which, in Eq. (11.152), was coinciding with the identity.

The generalized eigenvalue problem of Eq. (11.155) presents some caveats, which can make it ill conditioned from a numerical point of view. These heavily depend on the spectral properties of  $\mathbb{N}$ , also-called the *mass matrix*. If  $\mathbb{N}$  has zero eigenvalues, its kernel should be excluded. Moreover, if it possesses very small eigenvalues, the system can be very sensitive to numerical inaccuracies close to that subspace.

As we will show later in Sec. 11.7.2, the variational procedure described here is formally equivalent to a renormalization algorithm of the Hilbert space that goes under the name of the *density-matrix renormalization group* (DMRG).

### MPS and excited states

We have shown how to manipulate MPS states in order to minimize the energy cost functional, so to achieve a good representation of the ground state wave function  $|\psi_0\rangle = |\psi_0[A_0^{[1]}, A_0^{[2]}, \dots, A_0^{[L]}]\rangle$ ; here the subscript  $_0$  denotes the tensors resulting from each variational minimization of the type in Eq. (11.149). On top of this, it is also possible to target low-lying excited states  $|\psi_1\rangle, |\psi_2\rangle, \dots$  (e.g., in order to compute energy gaps) with a very similar procedure. To this purpose, it is first mandatory to find  $|\psi_0\rangle$ . Then the first excited state is obtained after another variational minimization of the energy:

$$|\psi_1\rangle = \min_{A^{[1]}, \dots, A^{[L]}} \frac{\langle \psi[A^{[1]}, \dots, A^{[L]}] | H | \psi[A^{[1]}, \dots, A^{[L]}] \rangle}{\langle \psi[A^{[1]}, \dots, A^{[L]}] | \psi[A^{[1]}, \dots, A^{[L]}] \rangle} \quad (11.156)$$

with the constraint

$$\langle \psi[A^{[1]}, \dots, A^{[L]}] | \psi[A_0^{[1]}, \dots, A_0^{[L]}] \rangle = 0, \quad (11.157)$$

which expresses the orthogonalization of  $|\psi_1\rangle$  with respect to  $|\psi_0\rangle$ . The minimization can be done similarly as before, applying steps (1)–(3) iteratively. One only

needs to modify the algorithm for the eigenvalue problem. This can be done with a standard extension of the usually employed iterative eigensolvers: for example, using the Lanczos iterations, one needs to take care that the next generated Lanczos state is orthonormalized not only with respect to the previous Lanczos states, but also to the already constructed ground state of the effective Hamiltonian.

This procedure can be sequentially iterated, so to build a hierarchy of excited states. However it becomes rather unpractical to target a high-lying excited state, unless such state is known to be the ground state of another sector of the Hilbert space decomposed according to some good quantum number. Then the calculation is just a ground-state calculation in that different sector.<sup>11</sup> Moreover and most importantly, the MPS representation is believed to be valid only for the ground state and the low-lying states, where the area-law scaling of entanglement is respected. In Sec. 11.8 we will describe an alternative strategy to address the excited part of the spectrum, through the simulation of the real-time evolution.

### 11.7.1 Density-matrix renormalization group

The DMRG algorithm has been devised by White (1992) as a clever renormalization procedure to describe ground states of strongly correlated quantum systems on 1D chains. Below we first provide details of this algorithm, and later in Sec. (11.7.2) demonstrate its substantial equivalence with the variational ground-state search over the MPS Ansatz described above.

The basic strategy of DMRG is to construct a portion of the system (called the *system block*) and then recursively enlarge it, until the desired system size is reached. At every step, the basis of the corresponding Hamiltonian is truncated, so that the size of the Hilbert space is kept manageable as the physical system grows. The truncation of the Hilbert space is performed by retaining the eigenstates corresponding to the  $\chi$  highest eigenvalues of the block reduced density matrix.

#### Infinite-system DMRG

Keeping in mind the main idea, let us now formulate the structure of the so-called *infinite-system DMRG*, where the system is iteratively enlarged until the ground-state properties one is interested in (e.g., its energy per site) have converged. Being this an iterative algorithm, we will proceed step by step.

- (1) Start with a block composed of one site,  $\mathcal{B}_{l/r}(1, d)$ , where the arguments of  $\mathcal{B}$  refer to the number of sites it embodies, and to the number of states used to describe it. The subscript  $l/r$  stands for the block corresponding to the left/right edge of the chain; for a system having global reflection symmetry, one can safely take  $\mathcal{B}_l(1, d) = \mathcal{B}_r(1, d)$ . From the computational point of view, with

<sup>11</sup>We should mention that it is possible to construct a MPS variational Ansatz which exploits quantum numbers, in the sense that the corresponding symmetry of the problem is automatically enforced in the structure of the various tensors  $A^{[k]}$ . This procedure is particularly simple for Abelian symmetries (see, for example, McCulloch (2007)).

$\mathcal{B}_{l/r}(\ell, \chi_\ell)$  we intend to store all the information about a given block of  $\ell$  sites: its Hamiltonian  $H_B$ , the basis, and other operators that will be introduced later. Notice that  $H_B$  includes only local terms (i.e., local and interaction terms where only sites belonging to the block are involved).

- (2) Add a site to the right of the previously created block (at the first stage  $\ell = 1$ ), thus creating the so-called *left-enlarged block*:  $[\mathcal{B}_l(\ell, \chi_\ell) \ominus \bullet]$ , where the symbol  $\bullet$  denotes the added site and  $\ominus$  is a link between the two objects. The corresponding Hamiltonian  $H_E$  is composed by the local Hamiltonians of the block and the site, plus the interaction term:

$$H_E = H_B + H_s + H_{Bs}. \quad (11.158)$$

- (3) Couple the enlarged block to a similarly constructed right-enlarged block, namely:  $[\bullet \ominus \mathcal{B}_r(\ell, \chi_\ell)]$  (for a system with reflection symmetry, the right-enlarged block Hamiltonian  $H_{E'}$  can be obtained by just reflecting the left-enlarged block), thus creating a *super-block*:  $[\mathcal{B}_l(\ell, \chi_\ell) \ominus \bullet \ominus \bullet \ominus \mathcal{B}_r(\ell, \chi_\ell)]$ . The super-block Hamiltonian  $H_{SB}$  is thus given by:

$$H_{SB} = H_E + H_{E'} + H_{ss'}. \quad (11.159)$$

The sites  $s$  and  $s'$  are often referred to as the *free sites*.

- (4) Diagonalize the matrix  $H_{SB}$ , and find its ground state  $|\psi_{GS}\rangle$ , which can be written as:

$$|\psi_{GS}\rangle = \sum_{\alpha, \beta} \sum_{a, b} \psi_{\alpha, a, b, \beta} |\alpha\rangle \otimes |a\rangle \otimes |b\rangle \otimes |\beta\rangle, \quad (11.160)$$

where Latin indexes refer to free sites, while Greek indexes are for the blocks.

- (5) From  $|\psi_{GS}\rangle$ , evaluate the reduced density matrix of the left-enlarged block, by tracing out the right-enlarged block:

$$\rho_l = \text{Tr}_r(|\psi_{GS}\rangle \langle \psi_{GS}|) = \sum_{\alpha, a} \sum_{\alpha', a'} \left( \sum_{\beta, b} \psi_{\alpha, a, b, \beta} \psi_{\alpha', a', b, \beta}^* \right) |\alpha, a\rangle \langle \alpha', a'|. \quad (11.161)$$

- (6) Find an approximate representation of  $\rho_l$ , in terms of a reduced basis of (at most)  $\chi$  elements. This amounts to a truncation of the Hilbert space of the enlarged block, since  $\chi_{\ell+1} = \min(\chi_\ell d, \chi)$ . The truncated Hilbert basis is made up of the first  $\chi_{\ell+1}$  eigenstates of  $\rho_l$ , which correspond to the largest eigenvalues. It is thus possible to construct the  $(\chi_\ell d) \times (\chi_{\ell+1})$  isometric change-of-basis matrix  $\mathcal{O}_{\ell \rightarrow \ell+1}$  containing such  $\chi_{\ell+1}$  selected eigenstates, stored in columns. The core of the DMRG algorithm stands in the renormalization procedure of the enlarged block, which produces, as an output, a truncated enlarged block  $\mathcal{B}_l(\ell+1, \chi_{\ell+1})$ . This consists in the matrix  $\mathcal{O}_{\ell \rightarrow \ell+1}$ , the new block Hamiltonian

$$H'_B = \mathcal{O}_{\ell \rightarrow \ell+1}^\dagger H_E \mathcal{O}_{\ell \rightarrow \ell+1}, \quad (11.162)$$

and all the relevant local operators

$$S'_{\ell+1} = \mathcal{O}_{\ell \rightarrow \ell+1}^\dagger S_{\ell+1} \mathcal{O}_{\ell \rightarrow \ell+1} \quad (11.163)$$

written in the new basis. Notice that an analogous renormalization procedure can be performed for the right-enlarged block, from  $\rho_r = \text{Tr}_l(|\psi_{GS}\rangle \langle \psi_{GS}|)$ .

- (7) Return to point (2), using the truncated enlarged block(s)  $\mathcal{B}_{l/r}(\ell+1, \chi_{\ell+1})$  as the new starting block(s) for the next DMRG iteration. The updated operators  $S'_{\ell+1}$  are necessary for constructing the interaction  $H_{Bs}$  between the rightmost site in the left block and the free site (and conversely for the right block).

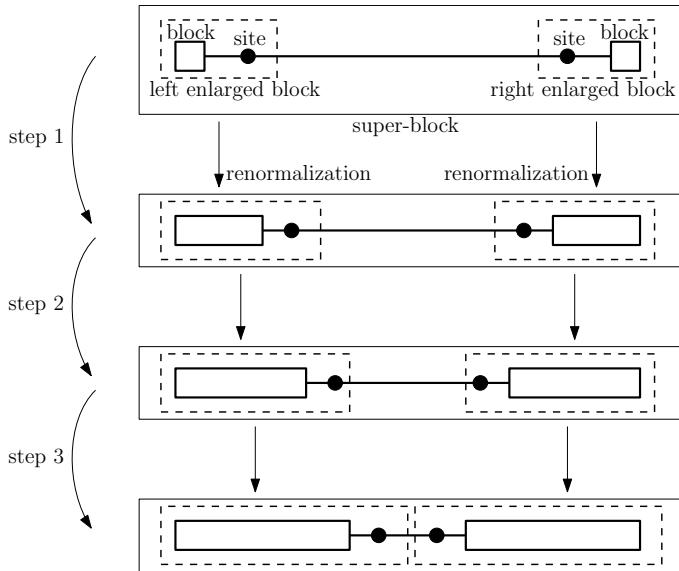


Fig. 11.19 Schematic procedure for the infinite-system DMRG algorithm. Starting from a system block  $\mathcal{B}_l(\ell, \chi_\ell)$  and adding one free site to it, the left-enlarged block  $[\mathcal{B}_l(\ell, \chi_\ell) \ominus \bullet]$  is formed. Analogously, the right-enlarged block is constructed. Then, after having created the super-block  $[\mathcal{B}_l(\ell, \chi_\ell) \ominus \bullet \ominus \bullet \ominus \bullet \ominus \mathcal{B}_r(r, \chi_r)]$ , a renormalization procedure is applied in order to get the new block for the next DMRG iteration.

The full infinite-system DMRG procedure is illustrated in Fig. 11.19. We point out that, following this iterative procedure, it is possible to increase the system size without increasing the number of states which are kept to describe it. At each step, the number of sites in the super-block goes from  $2\ell$  to  $2\ell + 2$ , thus the simulated system size increases by 2 sites. The procedure stops once a satisfactory convergence, e.g., in the ground-state energy of the super-block, is reached. The net output is the (approximate) ground state of a 1D chain in the thermodynamic limit.

### Finite-system DMRG

In most situations, the infinite-system DMRG may not yield accurate results up to the wanted precision. This is the case, e.g., for systems which are not translational invariant: the inhomogeneities in the Hamiltonian cannot be properly accounted for, during the intermediate steps. Moreover, it is often likely that the algorithm gets trapped in a metastable state favoured for small sizes.

For these reasons, White (1993) devised the *finite-system DMRG* in order to overcome these problems. The idea is to stop the infinite-system algorithm at some preselected super-block length  $L$ , which is subsequently kept fixed. In the following DMRG steps, only one block is increased in size, while the other is shrunk, thus keeping the super-block constant in size. The reduced basis transformation is carried out only for the growing block; the shrinking block is simply taken from memory, as it has been built and saved in a previous step of the infinite procedure.

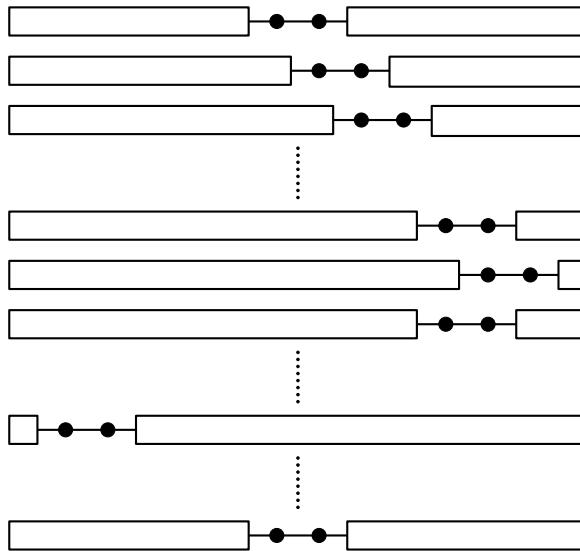


Fig. 11.20 Schematic procedure for the finite-system DMRG algorithm.

In practice, the infinite-system procedure stops when the system is formed by two blocks  $\mathcal{B}_{l,r}(L/2-1, \chi)$  and two free sites, as shown in the first row of Fig. 11.20. Later one starts with the finite-system procedure, so that, after the next step, the system configuration is given by  $[\mathcal{B}_l(L/2, \chi) \ominus \bullet \ominus \bullet \ominus \mathcal{B}_r(L/2-2, \chi)]$ . The left block is constructed by enlarging  $\mathcal{B}_l(L/2-1, \chi)$  with the procedure detailed above. The right block is retrieved from memory: at each previous step, one should have saved the matrices  $\mathcal{O}_{i \rightarrow i+1}$ , the block Hamiltonians  $H_{B,\ell}$  and the interaction operators  $S_\ell$ , for  $\ell = 1, \dots, L/2 - 1$ . Eventually, when the left block has reached the length  $L - 4$ , a right block  $\mathcal{B}_r(1, d)$  with one site has to be constructed from scratch. At this stage, the role of the left and right block are switched, and the free sites start sweeping from right to left. Generally at each sweep the approximation of the ground state for fixed size  $L$  improves. This procedure is illustrated in Fig 11.20.

### 11.7.2 \* DMRG as a variational optimization over the MPS class

We now explicitly show that the finite-system DMRG algorithm is basically equivalent to a variational optimization over the class of MPS wave functions, as originally conceived by Verstraete *et al.* (2004c). To this purpose, it is useful to focus on a given step of the algorithm and write the super-block structure (see Fig. 11.21 for a pictorial representation). After this step and using the notations of the figure, the many-body wave function of the  $L$ -site system can be written according to:

$$|\psi\rangle = \sum_{\alpha_{k-1}, \beta_{k+2}=1}^{\chi} \sum_{a_k, b_{k+1}=1}^d \psi_{\alpha_{k-1}, \beta_{k+2}}^{a_k, b_{k+1}} \underbrace{|\alpha_{k-1}\rangle}_{\text{left block}} \underbrace{|a_k\rangle}_{\text{free sites}} \underbrace{|b_{k+1}\rangle}_{\text{right block}} \underbrace{|\beta_{k+2}\rangle}_{\text{right block}}, \quad (11.164)$$

where  $\psi_{\alpha_{k-1}, \beta_{k+2}}^{a_k, b_{k+1}}$  denotes the super-block matrix representation on the block-site-site-block basis, and it can be seen as a four-rank tensor. In order to recover a MPS-like form for  $|\psi\rangle$ , we need to explicitly write the representation of the states in the left ( $\{\alpha_{k-1}\}$ ) and in the right ( $\{\beta_{k+2}\}$ ) block.

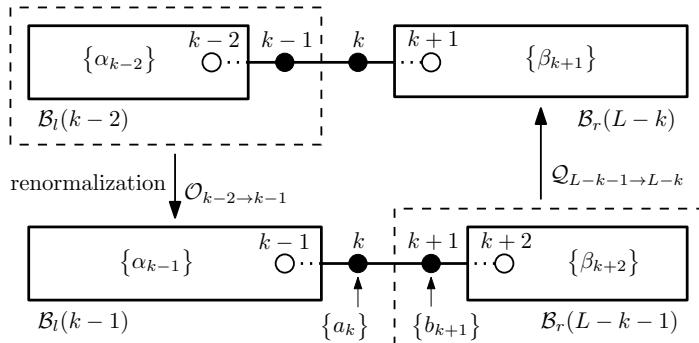


Fig. 11.21 Schematic procedure for the finite-system DMRG algorithm.

Considering first the left block, we can write its basis states  $\{\alpha_{k-1}\}$  in terms of the basis states of the “parent” enlarged block (i.e., the previous left block plus free site – upper line in Fig. 11.21). Namely, we have:

$$|\alpha_{k-1}\rangle = \sum_{\alpha_{k-2}} \sum_{a_{k-1}} \mathcal{O}_{\alpha_{k-2}, \alpha_{k-1}}^{[k-1]a_{k-1}} |\alpha_{k-2}\rangle |a_{k-1}\rangle, \quad (11.165)$$

where  $\mathcal{O}_{\alpha_{k-2}, \alpha_{k-1}}^{[k-1]a_{k-1}}$  are the elements of the  $(d\chi) \times \chi$  change-of-basis truncation matrix  $\mathcal{O}_{k-2 \rightarrow k-1}$ , with  $\alpha_{k-2}$  and  $a_{k-1}$  being the row indexes, and  $\alpha_{k-1}$  the column index. Equivalently this can be seen as a three-rank tensor, of the same type of the  $A^{[k]}$  tensors entering the MPS representation (11.90). It is now clear that this procedure can be iterated, by applying it sequentially to  $|\alpha_{k-2}\rangle$ ,  $|\alpha_{k-3}\rangle$ , and so on, up to  $|\alpha_1\rangle$ , and recursively using the definitions of the change-of-basis matrices  $\mathcal{O}_{j \rightarrow j+1}$ . Eventually one is able to unroll the basis element  $|\alpha_{k-1}\rangle$  on the linear combination

of states belonging to the sites running from 1 to  $k - 1$ :

$$|\alpha_{k-1}\rangle = \sum_{\alpha_1, \dots, \alpha_{k-1}} \sum_{a_1, \dots, a_{k-1}} \mathcal{O}_{\alpha_1}^{[1]a_1} \mathcal{O}_{\alpha_1, \alpha_2}^{[2]a_2} \cdots \mathcal{O}_{\alpha_{k-2}, \alpha_{k-1}}^{[k-1]a_{k-1}} |a_1, a_2, \dots, a_{k-1}\rangle, \quad (11.166)$$

where, by construction, each of the three-rank tensors  $\mathcal{O}^{[j]}$  satisfies the left isometric gauge condition:

$$\sum_{\alpha_{k-1}, a_k} (\mathcal{O}_{\alpha_{k-1}, \alpha'_k}^{[k]a_k})^* \mathcal{O}_{\alpha_{k-1}, \alpha_k}^{[k]a_k} = \delta_{\alpha_k, \alpha'_k}. \quad (11.167)$$

In the same way, it is possible to write the basis states  $\{\beta_{k+2}\}$  of the right block in terms of the states of the “parent” enlarged block:

$$|\beta_{k+2}\rangle = \sum_{\beta_{k+3}} \sum_{b_{k+2}} \mathcal{Q}_{\beta_{k+2}, \beta_{k+3}}^{[k+2]b_{k+2}} |\beta_{k+3}\rangle |b_{k+2}\rangle. \quad (11.168)$$

This time  $\mathcal{Q}_{\beta_{k+2}, \beta_{k+3}}^{[k+2]b_{k+2}}$  are the elements of the  $\chi \times (d\chi)$  change-of-basis truncation matrix  $\mathcal{Q}_{L-k-1 \rightarrow L-k}$ , with  $\beta_{k-2}$  and  $b_{k+2}$  being the column indexes, and  $\beta_{k+3}$  the row index. At the end one recursively unrolls them on the linear combination of states belonging to the sites running from  $k + 2$  to  $L$ :

$$|\beta_{k+2}\rangle = \sum_{\beta_{k+2}, \dots, \beta_L} \sum_{b_{k+2}, \dots, b_L} \mathcal{Q}_{\beta_{k+2}, \beta_{k+3}}^{[k+2]b_{k+2}} \mathcal{Q}_{\beta_{k+3}, \beta_{k+4}}^{[k+3]b_{k+3}} \cdots \mathcal{Q}_{\beta_L}^{[L]b_L} |b_{k+2}, b_{k+3}, \dots, b_L\rangle, \quad (11.169)$$

where by construction, analogously as before, these matrices satisfy the right isometric gauge

$$\sum_{\beta_{k+1}, b_k} (\mathcal{Q}_{\beta_k, \beta_{k+1}}^{[k]b_k})^* \mathcal{Q}_{\beta'_k, \beta_{k+1}}^{[k]b_k} = \delta_{\beta_k, \beta'_k}. \quad (11.170)$$

In conclusion, inserting the expressions (11.166) for  $|\alpha_{k-1}\rangle$ , and (11.169) for  $|\beta_{k+2}\rangle$  in the representation (11.164) of the many-body wave function, we obtain:

$$|\psi\rangle = \sum_{\vec{\alpha}, \vec{\beta}} \sum_{\vec{a}, \vec{b}} \mathcal{O}_{\alpha_1}^{[1]a_1} \mathcal{O}_{\alpha_1, \alpha_2}^{[2]a_2} \cdots \mathcal{O}_{\alpha_{k-2}, \alpha_{k-1}}^{[k-1]a_{k-1}} \psi_{\alpha_{k-1}, \beta_{k+2}}^{a_k b_{k+1}} \times \\ \times \mathcal{Q}_{\beta_{k+2}, \beta_{k+3}}^{[k+2]b_{k+2}} \cdots \mathcal{Q}_{\beta_L}^{[L]b_L} |a_1, \dots, a_k, b_{k+1}, \dots, b_L\rangle. \quad (11.171)$$

Graphically, this is represented in Fig. 11.22.

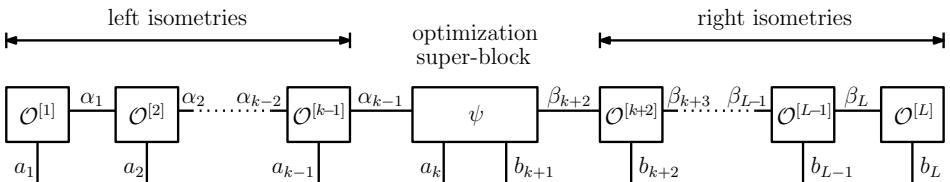


Fig. 11.22 Diagrammatic representation of the many-body wave function at a given DMRG step, which optimizes the central four-rank tensor. The latter represents the only formal difference with the variational optimization procedure over a standard MPS wave function.

It is important to stress that, due to the presence of the four-rank tensor  $\psi_{\alpha_{k-1}, \beta_{k+2}}^{a_k, b_{k+1}}$ , which is attached to the two physical sites  $k$  and  $k + 1$ , the state in Eq. (11.171) is not exactly in a MPS form. However, since all the three-rank tensors  $\mathcal{O}^{[k]}$  are in the isometric gauge, a SVD applied to the tensor  $\psi$  with respect to the indexes  $(\alpha_{k-1}, a_k)$  and  $(\beta_{k+2}, b_{k+1})$  is able to cast the state as a MPS. The finite-system DMRG is thus equivalent to an optimization procedure over the pseudo-MPS Ansatz of the form (11.171). We remark that there would be perfect equivalence between the variational MPS optimization and a modified DMRG procedure, in which the two blocks are connected by a single free site to form the superblock.

## 11.8 Time evolution of matrix product states

Up to now, the discussion has been focused only on the ground (or low-lying) states of many-body systems. In that case, the area-law considerations on the scaling of the bipartite entanglement entropy ensure that the MPS representation of the wave function is optimal, and thus is able to exponentially reduce the number of required variational parameters in the ground-state manifold. However, the techniques described here can be also adapted to simulate states at thermal equilibrium at finite temperature, as well as the time evolution of systems governed by local Hamiltonians (or even by local Liouvillians). This section is devoted to a description of all these approaches.

We first discuss the most paradigmatic one, dealing with the real-time evolution, from which all the others follow as a simple extension. For the sake of simplicity, we focus on the simplest conceptual situation. We start from an initial MPS wave function  $|\psi(0)\rangle$ , of the form in Eq. (11.90), and study its dynamics following the unitary time evolution operator  $U(t) = \exp(-iHt)$ , where  $H$  is a given local Hamiltonian:<sup>12</sup>

$$|\psi(t)\rangle_{\text{MPS}} = e^{-iHt} |\psi(0)\rangle_{\text{MPS}}. \quad (11.172)$$

For the sake of simplicity, we also suppose that  $H$  is time independent, and can be written as sum of nearest-neighbour interaction terms, such that

$$H = H_{\text{even}} + H_{\text{odd}}, \quad (11.173)$$

with

$$H_{\text{even}} = \sum_{j \text{ even}} h_{j,j+1}, \quad H_{\text{odd}} = \sum_{j \text{ odd}} h_{j,j+1}. \quad (11.174)$$

In this way, the terms contained in  $H_{\text{even}}$  only act between sites 2-3, 4-5, and so on, while the terms contained in  $H_{\text{odd}}$  only act between sites 1-2, 3-4, and so on. As will be clear below, our protocol can be easily extended to the more general case where  $H = H(t)$  depends explicitly on time, and thus the unitary evolution is ruled by the time-ordered operator:  $U_T(t) = \text{Texp}\left[-i \int_0^t H(t') dt'\right]$ .

---

<sup>12</sup>For the sake of simplicity, throughout this chapter we are working in units of  $\hbar = 1$  and  $k_B = 1$ .

The decomposition of the Hamiltonian (11.173) as a sum of two operators, grouping even and odd links respectively, ensures that each term contained in  $H_{\text{even}}$  commutes with all the others, and the same occurs for  $H_{\text{odd}}$ . Therefore the unitary evolution of each single part,  $U_{\text{even}}(t) = e^{-iH_{\text{even}}t}$  and  $U_{\text{odd}}(t) = e^{-iH_{\text{odd}}t}$ , can be straightforwardly written as the product of evolutions of single links. As explained below, each of these single-link evolutions can be performed by a local update of the MPS wave function. The difficulty of the many-body problem (11.172) has been thus moved to the fact that  $[H_{\text{even}}, H_{\text{odd}}] \neq 0$ , from which it follows that

$$e^{-iHt} \neq e^{-iH_{\text{even}}t} e^{-iH_{\text{odd}}t}. \quad (11.175)$$

In order to split the full evolution operator, it is useful to employ the Baker-Campbell-Hausdorff formula (A.101), which can be recast according to:

$$e^A e^B = e^Z, \quad \text{with } Z = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] + [B, [B, A]]) + \dots \quad (11.176)$$

Therefore by repeatedly applying it, one gets

$$e^{\tau(A+B)} = \prod_{j=1}^k e^{c_j \tau A} e^{d_j \tau B} + O(\tau^{n+1}), \quad (11.177)$$

where both  $k \in \mathbb{N}$  and  $c_j, d_j \in \mathbb{R}$  depend on the order  $n$  of the expansion. This expansion is usually referred to as the *Suzuki-Trotter decomposition*. Specifically we have, up to fourth order:

$$n = 1 \Rightarrow k = 1; \quad c_1 = d_1 = 1 \quad (11.178)$$

$$n = 2 \Rightarrow k = 2; \quad c_1 = c_2 = 1/2, \quad d_1 = 1, \quad d_2 = 0 \quad (11.179)$$

$$n = 4 \Rightarrow k = 4; \quad c_1 = c_4 = \frac{1}{2(2 - 2^{1/3})}, \quad c_2 = c_3 = (1 - 2^{1/3})c_1,$$

$$d_1 = d_3 = 2c_1, \quad d_2 = -2^{4/3}c_1, \quad d_4 = 0. \quad (11.180)$$

Notice that  $\sum_{j=1}^k c_j = \sum_{j=1}^k d_j = 1$ . The choice of the  $c_j$  and  $d_j$  parameters is not unique, but for even  $k$  values can be always taken in a symmetric way, such that  $c_j = c_{k-j+1}$ , ( $j = 1, \dots, k/2$ ) and  $d_j = d_{k-j}$ , ( $j = 1, \dots, k/2 - 1$ ), with  $d_k = 0$ .

If we now specialize the expansion (11.177) to the case of the unitary evolution operator  $U(t)$ , where  $A = H_{\text{even}}$  and  $B = H_{\text{odd}}$ , we obtain:

$$U(t) = e^{-iHt} \approx \left( \prod_{j=1}^k e^{-ic_j H_{\text{even}}dt} e^{-id_j H_{\text{odd}}dt} \right)^{t/dt}. \quad (11.181)$$

This expansion approximates  $U(t)$  as a product of many exponentials of operators that are composed of terms commuting each other. As anticipated above, each of these terms contains only a single-link evolution, and can be performed by a local update of the MPS wave function.

Let us focus, e.g., on the term  $e^{-ic_j H_{k,k+1}dt}$ , where  $k$  is a generic even site in the chain. The way in which this can be applied to a MPS is depicted in Fig. 11.23. Explicitly, the protocol works according to the following three steps:

- (1) One first applies the two site operator  $U_{k,k+1}(dt) = e^{-ic_j H_{k,k+1} dt}$  to the two-tensor structure  $A^{[k]} \times A^{[k+1]}$  in the MPS. This amounts to perform the following contraction:

$$\sum_{j_k, j_{k+1}=1}^d [U_{k,k+1}(dt)]_{j_k, j_{k+1}}^{i_k, i_{k+1}} A_{\alpha, \beta}^{[k] j_k} A_{\beta, \gamma}^{[k+1] j_{k+1}} M_{\alpha, \gamma}^{[k, k+1] i_k, i_{k+1}}. \quad (11.182)$$

- (2) Since the resulting four-rank tensor  $M^{[k, k+1]}$  is not locally a MPS, we need to split it into two three-rank tensors  $\tilde{A}^{[k]}$  and  $\tilde{A}^{[k+1]}$ , thus coming back to the original MPS structure of the wave function. This can be done by applying a SVD with respect to the indexes  $(\alpha, i_k)$  and  $(\beta, i_{k+1})$ , in a way similar to what has been described at the end of Sec. 11.7.2.
- (3) The splitting into the tensors  $\tilde{A}_{\alpha, \beta}^{[k] i_k}$  and  $\tilde{A}_{\beta, \gamma}^{[k+1] i_{k+1}}$  can be done simply by renormalizing the internal bond link  $\beta$ , keeping only the largest  $\chi'$  singular values (out of the  $d \times \chi$  total singular values) in the previous SVD. We stress that this cutting procedure works properly in the isometric gauge with respect to the working site  $k$ . The reason is due to the fact that the truncation of  $\beta$  can be done optimally only if the bipartite state over that link can be written in a Schmidt decomposition: indeed, in this gauge, the truncation optimizes the overlap between the untruncated and the truncated MPS.

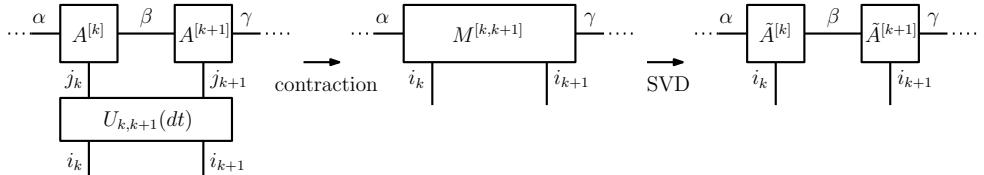


Fig. 11.23 Diagrammatic representation of a single step of the TEBD algorithm, which applies the operator  $U_{k,k+1}(dt) = e^{-ic_j H_{k,k+1} dt}$  locally on the MPS wave function.

It is thus possible to get  $|\psi(t)\rangle_{\text{MPS}}$  by iterating the above procedure through all the links of the lattice, sweeping back and forth as discussed in Sec. 11.7 for the variational ground-state search. In practice, following the Suzuki–Trotter expansion in Eq. (11.181), at each sweep one applies either  $U_{\text{even}}(dt)$  or  $U_{\text{odd}}(dt)$  sequentially. In this way, the  $n$ -th order evaluation of  $U(t)$  requires  $2k \times t/dt$  sweeps. In order to preserve the isometric gauge, during each sweep one alternately performs the protocol of Fig. 11.23 (to apply the evolution operator  $e^{-ic_j h_{k,k+1} dt}$  on the tensors  $A^{[k]} A^{[k+1]}$ ), followed by a SVD on the tensor to the right ( $A^{[k+1]}$ , for left-to-right sweeps) or to the left ( $A^{[k]}$ , for right-to-left sweeps) of the evolved link. The full procedure has been named *time-evolving block decimation* (TEBD) algorithm, by his inventor (Vidal, 2003, 2004).

It is important to stress that the TEBD procedure is not exempt of approximations. Indeed, in view of the area-law considerations of Sec. 11.4, it is not at all

guaranteed that, after a generic real-time evolution, the evolved state  $|\psi(t)\rangle$  can be faithfully expressed as a MPS. Specifically, we spotlight two main sources of errors:

- The first one comes from the Suzuki–Trotter decomposition (11.181), which is basically an approximate way to write the exponential of the sum of operators in terms of the product of exponentials. To improve its accuracy, one should either increase  $k$  or decrease the discretization time  $dt$ , which would both require an increasing number of sweeps (and thus of computational resources). Besides that, the Suzuki–Trotter error is also well understood analytically; this potentially allows for accurate extrapolations to account for it.
- The most dramatic error however comes from the renormalization, at each step, of the dimension  $d \times \chi \rightarrow \chi'$  of the bond link connecting the two tensors  $A^{[k]}$  and  $A^{[k+1]}$ , out of the tensor  $M^{[k,k+1]}$ . Of course the error decreases as long as  $\chi'$  increases, and vanishes for  $\chi' = d \times \chi$ . This generally implies that the dimension of the time evolved MPS would increase with the time.

Coming back to this last point, by employing the TEBD procedure for a generic time evolution of the type in Eq. (11.172), one would obtain that the resulting state  $|\psi(t)\rangle_{\text{MPS}}$  has a bond-link dimension that necessarily increases exponentially in time. There are however certain evolutions for which such dramatic growth does not occur or is highly reduced, e.g., if the evolution is adiabatic as in quantum annealing procedures (see Sec. 3.13).

Finally we point out that the TEBD algorithm can be applied as well for imaginary time evolutions:

$$|\psi(\tau)\rangle = e^{-\tau H} |\psi(0)\rangle, \quad (11.183)$$

after the replacement  $t \rightarrow \tau = -it$ . In the limit of infinite imaginary time,  $\tau \rightarrow \infty$ , and provided the initial state  $|\psi(0)\rangle$  has a finite overlap with the ground state, one will collapse into the latter. This procedure provides an alternative (non variational) way to find the ground state, during which the increase of the bond link discussed above does not occur.

### 11.8.1 Finite-temperature calculations

It is possible to adapt the TEBD algorithm in order to simulate equilibrium Boltzmann states at finite temperature  $T = 1/\beta$ :

$$\rho_\beta = \frac{1}{Z(\beta)} e^{-\beta H}, \quad (11.184)$$

where  $Z(\beta) = \text{Tr}_P[e^{-\beta H}]$  is the partition function, and the subscript  $P$  denotes the physical space. This can be achieved through a purification of the mixed state  $\rho_\beta$  into an enlarged space, where the ancillary state space can be taken as a copy of the original one (see Sec. 2.8), and then by a partial time evolution on such pure

state. Finite-temperature density matrices on a chain can be thus expressed as pure states on a ladder, as shown in Fig. 11.24. To this purpose, let us first write  $\rho_\beta$  as

$$\rho_\beta = \frac{1}{Z(\beta)} e^{-\beta H/2} \mathbb{I} e^{-\beta H/2}, \quad (11.185)$$

where the identity matrix  $\mathbb{I}$  is nothing but the unnormalized equilibrium state at infinite temperature:  $\mathbb{I} = Z(0)\rho_0$ . Suppose now that we know the purification of  $\rho_0$  in terms of a given MPS  $|\psi_{\beta=0}\rangle_{PQ}$ , where  $P$  denotes the physical space and  $Q$  the ancillary space. This means that  $\rho_0 = \text{Tr}_Q[|\psi_0\rangle\langle\psi_0|]$ . Therefore

$$\rho_\beta = \frac{Z(0)}{Z(\beta)} e^{-\beta H/2} \left\{ \text{Tr}_Q[|\psi_0\rangle\langle\psi_0|] \right\} e^{-\beta H/2} = \frac{Z(0)}{Z(\beta)} \text{Tr}_Q[e^{-\beta H/2} |\psi_0\rangle\langle\psi_0| e^{-\beta H/2}],$$

where the second equality follows from the fact that the Hamiltonian  $H$  acts only on the physical space. It is thus sufficient to perform an imaginary time evolution

$$|\psi_\beta\rangle = e^{-\beta H/2} |\psi_0\rangle \quad (11.186)$$

up to the imaginary time  $\tau = \beta/2$ , where  $H$  acts only on the real space  $P$ . In this way, we can recover the finite-temperature state

$$\rho_\beta = \frac{Z(0)}{Z(\beta)} \text{Tr}_Q[|\psi_\beta\rangle\langle\psi_\beta|] \quad (11.187)$$

and get the corresponding expectation value of any observable  $\mathcal{O}$  on it:

$$\langle \mathcal{O} \rangle_\beta = \text{Tr}_P[\mathcal{O} \rho_\beta] = \frac{Z(0)}{Z(\beta)} \text{Tr}_P[\mathcal{O} \text{Tr}_Q[|\psi_\beta\rangle\langle\psi_\beta|]] = \frac{Z(0)}{Z(\beta)} \langle \psi_\beta | \mathcal{O}_P \otimes \mathbb{I}_Q | \psi_\beta \rangle. \quad (11.188)$$

This can be written in a simpler form, after observing that

$$\langle \mathbb{I} \rangle_\beta = \frac{Z(0)}{Z(\beta)} \langle \psi_\beta | \mathbb{I}_P \otimes \mathbb{I}_Q | \psi_\beta \rangle \Rightarrow \frac{Z(0)}{Z(\beta)} = \frac{1}{\langle \psi_\beta | \psi_\beta \rangle}, \quad (11.189)$$

and therefore

$$\langle \mathcal{O} \rangle_\beta = \frac{\langle \psi_\beta | \mathcal{O}_P \otimes \mathbb{I}_Q | \psi_\beta \rangle}{\langle \psi_\beta | \psi_\beta \rangle}. \quad (11.190)$$

Note that it is even possible to carry out the real-time evolution of the purified state  $|\psi_\beta\rangle$ , such to treat time-dependent problems at finite temperature:

$$\langle \mathcal{O} \rangle_\beta(t) = \frac{\langle \psi_\beta(t) | \mathcal{O}_P \otimes \mathbb{I}_Q | \psi_\beta(t) \rangle}{\langle \psi_\beta(t) | \psi_\beta(t) \rangle}, \quad \text{with } |\psi_\beta(t)\rangle = e^{-iHt} |\psi_\beta\rangle. \quad (11.191)$$

The last issue concerns the way in which we purify the initial state at infinite temperature  $\rho_0$ . Since  $\rho_0 = (\mathbb{I}_j/d)^{\otimes L}$ , the purified MPS in the  $PQ$  space is trivially a product state with bond link  $\chi = 1$ :  $|\psi_0\rangle_{PQ} = |\psi_0\rangle_1 |\psi_0\rangle_2 \cdots |\psi_0\rangle_L$ . A single state on site  $j$  can now be purified as a state living on a rung  $j$  of the ladder, where the physical and the auxiliary leg are linked by a maximally entangled state:

$$|\psi_0\rangle_j = \frac{1}{\sqrt{d}} \sum_{\sigma=1}^d |\sigma\rangle_P |\sigma\rangle_Q. \quad (11.192)$$

The idea of the ladder is depicted in Fig. 11.24.

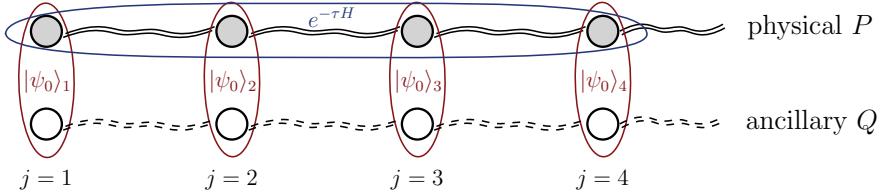


Fig. 11.24 Diagrammatic representation of finite-temperature simulations: one sets up a ladder, where the physical sites (upper leg) have the same Hilbert space as for the ancillary sites (lower leg). Equivalent sites on the physical and auxiliary leg are first linked by maximally entangled states (red). To reach the inverse temperature  $\beta$ , an imaginary time evolution of the upper leg (blue) is carried out using the TEBD protocol, up to time  $\tau = \beta/2$ .

### 11.8.2 Mixed-state time evolution

The TEBD algorithm can be also adapted to the more general case of non-unitary time evolutions, as firstly devised by Verstraete *et al.* (2004b); Zwolak and Vidal (2004). Here we discuss the evolution of a mixed state according to the GKLS master equation of Eq. (7.169). There is a straightforward way to manipulate general mixed states, which closely resembles the purification that we saw above. This goes under the formalism of the vectorization of superoperators. Let us fix a basis  $\{|i\rangle\}_{i=1,\dots,d}$  for our working Hilbert space  $\mathcal{H}$ . The idea is to map the space  $\mathcal{B}(\mathcal{H})$  of the linear operators acting on  $\mathcal{H}$  into the space  $\mathcal{H} \otimes \mathcal{H}$ , which is spanned by a basis  $\{|i,j\rangle\}_{i,j=1,\dots,d}$ . Every linear operator  $A \in \mathcal{B}(\mathcal{H})$  can be thus associated to a vector (also named super-ket)  $|A\rangle\langle\cdot| \in \mathcal{H} \otimes \mathcal{H}$  in the superoperator space:

$$A = \sum_{i,j} A_{ij} |i\rangle\langle j| \longrightarrow |A\rangle\langle\cdot| \equiv \sum_{i,j} A_{ij} |i\rangle\langle j| = \sum_{i,j} A_{ij} |i,j\rangle\langle j|. \quad (11.193)$$

The scalar product is defined by

$$\langle\langle A | B \rangle\rangle = \sum_{i,j} A_{ij}^* B_{ij} = \text{Tr}(A^\dagger B). \quad (11.194)$$

We introduce the identity superoperator  $\mathbb{I} \in \mathcal{B}(\mathcal{H})$  as

$$|\mathbb{I}\rangle\langle\cdot| \equiv \sum_i |i\rangle\langle i| = \sum_i |i,i\rangle\langle i|, \quad (11.195)$$

from which the following useful relations for any  $A, B, C \in \mathcal{B}(\mathcal{H})$  follow:

$$|A\rangle\langle\cdot| = (A \otimes \mathbb{I})|\mathbb{I}\rangle\langle\cdot| = (\mathbb{I} \otimes A^T)|\mathbb{I}\rangle\langle\cdot| \quad (11.196)$$

$$\begin{aligned} |AB\rangle\langle\cdot| &= (AB \otimes \mathbb{I})|\mathbb{I}\rangle\langle\cdot| = (A \otimes \mathbb{I})(B \otimes \mathbb{I})|\mathbb{I}\rangle\langle\cdot| \\ &= (A \otimes \mathbb{I})|B\rangle\langle\cdot| = (\mathbb{I} \otimes B^T)|A\rangle\langle\cdot|, \end{aligned} \quad (11.197)$$

$$|ABC\rangle\langle\cdot| = (A \otimes C^T)|B\rangle\langle\cdot|. \quad (11.198)$$

**Exercise 11.5** Prove Eqs. (11.196), (11.197), and (11.198).

If we now come back to the GKLS master equation (7.169) and write it in a vectorized form, we formally get:

$$\frac{d|\rho\rangle\langle\cdot|}{dt} = \mathbb{L}|\rho\rangle\langle\cdot|, \quad (11.199)$$

where  $\mathbb{L}$  is called the *Liouville superoperator* (or Liouvillian) of the Markovian dynamics, and acts over the vectorized density matrix  $|\rho\rangle\langle\rho|$ . With the help of the relations (11.197)-(11.198), it is not difficult to prove that

$$\mathbb{L}|\rho\rangle\langle\rho| = \left\{ \left[ -iH - \sum_k L_k^\dagger L_k \right] \otimes \mathbb{I} + \mathbb{I} \otimes \left[ iH^T - \sum_k L_k^T L_k^* \right] + \sum_k L_k \otimes L_k^* \right\} |\rho\rangle\langle\rho|. \quad (11.200)$$

Assuming now that, without loss of generality, the Liouvillian  $\mathbb{L}$  does not explicitly depend on time, and that it can be decomposed as a sum of terms involving at most nearest-neighbour operators:

$$\mathbb{L} = \sum_j \mathbb{L}_{j,j+1}, \quad (11.201)$$

we can formally write the time-evolved state as

$$|\rho(t)\rangle\langle\rho(0)| = e^{\mathbb{L}t} |\rho(0)\rangle\langle\rho(0)| = \exp \left[ \sum_j \mathbb{L}_{j,j+1} t \right] |\rho(0)\rangle\langle\rho(0)|. \quad (11.202)$$

At this point, it is thus straightforward to apply a Suzuki–Trotter decomposition on the above exponential, in order to split it into the product of many local exponentials which can be suitably treated with the TEBD method.

The last ingredient that we need is a MPS-like representation of the state  $\rho$ , which generalizes the form in Eq. (11.90) to operators, and goes under the name of *matrix product operator* (MPO). Using the standard vectorized form, this can be expressed in a compact form by

$$|\rho\rangle\langle\rho| = \sum_{\vec{i}, \vec{j}} \text{Tr}[B^{[1]i_1, j_1} B^{[2]i_2, j_2} \dots B^{[L]i_L, j_L}] |i_1, j_1\rangle\langle i_1, j_1| \otimes |i_2, j_2\rangle\langle i_2, j_2| \otimes \dots \otimes |i_L, j_L\rangle\langle i_L, j_L|, \quad (11.203)$$

where the trace has to be always intended as a contraction over all the implicit greek indexes  $\beta_j = 1, \dots, \chi$  (with  $j = 1, \dots, L$  for PBC, and  $j = 1, \dots, L-1$  for OBC) standing for the horizontal bond links. Graphically, this is represented as in Fig. 11.25. Note that the structure is basically identical to that of a MPS, with the only difference that there are now two vertical legs per tensor, instead of a single one, representing the “bra” and the “ket”.

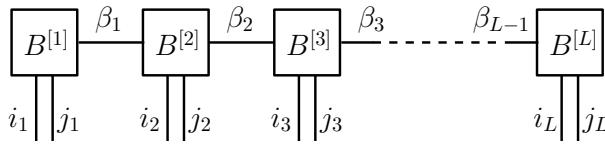


Fig. 11.25 Graphical representation of a MPO for  $L$  physical sites, with OBC.

The normalization condition  $\text{Tr}[\rho] = \langle\langle \mathbb{I} | \rho \rangle\rangle$  can be ensured with a local constraint on the MPO structure of Eq (11.203). Indeed we have:

$$\langle\langle \mathbb{I} | \rho \rangle\rangle = \sum_{\vec{j}} \sum_{\vec{\beta}} B_{\beta_1}^{[1]j_1, j_1} B_{\beta_1, \beta_2}^{[2]j_2, j_2} \dots B_{\beta_{L-1}}^{[L]j_L, j_L}, \quad (11.204)$$

which can be satisfied, during the sweeping procedure along the chain, with a simple sequential book keeping. Specifically, the following vectors  $v^{[k]}$  (when going from left to right) and  $\omega^{[k]}$  (when going from right to left) need to be stored in memory (for  $k = 2, \dots, L - 1$ ):

$$v_{\beta_k}^{[k]} = \sum_{i_k} \sum_{\beta_{k-1}} v_{\beta_{k-1}}^{[k-1]} B_{\beta_{k-1}, \beta_k}^{[k] i_k, i_k}, \quad (11.205)$$

$$\omega_{\beta_{k-1}}^{[k]} = \sum_{i_k} \sum_{\beta_k} B_{\beta_{k-1}, \beta_k}^{[k] i_k, i_k} \omega_{\beta_k}^{[k+1]}, \quad (11.206)$$

with the boundary conditions  $v_{\beta}^{[1]} = \sum_{i_1} B_{\beta}^{[1] i_1, i_1}$  and  $\omega_{\beta}^{[L]} = \sum_{i_L} B_{\beta}^{[L] i_L, i_L}$ . Therefore, at the  $k$ -th step, the normalization (11.204) translates into

$$\text{Tr}[\rho] = \langle\langle \mathbb{I} | \rho \rangle\rangle = \sum_{\beta} v_{\beta}^{[k]} \omega_{\beta}^{[k+1]}, \quad (11.207)$$

and can be enforced by simply dividing the elements of the tensor  $B^{[k]}$  by  $\text{Tr}[\rho]$ .

Unfortunately, the complete positivity of  $\rho(t)$ , which is formally preserved under the time evolution of Eq. (11.199), is not easy to enforce in a TEBD algorithm, since it is a non-local condition. Indeed, due to the approximations induced by TEBD (especially the truncation error discussed above), we cannot guarantee that the time evolved MPO state will satisfy such constraint. The most convenient way to control this positivity loss is to keep track of the unnormalized value of  $\text{Tr}[\rho]$  during the sweeping. When this acquires an imaginary part that becomes non negligible, it means that the MPO representation of the super-ket  $|\rho\rangle\rangle$  is deviating from the manifold of quantum states, and thus the bond link  $\chi$  has to be increased.

Finally we mention that observables onto a given MPO state can be easily calculated using the notion of scalar product in the superoperator space:  $\text{Tr}[A\rho] = \langle\langle A | \rho \rangle\rangle$ . The needed contractions can be speeded up by cleverly using the vectors  $v^{[k]}$  and  $\omega^{[k]}$  defined above, in a way similar to what has been done in Eq. (11.207) in order to fix the normalization condition

## 11.9 \* General tensor-network structures

Up to now, we have seen that the MPS Ansatz works very well for simulating states that satisfy an area-law scaling of the bipartite entanglement for 1D systems. The ground states of local Hamiltonians belong to this category. The purpose of this section is to go one step forward and introduce more complicated tensor-network structures that are able to formulate more general Ansatzes for different situations, such as for higher dimensional systems which satisfy again an area law, or for critical systems which admit a logarithmic violation of the area law.

### 11.9.1 \* Projected entangled pair states

Let us focus on the two dimensional (2D) scenario. A naive idea to simulate 2D systems would be to use again MPS wave functions, after choosing an artificial

1D ordering of the sites in the lattice. While this approach has been numerically applied in several occasions, it cannot physically reproduce the area-law scaling of entanglement in two dimensions for increasing the system size, since the entanglement entropy would increase as the square root of the number of sites (see Sec. 11.4). More cleverly, Verstraete and Cirac (2004a) proposed to generalize the mapping of Eq. (11.87) to a higher dimensional construction, giving rise to the so-called *projected entangled pair states* (PEPS). Here we will focus on a square lattice, nonetheless this construction can be safely adapted to more general lattice structures, with a generic number of  $z$  nearest neighbours per site.

We first define a many-body system living on a square lattice of  $L \times L$  physical sites. We then associate each site with  $z = 4$  ancillary subsystems with dimension  $\chi$  ( $\chi$  will denote the bond link of the PEPS wave function that we are going to construct). These are placed in a maximally entangled state  $|\omega_\chi\rangle$  with the corresponding subsystem of each of the adjacent sites, as in Eq. (11.85). Then for each physical site that can be located on the lattice by the coordinates  $(x, y)$ , we define a linear map  $M^{[x,y]} : \mathbb{C}^\chi \otimes \mathbb{C}^\chi \otimes \mathbb{C}^\chi \otimes \mathbb{C}^\chi \rightarrow \mathbb{C}^d$ , which generalizes Eq. (11.87) to the 2D case. The action of  $M^{[x,y]}$  is to project the states of the four auxiliary sites onto the states of the physical site. This can be written as

$$M^{[x,y]} = \sum_{i=1}^d \sum_{\alpha,\beta,\gamma,\delta=1}^{\chi} A_{\alpha,\beta,\gamma,\delta}^{[x,y]i} |i\rangle \langle \alpha, \beta, \gamma, \delta|, \quad (11.208)$$

where  $A_{\alpha,\beta,\gamma,\delta}^{[x,y]i}$  denotes a five-rank tensor in which the Roman index  $i$  runs over the  $d$  states  $\{|i\rangle\}_{i=1,\dots,d}$  of the local Hilbert basis of site  $(x, y)$ , while each of the four Greek indexes runs over the  $\chi$  states  $\{|\alpha\rangle\}_{i=1,\dots,\chi}$  of the Hilbert basis of one ancillary system. The construction is illustrated in Fig. 11.26. Similarly to the 1D case, it is possible to show that PEPS are able to faithfully approximate ground states of local Hamiltonians living on a lattice with a coordination number  $z$ . Moreover as for MPS states, such tensor-network Ansatz necessarily forces correlation functions to be either constant or to decay exponentially with the distance, while it cannot support decays over larger scales (power-law scalings).

### Examples of PEPS wave functions

A PEPS with bond link  $\chi = 1$  trivially represents a product state. The simplest non-trivial example is probably represented by the GHZ-state in 2D, which is formally given by the same expression of Eq. (11.93). In order to build up the PEPS representation of such state, it is sufficient to consider the two following (translationally invariant) tensors ( $i = 0, 1$ ) with a bond link  $\chi = 2$ :

$$A_{\alpha\beta\gamma\delta}^i = \begin{cases} i & \text{if } \alpha = \beta = \gamma = \delta = i \\ 0 & \text{otherwise.} \end{cases} \quad (11.209)$$

Indeed these realize the mapping operation  $M = |0\rangle\langle 0000| + |1\rangle\langle 1111|$  which selects either the physical state  $|00\dots 0\rangle$  or  $|11\dots 1\rangle$ .

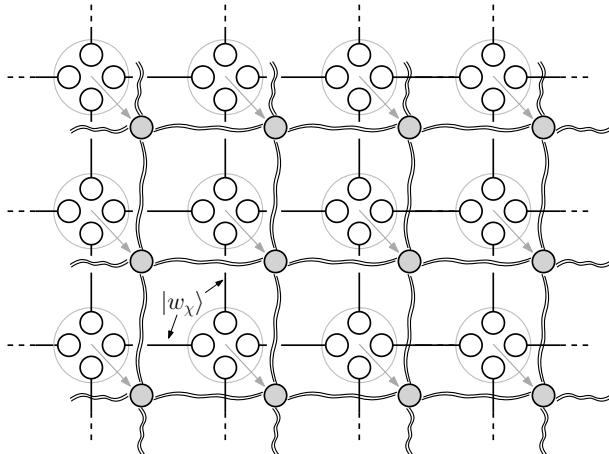


Fig. 11.26 Schematic representation of how to derive a PEPS representation for a 2D system on a square lattice with  $L^2$  sites (grey circles), from the mapping of a pairwise entangled state of an ancillary square lattice with  $4L^2$  sites (white circles).

Another simple generalization of the 1D case is the two-dimensional AKLT state of spin-2 particles. This is constructed by considering a bond link  $\chi = 2$ , thus associating a spin-1/2 to each site of the auxiliary lattice. Each pair of neighbouring sites in such space (note that we now have  $z = 4$  bonds per site) is placed in a maximally entangled singlet state  $|w_2\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ . The auxiliary space is finally mapped into the physical spin-2 lattice, by symmetrizing it in such a way as to keep the five fully symmetric states, out of the  $2^z = 16$  possible states of the four auxiliary spin-1/2 particles. This construction can be equally performed on a generic lattice of  $z$  nearest neighbours, by projecting the spin-1/2 singlet bonds of the auxiliary lattice into the  $(2z+1)$  completely symmetrized states of the physical spin- $z/2$  particles, thus creating a spin- $z/2$  AKLT state.

#### Contraction of PEPS wave functions

Despite the apparent simplicity of the higher dimensional tensor-network structures that can be formed using the reasoning depicted above, it is not so easy to extrapolate physical information out of them. Indeed a crucial ingredient required to evaluate, for example, any expectation value of local observables is the ability to perform contractions. Without loss of generality, below we focus on the calculation of the normalization  ${}_{\text{PEPS}}\langle\psi|\psi\rangle_{\text{PEPS}}$ , which is obtained by sandwiching the corresponding ‘ket’ and ‘bra’ tensor networks of  $|\psi\rangle_{\text{PEPS}}$ . As for the 1D MPS (see Sec. 11.6.1), one can use the transfer matrix formalism, by defining

$$\mathcal{E}_O^{[x,y]} = \sum_{i,j} (A^{[x,y]i}) \otimes (A^{[x,y]j})^* \langle j|O|i\rangle, \quad (11.210)$$

according to the diagrammatic representation of Fig. 11.27.

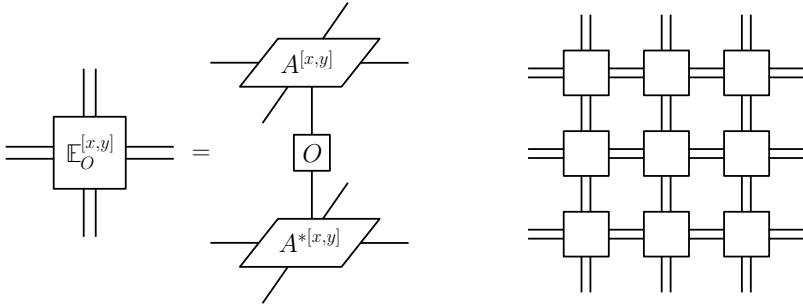


Fig. 11.27 Schematic representation of a PEPS transfer matrix (left panel), and of the brute-force contraction scheme over the whole lattice (right panel).

Unfortunately, differently from the case of MPS, here there is no 1D structure which can be used to reduce the contraction problem of Fig. 11.27 (right panel) to a chain of matrix multiplications. In fact, for any possible kind of exact contraction scheme, the cluster of tensors that needs to be stored during the contraction (e.g., a rectangle) will at some point have a boundary of a length that increases as a power law of the system size. This means that an intermediate object with a number of indices that is roughly proportional to  $L$  is being formed, and thus an exponential number of configurations in the Hilbert space would be generated. This immediately shows that it is impossible to exactly contract a PEPS efficiently, and some approximate contraction scheme needs to be invoked. We now describe the simplest of such schemes, which is based on a MPS-like approach.

Let us consider the contraction of a finite  $L \times L$  PEPS with OBC, as in Fig. 11.28. We take the first two rows and block the two tensors in each row into a new tensor  $\mathbb{F}$ , with horizontal bond dimension  $\chi^4$ , thus reducing the number of rows by one. Since the bond dimension of the first row (which can be seen as an effective MPS  $|\psi_{\text{eff}}\rangle$  of size  $L$  and bond link  $\chi' = \chi^4$ ) has been squared, in order to keep the amount of resources limited, we need to truncate it to some fixed value  $\alpha\chi^2$ . This procedure can be iterated, row by row, thus contracting the whole PEPS in such a way that the size of the intermediate tensors stays bounded at any point. The truncation of the effective MPS into another one  $|\phi_{\text{eff}}\rangle$  such that  $\chi^4 \rightarrow \alpha\chi^2$  can be done efficiently. As a matter of fact, the overlap  $\langle\phi_{\text{eff}}|\psi_{\text{eff}}\rangle$  is quadratic in each tensor  $A^{[s]}$  of  $|\psi_{\text{eff}}\rangle$ , and thus, maximizing it can be again reduced to solving a generalized eigenvalue problem, as for the variational minimization of energy in 1D.

This shows that PEPS can be contracted approximately in an efficient way. It can further be proved that the full MPS-based contraction scheme described here for OBC scales as  $\chi^{12}$  (which using some tricks can be improved down to  $\chi^8$ ). Due to the much unfavourable scaling as compared with the MPS, the utility of the PEPS Ansatz is typically limited to much smaller  $\chi$ . It should be however noted that the error which is committed along the contraction can be controlled by the parameter  $\alpha$ : as for the MPS Ansatz, it turns out that the approximation is very accurate as long as the system is short-range correlated.

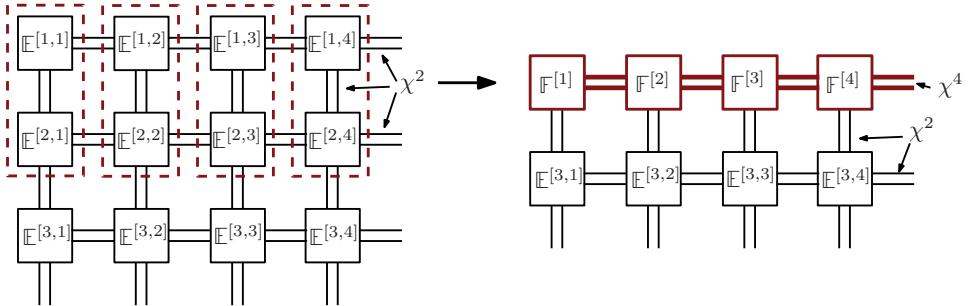


Fig. 11.28 PEPS contraction scheme using a MPS-based approach. Each tiny line connecting the different transfer matrices  $\mathbb{E}[x,y]$  carries a number of  $\chi$  states, while each bold line connecting the contracted transfer matrices  $\mathbb{F}^{[x,y]}$  carries a number of  $\chi^2$  states. At each step, the  $\mathbb{E}$  matrices of the first two rows, grouped as in the left figure, have to be contracted such to form the  $\mathbb{F}$  matrices.

Using all the machinery described so far, it is possible to generalize to the 2D case (at least from a conceptual point of view) many of the algorithms that have been described previously in 1D. In particular, it can be shown that PEPS can be used for variational ground-state and low-energy lying states calculations, as well as for the simulation of (real or imaginary) time evolutions. Given the conceptual problems discussed above on the PEPS contraction, some tricks are often required in order to speed up their manipulation, so to require a reasonable amount of resources. Here we will not enter the details of such algorithms, but will only sketch the protocol for the search of the ground state using a local variational approach, referring the interested reader to the bibliography for further details.

#### Variational ground-state search

The PEPS form of 2D wave functions can be manipulated to minimize the energy cost functional of a local Hamiltonian, thus admitting the possibility to find the best representation of the ground state in the variational Ansatz. The procedure works analogously to the 1D case with MPS, that was discussed thoroughly in Sec. 11.7. Specifically, we can start from a randomly chosen PEPS with a finite bond link  $\chi$ , that is,  $|\psi\rangle = |\psi[A^{[1,1]}, A^{[1,2]}, \dots, A^{[L,L]}]\rangle$ . Then we pick one site of the lattice, say at position  $(x, y)$ , and find the minimum of the functional

$$E(\mathbf{X}) = \frac{\langle\psi(\mathbf{X})|H|\psi(\mathbf{X})\rangle}{\langle\psi(\mathbf{X})|\psi(\mathbf{X})\rangle} = \frac{\mathbf{X}^\dagger \mathbb{H} \mathbf{X}}{\mathbf{X}^\dagger \mathbb{N} \mathbf{X}}, \quad (11.211)$$

as a function of the  $(d\chi^4)$  entries of the tensor  $A^{[x,y]} = \mathbf{X}$ . Finally we iterate the minimization over all the lattice sites, until the energy converges to a given value.

We should emphasize that the technical details behind this minimization require advanced manipulation tools which lie beyond the purpose of this introduction, and will not be mentioned here. To recall the main difficulties of this operation, let us remind that, as is the case of MPS with PBC, the normalization of a PEPS wave function cannot be simply enforced by a suitable gauge condition, and thus the

minimization involves a generalized eigenvalue problem of the type in Eq. (11.155). Moreover, the evaluation of the effective Hamiltonian  $\mathbb{H}$  and the effective norm  $\mathbb{N}$  invokes some approximate scheme which cleverly exploits the PEPS contraction in an efficient way, as discussed above.

### 11.9.2 \* Hierarchical tensor networks

We conclude by mentioning the possibility to devise a particular form of tensor networks, that are suited to describe systems exhibiting a logarithmic violation of the area law. We are interested in a scale-invariant Ansatz, which describes the ground states of gapless Hamiltonians and, as PEPS can do, sustains algebraically decaying correlation functions. For the sake of simplicity, here we focus on the 1D case, but the constructions provided below can be extended to 2D and beyond.

A first tentative toward such an Ansatz is a *tree-tensor network* (TTN), which is represented by the hierarchical structure shown in Fig. 11.29. Each horizontal sequence of tensors in the tree is called a layer. The various bond links which connect tensors belonging to different layers encode a number of states of the order of  $\chi$ , except for the lines going out of the bottom layer, which encode the  $d$  states of the physical sites of the chain. The TTN approach is analogous to a real-space renormalization group scheme, where the system's ground state is composed of energetically low lying states, which live on smaller subsystems.

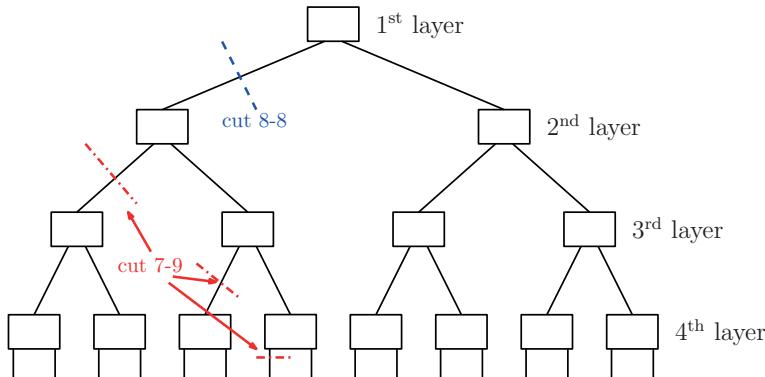


Fig. 11.29 Schematic representation of a TTN with four layers, representing the many-body wave function of a chain with  $L = 16$  sites. Dashed-dotted lines denote the required bond-link cuts that are needed to create a bipartition of the first 8 sites with the remaining 8 sites (dashed blue line), and of the first 7 sites with the remaining 9 sites (dotted-dashed red lines).

The TTN wave functions can sustain logarithmic corrections to the area-law scaling of entanglement, but only for certain bipartitions; this enables them to capture power-law correlations and to describe critical states. However, unfortunately, this Ansatz presents large fluctuations of the maximal bipartite entanglement entropy with respect to the cut in the bipartition, ranging from almost no entanglement to

critical entanglement. Indeed the amount of maximal bipartite entanglement that can be sustained by the tree is proportional to the number of bond cuts required in order to disconnect the two groups of lattice sites in the bipartition. For example, in a system with OBC and  $L = 16$  sites, the maximum admissible entanglement entropy of the leftmost 8 sites is  $O(\log \chi)$  (only the upper bond of the tree has to be cut, in order to disconnect the chain in the middle); for the first 7 sites is  $O(3 \log \chi)$  (three bonds have to be cut, in order to make a bipartition between sites 1-7 and sites 8-16); and so on.

To overcome this problem, Vidal (2008) introduced a more complicated tensor network, the *multiscale entanglement renormalization Ansatz* (MERA). This structure adds some “disentanglers” in such a way to remove the entanglement between different blocks (see Fig. 11.30). It can be shown that such hierarchical configuration can admit a bipartite entanglement which scales logarithmically with the size; indeed, in order to make a bipartition which isolates  $\ell$  sites from the rest of the chain, one should always cut a number of bonds equal to the number of layers that connect the two parts, which scales logarithmically with  $\ell$ . If  $\chi$  is the typical dimension of the tensors indexes, it is thus possible to achieve an efficient parametrization of a class of states with built-in scale invariance, using  $O(L \times \chi^4)$  parameters.

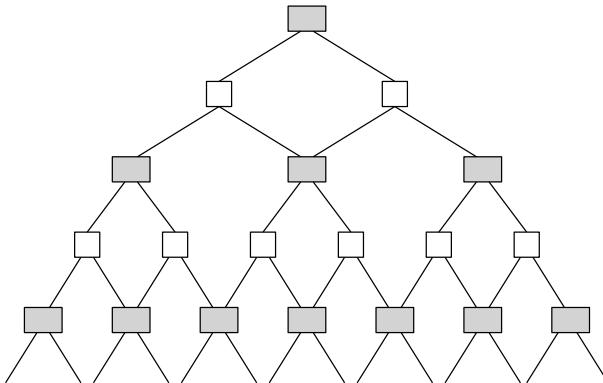


Fig. 11.30 Schematic representation of a MERA with five layers, representing the many-body wave function of a chain with  $L = 16$  sites. White tensors are isometries, and grey tensors are unitaries. In the original formulation, grey boxes are the so-called “disentanglers”.

Contrary to TTN schemes, the MERA wave function cannot be gauged into the useful isometric gauge and in general is not efficiently contractable, unless the tensors represented by grey (resp. white) boxes in Fig. 11.30 are forced to be unitaries (resp. isometries) along the vertical direction. The ultimate reason for that is the presence of “closed loops” in the tree, which forbids to have a starting boundary from which a sequence of SVD would be able to put the tensor in the isometric gauge. Imposing the isometric conditions of the various tensors, one can however see that, for any local operator  $\mathcal{O}$ , the contractions required to perform  $\text{MERA} \langle \psi | \mathcal{O} | \psi \rangle_{\text{MERA}}$  scale as  $O(\text{poly}\chi \times \log L)$ . The causal cone of  $\mathcal{O}$  has always a fixed width in terms

of tensors (but not in real space, since the width grows exponentially), and depth  $L$ . In order to evaluate two-point correlation functions such as  $\langle \mathcal{O}_j \mathcal{O}_{j+r} \rangle$ , one can see that two causal cones of fixed width eventually merge into a single causal cone, and thus can be contracted efficiently. The intersection occurs after  $\log r$  layers. If the tensors for a given layer are the same (scale invariance), each layer reduces the correlation by some factor  $\lambda$ , and thus  $\langle \mathcal{O}_j \mathcal{O}_{j+r} \rangle \sim \lambda^{2\log r} = r^{2\log \lambda}$ , which give algebraic decay of correlations. Finally we mention that the variational optimization of a MERA is much more complicated to be performed, with respect to a MPS or a PEPS. Indeed, although the energy is bilinear in each tensor, a local minimization over a single tensor cannot be mapped into an eigenvalue problem, due to the unitarity constraint discussed above. For this reason, some clever strategies need to be adopted. For further reading, we refer to the bibliography.

## 11.10 A guide to the bibliography

Nowadays Feynman's idea of realizing a quantum simulator in the lab is routinely experienced by means of several experimental platforms. An introductory paper on quantum simulators is Buluta and Nori (2009), and a more complete review can be found in Georgescu *et al.* (2014). The Nature Publishing Group dedicated an entire Nature Physics Insight on Quantum Simulation (Trabesinger *et al.*, 2012).

The following are specific reviews on the simulation of many-body physics with different approaches. Ultracold atomic gases are discussed in Lewenstein *et al.* (2007) and Bloch *et al.* (2008). The use of trapped ions as quantum simulators is reviewed in Schneider *et al.* (2012). A good reference on arrays of Josephson junctions is Fazio and van der Zant (2001), while more focused recent overviews on quantum electrodynamics with coupled QED cavities and superconducting circuits are Houck *et al.* (2012), Noh and Angelakis (2016), and Le Hur *et al.* (2016). Information on the company D-Wave Systems Inc. can be found at the website: <http://www.dwavesys.com>. See also the Conclusions and Prospects for references on other companies which are involved in quantum computing, such as IBM, Intel, Google, and Microsoft.

Various aspects of the physics of quantum strongly correlated systems, with a condensed matter perspective, are discussed in several textbooks with different flavors: Sachdev (2011) focuses on the onset of quantum phase transition and critical phenomena, Nagaosa (1999) uses a quantum field theory language, while Essler *et al.* (2005) provides a comprehensive discussion of the integrable Hubbard chain. Several aspects of the Ising chain can be also found in Suzuki *et al.* (2013). An introductory description of the emergence of Majorana physics in condensed matter systems can be found in Leijnse and Flensberg (2012).

The emergence of an area-law behaviour of entanglement in many-body systems is extensively discussed in Eisert *et al.* (2010). A review on the phenomenology and the main features of the many-body localization is Nandkishore and Huse (2015).

A fairly complete discussion on all the techniques and tricks to manipulate matrix product states is contained in Schollwöck (2011). A broad overview on quantum tensor networks is Biamonte and Bergholm (2017), while more thorough discussions on tensor network schemes are provided in Cirac and Verstraete (2009), Verstraete *et al.* (2008) and, for a specific focus on projected entangled pair states, in Orus (2014). For an historical perspective on the original density-matrix renormalization group algorithm, we refer to Schollwöck (2005). A pedagogical introduction, that is useful for the beginners who want to program a basic implementation of DMRG, is provided in De Chiara *et al.* (2008). A compendium on numerical simulation techniques and various tricks to manipulate tensor-network geometries is contained in Silvi *et al.* (2017).

A recent textbook perspective on the cross fertilization between quantum information and condensed matter theory, centered on the emerging topological aspects of quantum matter, is provided in Zeng *et al.* (2015).

**This page intentionally left blank**

# Conclusions and prospects

Quantum information has literally invaded all the fields of science, from cosmology to biophysics. Concepts and tools from quantum information find applications in various fields, including many-body physics (as discussed in Chap. 11), quantum gravity, quantum thermodynamics, and quantum biology, to name but a few. Understanding the connections between all these fields may provide fruitful cross-fertilization of ideas. Furthermore, quantum information and more generally quantum mechanics impact on our perception of nature and life, including philosophical speculations on determinism, free will, and human consciousness. The legacy of quantum information is too broad to be covered in this textbook and many relevant issues were neglected. We wish to conclude our presentation by giving a flavour of some of them, and outlining possible short- and long-term developments in the field of quantum information.

## *Quantum computation and the postulates of quantum mechanics*

Quantum computers are based on the validity of the postulates of quantum mechanics. Basic ingredients are the superposition principle and entanglement. To these two features, it is necessary to add the validity of the postulate of the projection of the wave function at the time of measurement. This postulate is essential in the study of quantum error-correcting codes, where the post-measurement state of the quantum computer is assumed to be correctly predicted by quantum mechanics.

It is well known that all the experiments performed so far are in accordance with the principles of quantum mechanics with up to 12 significant digits. However, a large-scale quantum computer capable of outperforming a classical computer on complex problems like integer factorization would require at least  $10^4 - 10^6$  qubits (including quantum error correction). Therefore, its feasibility assumes the validity of the postulates of quantum mechanics up to levels of accuracy never achieved.

While a large-scale quantum computer might remain a dream for a long time, the competition to develop prototypes of such device has already started. In this respect, nowadays the common belief is that the simplest approach to pursue resides in adiabatic quantum computing (see Sec. 3.13). On top

of D-Wave Systems, several other major companies are actively involved in this business, including IBM (see: <http://www.research.ibm.com/ibm-q/>), Google (see: <http://research.google.com/pubs/QuantumAI.html>), Intel (see: <http://newsroom.intel.com/press-kits/quantum-computing/>), and Microsoft (see: <http://www.microsoft.com/en-us/quantum/>). Besides of the D-Wave device, other examples of quantum processors made up to a few tens of superconducting qubits already exist. In 2014, Google demonstrated a nine-qubit design capable to perform digitized adiabatic quantum computing (Barends *et al.*, 2016), and in March 2018 revealed the fabrication of a 72-qubit quantum chip. In November 2017, IBM announced that also their researchers had successfully built a 50-qubit prototype. The main limitations of such implementations are that quantum coherence is preserved for a short period of time (order of 100 microseconds) and that it is generally not possible to apply two-qubit gates to all pairs of qubits. However, it can be foreseen that in a near future such prototypes will be sufficiently developed to address useful problems, and not just proofs of principle. As a matter of fact, a well controlled 50-qubit quantum computer could perform calculations that are extremely difficult to simulate without quantum technology, thus realizing the sought-after “quantum advantage” over any classical computation.

Coming back to the validity of the postulates of quantum mechanics, it should be noted that a hypothetical quantum computer capable of using 50–100 qubits would have a Hilbert space of size of  $2^{50}$ – $2^{100}$ , approximately equal to  $10^{15}$ – $10^{30}$ . This would also allow to test quantum mechanics up to a level of accuracy inaccessible to any experiment done in the past.

### *Quantum information and quantum gravity*

Since the dawn of quantum mechanics, many approaches tried to reconcile quantum principles and general relativity (for an historical account, see Rocci, 2013). On the other hand, only in the last decades the effects of the gravitational field on quantum mechanical phenomena have been concretely investigated, both theoretically and experimentally. Among these effects we can mention the gravitational red shift, the interferometry of neutrons and atoms in the gravitational field, and the recent proposals to investigate decoherence in quantum systems caused by gravitation, see Bassi *et al.* (2017).

The first ground-to-satellite quantum communication experiments (see Chap. 10) greatly enhanced the possibility of exploring the effects of large distances and gravity on quantum protocols. It is conceivable that in the not-too-far future one could perform quantum entanglement tests at the scale of interplanetary distance, and investigate how gravity and motion affect entanglement, quantum protocols, and the precision of quantum clocks. Experimental prospects in this direction are discussed in Rideout *et al.* (2012) and Howl *et al.* (2018).

It is well known that the standard quantum field theory (QFT) is based on the invariance of the theory under Lorentz (and Poincaré) transformations. However, in the presence of a gravitational field, space-time is no longer flat. Certainly, in

almost all the experiments performed on the Earth's surface, it can be assumed that gravitation has a negligible effect. However, the sensitivity reached in various experiments (ultra-cold neutron interferometers) is such as to reveal effects of Newtonian gravity. Recently, thanks to the possibility of using satellites and planetary probes, experiments that are sensitive to the space-time curvature (that is, to relativistic gravity) have been proposed, see Rideout *et al.* (2012), Howl *et al.* (2018), and Bassi *et al.* (2017).

To build a QFT in a curved space-time, a major problem is how to define the state of a particle. This problem has not yet been satisfactorily solved. A partial solution occurs in a space-time that is asymptotically flat for time  $t \rightarrow -\infty$  and  $t \rightarrow +\infty$ . In these cases, in the asymptotic regions it is possible to copy the standard formalism of QFT in a flat space-time. It follows that in general, starting at  $t \rightarrow -\infty$  from a vacuum state, for  $t \rightarrow +\infty$  a state with a non-zero particle number is obtained (the vacuum at  $t \rightarrow -\infty$  can be seen as a superposition of particle states at  $t \rightarrow +\infty$ ). The energy necessary for the creation of particles is given by the (accelerating) system that distorts space-time. Hence, the creation of particles closely follows the dynamical Casimir effect briefly described in Sec. 10.3.3.

#### *Gravitation, thermodynamics, and information theory*

We mention the physical reasons for which a connection between gravitation, thermodynamics and information theory is essential, even though, in spite of huge efforts, a quantum theory of gravitation is still not available. A link between thermodynamics and general relativity is possible since, according to the classical theory of gravitation, the gravitational field of a collapsed object (black hole) is described by few parameters (mass, electric charge, and angular momentum). It follows that if a thermal system (therefore bringing entropy) is thrown into a black hole, then the final system, after the relaxation time, is still described by few parameters. Hence, with regard to classical gravitation, we have the paradox that there is a destruction of entropy, because in the classical case a collapsed object cannot possess entropy.

This paradox, and the consequent conflict with the second law of thermodynamics, was highlighted by Wheeler, and heuristically solved by Bekenstein (1973), who hypothesized to identify the area of the collapsed object with entropy, thus saving the second principle of thermodynamics. Bekenstein introduced in his calculation an arbitrary constant, shortly afterwards calculated by Hawking (1975), thus establishing a link between gravitation and thermodynamics on rather firm grounds. Since entropy is deeply connected to information, a link between information theory and gravitation follows.

#### *Quantum information and cosmology*

The basic question in this context is whether it would be possible to consider the whole universe (including the observer) as a single quantum system, characterized by a “quantum state of the universe”. Clearly this is an extreme case, in which Bohr's separation between quantum objects and classical measurement apparatuses is not admitted.

Another point we would like to mention is that cosmic microwave background anisotropies are of quantum mechanical origin. This in principle opens up the possibility to observe quantum effects from the very early universe. However, there exist difficulties in realizing a Bell experiment on the cosmic background, discussed in Martin and Vennin (2017).

### *Quantum thermodynamics*

Quantum mechanics and thermodynamics have a deep connection, whose investigation started from thermodynamic studies by Planck (1901) and Einstein (1917). Nowadays quantum thermodynamics is a highly active field, also motivated by experimental advances and the potential of future nanoscale applications, requiring quantum mechanics for an accurate description. Many theoretical questions are being addressed, including thermalisation of quantum systems, the definition of work and heat in quantum mechanics, the role of coherence, entanglement, quantum fluctuations, and quantum measurements in thermal machines, the efficiency and power of quantum engines, and the minimum temperature achievable in small quantum chillers. These questions are vital for the development of quantum heat engines and refrigerators.

Useful reviews discussing these topics from different perspectives are Giazotto *et al.* (2006), Dubi and Di Ventra (2011), Campisi *et al.* (2011), Muhonen *et al.* (2012), Kosloff (2013), Sothmann *et al.* (2015), Gelbwaser-Klimovsky *et al.* (2015), Millen and Xuereb (2016), Goold *et al.* (2016), Vinjanampathy and Anders (2016), and Benenti *et al.* (2017). Reviews of some of the most recent developments in the field can be found in the book by Binder *et al.* (2018).

### *Quantum information and quantum biology*

The importance of quantum mechanics in biology was already highlighted by Schrödinger (1944), who tried to explain the stability of genetic information. For instance, he hypothesized the existence of an “aperiodic crystal” that encodes genetic information, an idea subsequently confirmed in a spectacular way by the discovery of the molecular structure of DNA. He assumed that genetic information is written in a large molecule, in which each atom, radical or heterocyclic ring, has its own particular task. The stability of molecules (and therefore genetic stability) cannot be explained by classical physics, but is due to the discrete nature of quantum mechanics. In Schrödinger’s theory, mutations are then due to quantum leaps in the gene.

The importance of quantum mechanics in biology has been repeatedly emphasized during the years, see for instance Davidov (1982) and Abbott *et al.* (2008) for a discussion on quantum aspects of life, in particular on the possibility of a quantum origin of life, and on the modeling of quantum decoherence in biomolecules. Today there is interest in going beyond the explanation of life with molecular “quantum chemistry” interactions, looking for the existence of effects due to overlap and entanglement in complex, noisy biological environments. Such effects are for instance

sought in photosynthetic light harvesting (see the collection of papers edited by Fleming *et al.*, 2011), in the strong coupling between living bacteria and light (see Krisnanda *et al.*, 2017, and references therein), in avian magnetoreception (Kattnig and Hore, 2017, and references therein), in quantum superposition in vision (Strini and Pizzi, 2009), in the discrimination of odorant molecules (Franco *et al.*, 2011, and references therein), and in the cooperation of living organisms (for instance, of two insects), by hypothesizing that they share a large number of quantum entangled qubit pairs, see Summhammer, 2006. For a brief review on quantum biology, see Lambert *et al.* (2013).

#### Neural networks and quantum machine learning

The possibility to exploit quantum information concepts in the fields of artificial intelligence and machine learning has been recently opening a new line of research, with the final purpose to devise and implement quantum software to program self learning machines. A prominent example in this context are *artificial neural networks*, inspired by biological brains, which model parallel information processing of a network of simple computational units (the neurons), see Schmidhuber (2015). The so-called “deep learning” approach fits in this paradigm: recurrent networks have loops, which allow feeding information from outputs of a (sub)-network back to its own input. One of the key issues in this framework is the use of arguably genuine quantum features, such as entanglement or squeezing, in order to speed up the computational power of such machines, as hinted by Deng *et al.* (2017).

Besides that, neural networks have been shown to be able to capture classical and quantum correlations, in such a way that they could form a useful ansatz for quantum states that are relevant in the context of condensed matter and many-body physics, as shown, e.g., by Carleo and Troyer (2017), and Carrasquilla and Melko (2017). Here the number of neurons in the inner layers may control the size of the representable subset of the Hilbert space, analogously to how the bond-link dimension controls the information content in a tensor-network ansatz. These results may pave the way to a near-future application of computer-science-based machine learning in the context of quantum many-body physics problems. Recent reviews on this subject are Biamonte *et al.* (2017) and Dunjko and Briegel (2018).

Let us conclude by commenting on the following question: are quantum computers the last frontier? It is useful to quote here Hameroff (1987), who discussed molecular computers, where components would be single molecules, as ultimates computers. Without neglecting the great interest and importance of the development of molecular electronic devices, we must consider that the attribute “ultimate” has been denied by quantum computers. We can then ask ourselves if the quantum computers discussed in this book must be considered as “ultimate” or if it is possible to imagine a further class of computers that could outperform them. With a minimum of trust in the fantasy of nature, we do not feel entitled to classify quantum computers as “ultimate”. Quoting Shakespeare [Hamlet (1.5.167-8)]: “*There are more things in heaven and earth, Horatio, Than are dreamt of in your philosophy.*”

**This page intentionally left blank**

# Appendix A

## Elements of linear algebra

In this Appendix, we review some elementary linear algebra. This section is self-contained and no previous knowledge of linear algebra is required. We use the Dirac’s “bra–ket” notation, which will be adopted throughout all the book. We shall neither be concerned with mathematical rigour nor completeness of exposition, but shall just provide the basic notions required to understand the fundamental principles of quantum mechanics. For a more detailed and comprehensive presentation, we refer the reader to standard textbooks. A useful reference is Lang (1996), for more advanced topics of (numerical) linear algebra see Golub and Van Loan (2013). For an operative introduction to the mathematical tools of quantum mechanics see Cohen-Tannoudji *et al.* (1977).

### A.1 Finite-dimensional vector spaces

A vector space (also called a linear space) is a set  $\mathcal{V}$  over a field  $\mathcal{F}$  where two operations can be defined, namely the vector addition and the scalar multiplication. The elements of  $\mathcal{V}$  are called *vectors*. The elements of  $\mathcal{F}$  are called *scalars*.

One prototypical example is given by the space  $\mathcal{V} \equiv \mathbb{C}^n$  over the complex field  $\mathcal{F} \equiv \mathbb{C}$ , in which a vector is singled out by an  $n$ -tuple of complex numbers  $(\alpha_1, \alpha_2, \dots, \alpha_n)$ , where  $n$  is the dimension of the vector space. Here we mostly focus on this situation, since, as we shall see below, this is the framework to study quantum mechanical systems with finite-dimensional spaces ( $n < +\infty$ ). Following Dirac’s notation, we write a vector using the symbol  $|\alpha\rangle$  and call it a “ket”.

Let us start, in the next section, with the definition of the two fundamental operations in a vector space and their corresponding basic properties.

#### A.1.1 Basic properties of vector spaces

Two vectors (kets)  $|\alpha\rangle, |\beta\rangle \in \mathcal{V}$  may be added to give a new vector

$$|\gamma\rangle = |\alpha\rangle + |\beta\rangle, \quad (\text{A.1})$$

residing in the same vector space  $\mathcal{V}$ . For example, in the vector space  $\mathbb{C}^n$ , we have, in terms of the vector components  $|\alpha\rangle = (\alpha_1, \dots, \alpha_n)$ ,  $|\beta\rangle = (\beta_1, \dots, \beta_n)$  and

$$|\gamma\rangle = (\gamma_1, \dots, \gamma_n),$$

$$\gamma_i = \alpha_i + \beta_i, \quad (i = 1, \dots, n). \quad (\text{A.2})$$

The vector addition is required to have the following properties:

$$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle, \quad (\text{commutativity}), \quad (\text{A.3a})$$

$$|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle, \quad (\text{associativity}). \quad (\text{A.3b})$$

It is also possible to multiply a vector  $|\alpha\rangle \in \mathcal{V}$  by a scalar  $c \in \mathcal{F}$ , to obtain a new vector  $c|\alpha\rangle$ , also written as  $|c\alpha\rangle$ . The following properties of the scalar multiplication hold for any  $c, d \in \mathcal{F}$  and  $|\alpha\rangle, |\beta\rangle \in \mathcal{V}$ :

$$c(|\alpha\rangle + |\beta\rangle) = c|\alpha\rangle + c|\beta\rangle, \quad (\text{A.4a})$$

$$(c + d)|\alpha\rangle = c|\alpha\rangle + d|\alpha\rangle, \quad (\text{A.4b})$$

$$(cd)|\alpha\rangle = c(d|\alpha\rangle), \quad (\text{A.4c})$$

denoting, respectively, distributivity with respect to vector addition and to field addition, and compatibility with field multiplication.

A vector space must contain the zero vector  $|o\rangle \in \mathcal{V}$ , which is defined by the following requirement: for any vector  $|\alpha\rangle$  belonging to the vector space,

$$|\alpha\rangle + |o\rangle = |\alpha\rangle. \quad (\text{A.5})$$

Moreover, for every vector  $|v\rangle \in \mathcal{V}$ , there exists an element  $|-v\rangle \in \mathcal{V}$ , called the additive inverse of  $|v\rangle$ , such that

$$|v\rangle + |-v\rangle = |o\rangle. \quad (\text{A.6})$$

Note that the zero vector  $|o\rangle$  has a distinct notation from the vector  $|0\rangle$ , since, as we shall see below, the latter denotes a state of the computational basis. It is of common habit, with abuse of notation, to indicate the zero vector simply by 0, thus identifying  $|o\rangle \equiv 0$ . In the remainder of the book we will use the latter convention.

We finally remark that, in a vector space, the following properties of the scalar multiplication hold:  $0|\alpha\rangle = 0$  and  $1|\alpha\rangle = |\alpha\rangle$ , where  $0, 1 \in \mathcal{F}$  respectively are the null and the identity element of the field.

### A.1.2 Inner product and norm of a vector

In many circumstances, a vector space can be enriched by additional structures, such as the inner product (also known as scalar product). The inner product of an ordered pair of vectors  $|\alpha\rangle, |\beta\rangle \in \mathcal{V}$  is a map over the field  $\mathcal{F}$ , denoted as  $\langle\alpha|\beta\rangle$ , with the following requirements:

$$(i) \quad \langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^* \quad (\text{skew symmetry}), \quad (\text{A.7a})$$

where, after specializing to the case  $\mathcal{F} = \mathbb{C}$ , for any complex number  $c = a + ib$  ( $a, b \in \mathbb{R}$ ),  $c^* = a - ib$  denotes its complex conjugate;

$$(ii) \quad \langle\alpha|c\beta + d\gamma\rangle = c\langle\alpha|\beta\rangle + d\langle\alpha|\gamma\rangle \quad (\text{linearity}), \quad (\text{A.7b})$$

with  $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in \mathcal{V}$  and  $c, d \in \mathcal{F}$ ;

$$(iii) \quad \langle\alpha|\alpha\rangle \geq 0 \quad (\text{positivity}), \quad (\text{A.7c})$$

for any  $|\alpha\rangle \in \mathcal{V}$ , with equality if and only if  $|\alpha\rangle$  is the zero vector.

Taking into account the previous properties, it is easy to check the following:

$$\langle c\alpha|\beta\rangle = c^*\langle\alpha|\beta\rangle. \quad (\text{A.8})$$

Note that  $\langle\alpha|$  is the *dual vector* (also called “bra”) to the vector  $|\alpha\rangle$ . The dual vector  $\langle\alpha|$  is a linear operator from the vector space  $\mathcal{V}$  to the field  $\mathcal{F}$ , defined by  $\langle\alpha|(|\beta\rangle) = \langle\alpha|\beta\rangle$ , for any  $|\beta\rangle \in \mathcal{V}$ .

As an example, we can define an inner product between two vectors  $|\alpha\rangle = (\alpha_1, \dots, \alpha_n)$  and  $|\beta\rangle = (\beta_1, \dots, \beta_n)$  in  $\mathbb{C}^n$  over the complex field, as follows:

$$\langle\alpha|\beta\rangle = \sum_{i=1}^n \alpha_i^* \beta_i. \quad (\text{A.9})$$

Once a vector space is equipped with an inner product, it is possible to associate any vector  $|\alpha\rangle \in \mathcal{V}$  with a scalar  $\|\alpha\| \in \mathcal{F}$ , called *norm* of the vector, defined by

$$\|\alpha\| = \sqrt{\langle\alpha|\alpha\rangle}. \quad (\text{A.10})$$

Any non-zero vector  $|\alpha\rangle$  can be normalized by dividing it by its norm  $\|\alpha\|$ . The normalized vector  $|\alpha\rangle/\|\alpha\|$  has unit norm unit and is therefore called a *unit vector*. Using the inner product in  $\mathbb{C}^n$  defined in Eq. (A.9), the norm of a vector  $|\alpha\rangle = (\alpha_1, \dots, \alpha_n)$  is given by

$$\|\alpha\| = \sqrt{\sum_{i=1}^n |\alpha_i|^2}, \quad (\text{A.11})$$

and a unit vector must satisfy the condition  $\sum_i |\alpha_i|^2 = 1$ .

In the finite-dimensional case, which is relevant for most of the quantum information theory illustrated in this book, a *Hilbert space* is defined as a complex vector space equipped with an inner product. As discussed in Sec. 2.3, quantum mechanics associates to a physical system a unit vector residing in a Hilbert space.

### The Cauchy–Schwarz inequality

For any two vectors  $|\alpha\rangle$  and  $|\beta\rangle$  belonging to a Hilbert space, the following inequality holds:

$$|\langle\alpha|\beta\rangle|^2 \leq \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle. \quad (\text{A.12})$$

**Proof.** The inner product is positive definite and therefore  $\langle\alpha - c\beta|\alpha - c\beta\rangle \geq 0$  holds for any  $|\alpha\rangle, |\beta\rangle \in \mathcal{V}$  and  $c \in \mathcal{F}$ . Owing to the linearity of the inner product, this relation is equivalent to  $\langle\alpha|\alpha\rangle - c\langle\alpha|\beta\rangle - c^*\langle\beta|\alpha\rangle + cc^*\langle\beta|\beta\rangle \geq 0$ . Taking  $c = \langle\beta|\alpha\rangle/\langle\beta|\beta\rangle$ , we obtain the Cauchy–Schwarz inequality (A.12).  $\square$

Note that, in the special case in which the inner product is real, that is  $\mathcal{F} = \mathbb{R}$ , the Cauchy–Schwarz inequality admits a simple geometrical interpretation. Indeed, since in this case

$$-1 \leq \frac{\langle \alpha | \beta \rangle}{\| |\alpha\rangle \| \| |\beta\rangle \|} \leq 1, \quad (\text{A.13})$$

we can write

$$\langle \alpha | \beta \rangle = \| |\alpha\rangle \| \| |\beta\rangle \| \cos \theta. \quad (\text{A.14})$$

This latter equation corresponds to the usual definition of the scalar product of two vectors  $|\alpha\rangle$  and  $|\beta\rangle$ , where  $\theta$  is the angle between the two vectors.

### A.1.3 Linear independence and the notion of basis

Two non-zero vectors  $|\alpha\rangle$  and  $|\beta\rangle$  are said to be *orthogonal* if their inner product is zero:

$$\langle \alpha | \beta \rangle = 0. \quad (\text{A.15})$$

A set of vectors  $|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle$  is said to be *orthonormal* if

$$\langle \alpha_i | \alpha_j \rangle = \delta_{ij} \quad (i, j = 1, 2, \dots, n), \quad (\text{A.16})$$

where  $\delta_{ij}$  is the Kronecker symbol, defined as  $\delta_{ij} = 1$  for  $i = j$  and  $\delta_{ij} = 0$  for  $i \neq j$ .

It is easy to see that the orthogonal vectors  $|\alpha_i\rangle$ , satisfying condition (A.16), also satisfy the following property of linear independence. Namely, a set of vectors  $|\alpha_1\rangle, \dots, |\alpha_m\rangle \in \mathcal{V}$  are said to be *linearly independent* if the relation

$$c_1 |\alpha_1\rangle + c_2 |\alpha_2\rangle + \dots + c_m |\alpha_m\rangle = 0, \quad (\text{A.17})$$

with  $c_1, c_2, \dots, c_m \in \mathcal{F}$ , holds if and only if  $c_1 = c_2 = \dots = c_m = 0$ .

The notion of linear independence enables to rigorously define the *dimension n* of a vector space, that is the maximum number of linearly independent vectors. A set of linearly independent vectors  $|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle$  in an  $n$ -dimensional vector space  $\mathcal{V}$  is said to be a *basis* for  $\mathcal{V}$ . Since any vector  $|\alpha\rangle$  can be expanded over a basis  $\{|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle\}$ ,

$$|\alpha\rangle = \sum_{i=1}^n a_i |\alpha_i\rangle \quad \text{with } a_i \in \mathcal{F}, \quad (\text{A.18})$$

we call the vectors  $|\alpha_i\rangle$  a *complete set* of vectors. The coefficients  $a_i$  are known as the components of the vector  $|\alpha\rangle$  with respect to the basis  $\{|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle\}$ . They are uniquely determined and, for an orthonormal basis, we have:

$$a_i = \langle \alpha_i | \alpha \rangle. \quad (\text{A.19})$$

The ordered ensemble of components  $\{a_1, a_2, \dots, a_n\}$  constitutes a *representation* of the vector  $|\alpha\rangle$ .

An example of special interest for us is the vector space  $\mathbb{C}^2$  over the complex field. A generic vector  $|\alpha\rangle \in \mathbb{C}^2$  can be written as

$$|\alpha\rangle = a_1|\alpha_1\rangle + a_2|\alpha_2\rangle, \quad (\text{A.20})$$

$a_1$  and  $a_2$  being two complex numbers, where the vectors  $|\alpha_1\rangle$  and  $|\alpha_2\rangle$  have components  $\alpha_1 = (1, 0)$  and  $\alpha_2 = (0, 1)$ . We shall use the following notation:

$$|\alpha_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\alpha_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (\text{A.21})$$

This basis has nothing special. A generic vector  $|\alpha\rangle$  can be expanded over any (orthonormal) basis. For instance, instead of Eq. (A.20) we can write

$$|\alpha\rangle = a'_1|\alpha'_1\rangle + a'_2|\alpha'_2\rangle, \quad (\text{A.22})$$

where

$$|\alpha'_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |\alpha'_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (\text{A.23})$$

It is easy to check that the coefficients  $a'_1$  and  $a'_2$  are related to the coefficients  $a_1$  and  $a_2$  of the expansion (A.20) as follows:

$$a'_1 = \frac{1}{\sqrt{2}}(a_1 + a_2), \quad a'_2 = \frac{1}{\sqrt{2}}(a_1 - a_2). \quad (\text{A.24})$$

We now compute the inner product of two generic vectors  $|\alpha\rangle$  and  $|\beta\rangle$ :

$$\langle\alpha|\beta\rangle = \left( \sum_i a_i^* \langle\alpha_i| \right) \left( \sum_j b_j |\alpha_j\rangle \right) = \sum_{i,j} a_i^* b_j \langle\alpha_i|\alpha_j\rangle = \sum_i a_i^* b_i. \quad (\text{A.25})$$

In particular, the norm of a generic vector  $|\alpha\rangle$  can be written as

$$\|\alpha\| = \sqrt{\langle\alpha|\alpha\rangle} = \sqrt{\sum_i |a_i|^2}. \quad (\text{A.26})$$

Note that the representation described above generalizes the expansion of a vector over orthogonal axes in three-dimensional Euclidean space. In particular, in this case the inner product becomes the usual scalar product of two vectors  $\mathbf{u} = (u_1, u_2, u_3)$  and  $\mathbf{v} = (v_1, v_2, v_3)$ :

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^3 u_i v_i, \quad (\text{A.27})$$

where  $(\mathbf{u}, \mathbf{v})$  is the angle between the vectors  $\mathbf{u}$  and  $\mathbf{v}$ .

### Gram–Schmidt decomposition

The Gram–Schmidt decomposition permits the construction of an orthonormal basis. Let us consider a basis  $\{|\alpha_1\rangle, |\alpha_2\rangle, \dots, |\alpha_n\rangle\}$  in an  $n$ -dimensional Hilbert space. It is easy to check that the vectors

$$|\beta_1\rangle = |\alpha_1\rangle \quad (\text{A.28a})$$

and

$$|\beta_2\rangle = |\alpha_2\rangle - \frac{\langle\beta_1|\alpha_2\rangle}{\|\beta_1\|^2}|\beta_1\rangle \quad (\text{A.28b})$$

are mutually orthogonal. We can define inductively, for any  $i = 2, 3, \dots, n$ , the vector

$$|\beta_i\rangle = |\alpha_i\rangle - \sum_{k=1}^{i-1} \frac{\langle\beta_k|\alpha_i\rangle}{\|\beta_k\|^2}|\beta_k\rangle. \quad (\text{A.28c})$$

It is easy to see that the vectors  $\{|\beta_1\rangle, |\beta_2\rangle, \dots, |\beta_n\rangle\}$  are mutually orthogonal. Thus, an orthonormal basis for the Hilbert space is given by

$$|\gamma_i\rangle = \frac{|\beta_i\rangle}{\|\beta_i\|}, \quad (i = 1, 2, \dots, n). \quad (\text{A.29})$$

#### A.1.4 Linear operators

An operator  $A$  is a map which takes a vector  $|\alpha\rangle \in \mathcal{V}$  into another vector  $|\beta\rangle \in \mathcal{V}$ :

$$|\beta\rangle = A|\alpha\rangle. \quad (\text{A.30})$$

This is said to be *linear* if, for any vectors  $|\alpha\rangle, |\beta\rangle \in \mathcal{V}$  and for any elements  $a, b$  of the field  $\mathcal{F}$ , the following fundamental property holds:

$$A(a|\alpha\rangle + b|\beta\rangle) = aA|\alpha\rangle + bA|\beta\rangle \quad (\text{additivity and homogeneity}). \quad (\text{A.31a})$$

We can produce new linear operators by defining the following operations between two linear operators  $A$  and  $B$ :

$$(i) \quad C|\alpha\rangle = A|\alpha\rangle + B|\alpha\rangle \quad (\text{with } C = A + B), \quad (\text{A.32a})$$

$$(ii) \quad D|\alpha\rangle = A(B|\alpha\rangle) \quad (\text{with } D = AB). \quad (\text{A.32b})$$

These respectively denote the sum and the product of linear operators. Note that the application of the operator  $D = AB$  to the vector  $|\alpha\rangle$  is equivalent to first applying  $B$  to  $|\alpha\rangle$  and then  $A$  to the vector  $B|\alpha\rangle$ .

Two operators  $A$  and  $B$  are said to be equal (and we write  $A = B$ ) if, for any vector  $|\alpha\rangle \in \mathcal{V}$ ,

$$A|\alpha\rangle = B|\alpha\rangle. \quad (\text{A.33})$$

It is easy to check that the addition (A.32a) of two linear operators obeys commutativity, that is,  $A + B = B + A$ . The same property however does not hold for the

multiplication (A.32b), that is, in general  $AB \neq BA$ . As we shall see later, it is only in special cases that the two operators commute, that is,  $AB = BA$ .

The simplest example of a linear operator is the *identity* operator  $I$ :

$$I|\alpha\rangle = |\alpha\rangle. \quad (\text{A.34})$$

Another simple example is the *zero* operator  $N$ , which maps any vector  $|\alpha\rangle \in \mathcal{V}$  into the zero vector 0:

$$N|\alpha\rangle = 0. \quad (\text{A.35})$$

*Remark:* From now on, when dealing with finite-dimensional spaces, we will always consider the Hilbert space  $\mathcal{H} = \mathbb{C}^n$  (for a given  $n$ ) over the complex field. Quantum mechanical theory for finite-dimensional systems is naturally framed in this space.

### Projectors

An important class of operators is given by the *projectors*. If  $|\alpha\rangle \in \mathcal{H}$  is a unit vector, one can define the unidimensional projector  $P_\alpha$  as follows:

$$|\beta\rangle = P_\alpha|\gamma\rangle = |\alpha\rangle\langle\alpha|\gamma\rangle = \langle\alpha|\gamma\rangle|\alpha\rangle, \quad (\text{A.36})$$

where  $|\gamma\rangle$  denotes any vector in the space  $\mathcal{H}$ . Note that  $P_\alpha = |\alpha\rangle\langle\alpha|$  correctly defines an operator, since it maps a vector into a vector. This linear operator is called a projector since it projects a generic vector  $|\gamma\rangle$  along the direction  $|\alpha\rangle$ . In particular,  $P_\alpha|\alpha\rangle = |\alpha\rangle$  and  $P_\alpha|\gamma\rangle = 0$  for any  $|\gamma\rangle$  orthogonal to  $|\alpha\rangle$ . A projector satisfies the following property:

$$P_\alpha^2 = P_\alpha. \quad (\text{A.37})$$

This property is easy to check by taking into account that  $P_\alpha|\alpha\rangle = |\alpha\rangle$ .

Definition (A.36) is readily extended to projectors over multi-dimensional subspaces. In such cases we have

$$P = \sum_{l=1}^k |\alpha_l\rangle\langle\alpha_l|, \quad (\text{A.38})$$

where  $k$  is the dimension of the subspace over which the operator  $P$  projects. Again, it is easy to check that  $P^2 = P$ . We note that it is also possible to prove that a linear operator  $P$ , satisfying  $P^2 = P$ , univocally defines a projector. Therefore this property can be taken as the definition of a projector.

### Completeness relation

Starting from the relation  $a_i = \langle\alpha_i|\alpha\rangle$  ( $i = 1, \dots, n$ ) of Eq. (A.19), we obtain

$$\left( \sum_i |\alpha_i\rangle\langle\alpha_i| \right) |\alpha\rangle = \sum_i |\alpha_i\rangle\langle\alpha_i|\alpha\rangle = \sum_i a_i |\alpha_i\rangle = |\alpha\rangle. \quad (\text{A.39})$$

Note that  $\sum_i |\alpha_i\rangle\langle\alpha_i|$  is a multi-dimensional projector over the basis  $\{|\alpha_1\rangle, \dots, |\alpha_n\rangle\}$  for the vector space. Since relation (A.39) applies for any vector  $|\alpha\rangle$ , we have the so called completeness relation (also known as closure relation)

$$\sum_i |\alpha_i\rangle\langle\alpha_i| = I, \quad (\text{A.40})$$

where  $I$  is the identity operator defined by Eq. (A.34).

### Hermitian operators

For any linear operator  $A$  on a Hilbert space  $\mathcal{H}$ , it is possible to show that there exists a unique linear operator  $A^\dagger$  on  $\mathcal{H}$ , called the *adjoint* or *Hermitian conjugate* of  $A$ , such that, for all vectors  $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$ ,

$$\langle \alpha | A\beta \rangle = \langle A^\dagger \alpha | \beta \rangle. \quad (\text{A.41})$$

Starting from the definition (A.41), it is easy to see that  $\langle A\alpha | \beta \rangle = \langle \alpha | A^\dagger \beta \rangle$ . (Indeed,  $\langle A\alpha | \beta \rangle = \langle \beta | A\alpha \rangle^* = \langle A^\dagger \beta | \alpha \rangle^* = \langle \alpha | A^\dagger \beta \rangle$ .)

A particularly interesting case is that in which  $A$  is *Hermitian* or *self-adjoint*; that is, it is equal to its own adjoint:

$$A^\dagger = A. \quad (\text{A.42})$$

In this case, the scalar product  $\langle \alpha | A\alpha \rangle$  is real (since  $\langle \alpha | A\alpha \rangle^* = \langle A\alpha | \alpha \rangle = \langle \alpha | A\alpha \rangle$ ).

**Exercise A.1** Show that  $(A + B)^\dagger = A^\dagger + B^\dagger$ ,  $(AB)^\dagger = B^\dagger A^\dagger$  and  $(A^\dagger)^\dagger = A$ .

### Inverse operator

Let us consider a linear operator  $A$ . If there exists an operator  $B$  such that

$$AB = BA = I, \quad (\text{A.43})$$

we call  $B$  the inverse of  $A$  and write  $B = A^{-1}$ . If we have  $|\beta\rangle = A|\alpha\rangle$ , then  $|\alpha\rangle = A^{-1}|\beta\rangle$ . It is possible to show that the inverse of an operator  $A$  exists if and only if the equation  $A|\alpha\rangle = 0$  implies that  $|\alpha\rangle$  is the zero vector. As we shall see below, it is immediate to conclude that the inverse of an operator  $A$  exists if and only if the determinant of its matrix representation is different from zero.

**Exercise A.2** Show that a projector  $P$  is Hermitian and can be inverted if and only if  $P = I$ .

### Unitary operators

An operator  $U$  is said to be unitary if

$$UU^\dagger = U^\dagger U = I. \quad (\text{A.44})$$

From this definition, we have that the adjoint of a unitary operator coincides with its inverse,

$$U^\dagger = U^{-1}, \quad (\text{A.45})$$

and that  $U^\dagger$  is unitary. The product  $UV$  of two unitary operators is unitary, since

$$(UV)(UV)^\dagger = UVV^\dagger U^\dagger = I. \quad (\text{A.46})$$

Unitary operators have the important property that they preserve the inner product between vectors. To see this, let us consider any two vectors  $|\alpha\rangle$  and  $|\beta\rangle$ . If we define  $|\gamma\rangle = U|\alpha\rangle$  and  $|\nu\rangle = U|\beta\rangle$ , then

$$\langle \gamma | \nu \rangle = \langle U\alpha | U\beta \rangle = \langle \alpha | U^\dagger U | \beta \rangle = \langle \alpha | \beta \rangle. \quad (\text{A.47})$$

If we take  $|\alpha\rangle = |\beta\rangle$ , we see that a unitary operator does not change the norm of a vector. Therefore, unitary operators act on vectors in Hilbert space in a way analogous to rotations in Euclidean space, which preserve both the length of a vector and the angle between two vectors.

### Matrix representation

A linear operator  $A$  can be represented as a square matrix by means of a complete set of vectors. Hereafter we will always denote the matrix representation with the same symbol  $A$  used for the operator. Let us consider the action of the operator on a generic vector  $|\alpha\rangle \in \mathcal{H}$ , namely,  $A|\alpha\rangle = |\beta\rangle$ . We expand the two vectors  $|\alpha\rangle$  and  $|\beta\rangle$  over an orthonormal basis  $\{|\gamma_1\rangle, |\gamma_2\rangle, \dots, |\gamma_n\rangle\}$ :

$$|\alpha\rangle = \sum_i a_i |\gamma_i\rangle, \quad |\beta\rangle = \sum_i b_i |\gamma_i\rangle, \quad (\text{A.48})$$

and therefore

$$b_i = \langle \gamma_i | \beta \rangle = \langle \gamma_i | A\alpha \rangle = \sum_j \langle \gamma_i | A\gamma_j \rangle a_j \equiv \sum_j A_{ij} a_j \quad (i = 1, 2, \dots, n), \quad (\text{A.49})$$

where we have defined

$$A_{ij} = \langle \gamma_i | A\gamma_j \rangle. \quad (\text{A.50})$$

Note that we shall also use the notation  $\langle \gamma_i | A | \gamma_j \rangle \equiv \langle \gamma_i | A\gamma_j \rangle$ . The system of equations (A.49) reads

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad (\text{A.51})$$

where a generic vector  $|\alpha\rangle$  is represented as a column vector:

$$|\alpha\rangle = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}. \quad (\text{A.52})$$

Thus, if we know all the matrix elements  $A_{ij}$ , we can compute the action of the operator  $A$  on a generic vector  $|\alpha\rangle \in \mathcal{H}$ , using relation (A.51). Note that, by means of this formalism, it is also possible to represent the inner product. For example, using Eq. (A.9), we have:

$$\langle \alpha | \beta \rangle = \sum_i a_i^* b_i = [a_1^*, a_2^*, \dots, a_n^*] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \quad (\text{A.53})$$

We recall that, given the matrix representation of  $A$  over a basis according to Eq. (A.50), the corresponding representation of the adjoint operator  $A^\dagger$  is well determined. Indeed, from the definition (A.41) we have that  $\langle A\gamma_i | \gamma_j \rangle = \langle \gamma_i | A^\dagger \gamma_j \rangle$  and this relation can be written as

$$(A_{ji})^* = (A^\dagger)_{ij}, \quad (\text{A.54})$$

where  $\star$  denotes the complex conjugate operation. Therefore, if we define the *transpose matrix* by  $A_{ij}^T = A_{ji}$ , the matrix elements of  $A^\dagger$  are the complex conjugates of the matrix elements of  $A^T$ , that is:

$$A^\dagger = (A^T)^\star. \quad (\text{A.55})$$

Note that, by definition, for a Hermitian operator we have

$$A = (A^T)^\star. \quad (\text{A.56})$$

As a consequence, the diagonal matrix elements of a Hermitian operator are real:  $A_{ii} = (A_{ii}^T)^\star = (A_{ii})^\star$ .

### Matrix exponential

It is possible to define a function that maps a linear operator into another linear operator acting on the same space, in a way analogous to the ordinary exponential function. For any matrix representation  $A$ , its exponential is a matrix of the same dimension, given by the power series

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \quad (\text{A.57})$$

where, by definition,  $A^k = A \times A \times \cdots \times A$  for  $k$  times (here  $\times$  denotes the usual matrix multiplication). The matrix exponential satisfies the following properties:

$$(i) \quad e^0 = I, \quad (\text{A.58a})$$

$$(ii) \quad e^{cA} e^{dA} = e^{(c+d)A}, \quad (\text{A.58b})$$

$$(iii) \quad e^{BAB^{-1}} = Be^A B^{-1}, \quad \forall B \text{ invertible}, \quad (\text{A.58c})$$

$$(iv) \quad e^{A^\star} = (e^A)^\star. \quad (\text{A.58d})$$

### Pauli matrices

A useful class of linear operators which will be often used throughout this book, is given by the so called Pauli matrices  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$ , which act on the  $\mathbb{C}^2$  vector space. Using the matrix representation, they are defined as follows:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (\text{A.59})$$

These matrices have the following relevant properties:

$$(i) \quad \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I, \text{ where } I \text{ is the identity matrix in } \mathbb{C}^2;$$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad (\text{A.60})$$

$$(ii) \quad \sigma_x \sigma_y = i\sigma_z, \quad \sigma_y \sigma_z = i\sigma_x, \quad \sigma_z \sigma_x = i\sigma_y.$$

**Exercise A.3** Show that the Pauli matrices are both Hermitian and unitary.

### A.1.5 Tensor product

Let us consider two generic vector spaces  $\mathcal{V}$  and  $\mathcal{W}$  over a field  $\mathcal{F}$ , of dimension  $m$  and  $n$ , respectively. We say that the vector space  $\mathcal{Z}$  is the tensor (or outer) product of  $\mathcal{V}$  and  $\mathcal{W}$ , and we write  $\mathcal{Z} = \mathcal{V} \otimes \mathcal{W}$ , if we can associate with each pair of vectors  $|\alpha\rangle \in \mathcal{V}$  and  $|\beta\rangle \in \mathcal{W}$  a vector belonging to  $\mathcal{Z}$ , denoted by  $|\alpha\rangle \otimes |\beta\rangle$  and called the tensor product of  $|\alpha\rangle$  and  $|\beta\rangle$ . By definition, the vectors in  $\mathcal{Z}$  are linear superpositions of the above vectors  $|\alpha\rangle \otimes |\beta\rangle$  and the following properties are satisfied:

- (i) for any  $|\alpha\rangle \in \mathcal{V}$ ,  $|\beta\rangle \in \mathcal{W}$  and  $c \in \mathcal{F}$ ,

$$c(|\alpha\rangle \otimes |\beta\rangle) = (c|\alpha\rangle) \otimes |\beta\rangle = |\alpha\rangle \otimes (c|\beta\rangle); \quad (\text{A.61a})$$

- (ii) for any  $|\alpha_1\rangle, |\alpha_2\rangle \in \mathcal{V}$  and  $|\beta\rangle \in \mathcal{W}$ ,

$$(|\alpha_1\rangle + |\alpha_2\rangle) \otimes |\beta\rangle = |\alpha_1\rangle \otimes |\beta\rangle + |\alpha_2\rangle \otimes |\beta\rangle; \quad (\text{A.61b})$$

- (iii) for any  $|\alpha\rangle \in \mathcal{V}$  and  $|\beta_1\rangle, |\beta_2\rangle \in \mathcal{W}$ ,

$$|\alpha\rangle \otimes (|\beta_1\rangle + |\beta_2\rangle) = |\alpha\rangle \otimes |\beta_1\rangle + |\alpha\rangle \otimes |\beta_2\rangle. \quad (\text{A.61c})$$

Note that, instead of  $|\alpha\rangle \otimes |\beta\rangle$ , we shall often use the shorthand notations  $|\alpha\rangle|\beta\rangle$ ,  $|\alpha, \beta\rangle$  or  $|\alpha\beta\rangle$ . Vectors belonging to  $\mathcal{Z}$  are also referred to as tensors.

The dimension of the vector space  $\mathcal{Z}$  is given by the product  $m \times n$  of the dimensions of  $\mathcal{V}$  and  $\mathcal{W}$ . Indeed, if  $|i\rangle$  and  $|j\rangle$  are orthonormal bases for  $\mathcal{V}$  and  $\mathcal{W}$ , then an orthonormal basis for  $\mathcal{Z} = \mathcal{V} \otimes \mathcal{W}$  is given by  $|i\rangle \otimes |j\rangle$ . For instance, let us consider two bidimensional Hilbert spaces ( $m = n = 2$ ), denoted with  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , with basis vectors  $|0\rangle$  and  $|1\rangle$ . The above defined tensor product  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$  has dimension  $mn = 4$  and basis vectors  $|0\rangle \otimes |0\rangle$ ,  $|0\rangle \otimes |1\rangle$ ,  $|1\rangle \otimes |0\rangle$  and  $|1\rangle \otimes |1\rangle$ . Therefore, a generic vector  $|\Psi\rangle \in \mathcal{H}$  can be expanded over this basis as follows:

$$|\Psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle, \quad (\text{A.62})$$

where  $c_{ij} \equiv \langle ij|\Psi\rangle$ , with  $i, j = 0, 1$ .

We point out that the tensor product can be iterated on more than two input vector spaces, that is, one can define  $\mathcal{Z} = \mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes (\mathcal{V}_3 \otimes \dots \otimes \mathcal{V}_N))$ . The resulting multi-tensor space has dimension  $m_1 \times m_2 \times \dots \times m_N$ , where  $m_j$  is the dimension of the  $j$ -th input space  $\mathcal{V}_j$ . The tensor product is associative:

$$\mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes \mathcal{V}_3) = (\mathcal{V}_1 \otimes \mathcal{V}_2) \otimes \mathcal{V}_3 \quad (\text{A.63})$$

therefore we can omit the parentheses and simply write  $\mathcal{V}_1 \otimes \mathcal{V}_2 \otimes \mathcal{V}_3$ .

#### Tensor products of linear maps

The tensor product also operates on linear maps between vector spaces. In the specific, if  $A$  and  $B$  are linear operators acting on the spaces  $\mathcal{V}$  and  $\mathcal{W}$ , respectively, then the action of  $A \otimes B$  on a generic vector

$$|\Psi\rangle = \sum_{ij} c_{ij}|i\rangle \otimes |j\rangle \quad (\text{A.64})$$

residing in  $\mathcal{Z}$  is defined by another linear operator

$$(A \otimes B) \left( \sum_{ij} c_{ij} |i\rangle \otimes |j\rangle \right) = \sum_{ij} c_{ij} A |i\rangle \otimes B |j\rangle. \quad (\text{A.65})$$

It is possible to show that a generic linear operator  $O$  acting on  $\mathcal{Z}$  can be written as a linear superposition of tensor products of linear operators  $A_i$  acting on  $\mathcal{V}$  and  $B_j$  acting on  $\mathcal{W}$ :

$$O = \sum_{ij} \gamma_{ij} A_i \otimes B_j. \quad (\text{A.66})$$

In the case where  $\mathcal{V}$  and  $\mathcal{W}$  are Hilbert spaces, one can also define the inner product of two generic vectors belonging to the tensor Hilbert space  $\mathcal{Z}$ ,  $|\Psi\rangle = \sum_{ij} c_{ij} |ij\rangle$  and  $|\Phi\rangle = \sum_{ij} d_{ij} |ij\rangle$ , according to the following:

$$\langle \Psi | \Phi \rangle = \sum_{ij} c_{ij}^* d_{ij}. \quad (\text{A.67})$$

It is easy to show that this definition satisfies the properties of an inner product.

The matrix representation of the operator  $A \otimes B$  in the basis  $|K\rangle \equiv |ij\rangle$ , labelled by the single index  $K = 1, 2, \dots, mn$ , with  $K = (i-1)n + j$ , is given by

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1m}B \\ A_{21}B & A_{22}B & \cdots & A_{2m}B \\ \vdots & \vdots & & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mm}B \end{bmatrix}, \quad (\text{A.68})$$

where the terms  $A_{ij}B$  denote sub-matrices of size  $n \times n$ , with  $A$  and  $B$  matrix representations of the operators  $A$  and  $B$  ( $A$  and  $B$  are  $m \times m$  and  $n \times n$  matrices, respectively). For instance, let us consider the matrix representation of the tensor product of the Pauli matrices  $\sigma_x$  and  $\sigma_z$ :

$$\sigma_x \otimes \sigma_z = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 \cdot \sigma_z & 1 \cdot \sigma_z \\ 1 \cdot \sigma_z & 0 \cdot \sigma_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}. \quad (\text{A.69})$$

**Exercise A.4** Compute the tensor products  $\sigma_x \otimes \sigma_y$  and  $I \otimes \sigma_x$ .

As a further example, let us consider the vectors  $|\alpha\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  and  $|\beta\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , and compute their tensor product  $|\alpha\rangle \otimes |\beta\rangle$ . It has matrix representation, with respect to the basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ , given by

$$|\alpha\rangle \otimes |\beta\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \cdot |\beta\rangle \\ -1 \cdot |\beta\rangle \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}. \quad (\text{A.70})$$

### A.1.6 Matrix decompositions

A given matrix can be decomposed by factorizing it into a product of matrices. This is of crucial importance in linear algebra of finite-dimensional vector spaces, since in a given basis any linear operator has a unique matrix representation. There are several decompositions, which are useful in different classes of problems. Here we give an overview of the most important ones, which will be used throughout this book.

#### Change of basis

An issue closely related to a certain class of matrix representation is the change of basis. In the specific, it is possible to pass from an orthonormal basis ( $|\gamma_i\rangle$ ) to another ( $|\gamma'_i\rangle$ ) by means of a so called unitary transformation  $S$ :

$$|\gamma'_i\rangle = \sum_j S_{ji} |\gamma_j\rangle \quad (i = 1, 2, \dots, n). \quad (\text{A.71})$$

This means that a generic vector

$$|\alpha\rangle = \sum_i a_i |\gamma_i\rangle \quad (a_i \equiv \langle \gamma_i | \alpha \rangle), \quad (\text{A.72})$$

can be expressed in the new basis as

$$|\alpha\rangle = \sum_j a'_j |\gamma'_j\rangle = \sum_{ij} a'_j S_{ij} |\gamma_i\rangle \quad (a'_j \equiv \langle \gamma'_j | \alpha \rangle), \quad (\text{A.73})$$

where we have used Eq. (A.71). Thus, the old and new vector components are linked by the relation

$$a_i = \sum_j S_{ij} a'_j. \quad (\text{A.74})$$

**Exercise A.5** Show that the matrix representations  $A$  and  $A'$  of an operator  $A$  with respect to the bases  $|\gamma_i\rangle$  and  $|\gamma'_i\rangle$  are connected as follows:

$$A' = S^{-1} A S. \quad (\text{A.75})$$

Among all the possible change-of-basis transformations of the type in Eq. (A.75), there is one which covers a very important role. This leads to the so called *diagonal representation* for the operator  $A$ , in which the new basis  $|\gamma'_i\rangle$  is given by the eigenvectors of  $A$ . Such transformation is based on the theory of the spectral decomposition. As we can see in Chap. 2 this decomposition can be given a physical relevance in the context of quantum mechanics. Let us start in introducing the concept of eigenvalues and eigenvectors of a square matrix.

### Eigenvalues and eigenvectors

An *eigenvector* of a linear operator  $A$  is a non-zero vector  $|\alpha\rangle$  such that

$$A|\alpha\rangle = \alpha|\alpha\rangle, \quad (\text{A.76})$$

where  $\alpha$  is a complex number called the *eigenvalue* of  $A$  corresponding to the eigenvector  $|\alpha\rangle$ . The eigenvalue equation (A.76) always has a solution. Indeed, we can expand the vectors  $|\alpha\rangle$  and  $A|\alpha\rangle$  over an orthonormal basis  $\{|\gamma_1\rangle, |\gamma_2\rangle, \dots, |\gamma_n\rangle\}$  as follows:

$$|\alpha\rangle = \sum_{i=1}^n a_i |\gamma_i\rangle \quad (a_i \equiv \langle \gamma_i | \alpha \rangle), \quad (\text{A.77})$$

$$A|\alpha\rangle = \sum_{i=1}^n c_i |\gamma_i\rangle, \quad (\text{A.78})$$

where

$$c_i = \langle \gamma_i | A | \alpha \rangle = \sum_j \langle \gamma_i | A | \gamma_j \rangle a_j = \sum_j A_{ij} a_j. \quad (\text{A.79})$$

If we insert these expansions into Eq. (A.76) we readily obtain

$$\sum_{i=1}^n \left( \sum_{j=1}^n A_{ij} a_j - \alpha a_i \right) |\gamma_i\rangle = 0, \quad (\text{A.80})$$

which is satisfied if and only if

$$\sum_{j=1}^n A_{ij} a_j - \alpha a_i = \sum_{j=1}^n (A_{ij} - \alpha \delta_{ij}) a_j = 0, \quad (i = 1, 2, \dots, n). \quad (\text{A.81})$$

This system of homogeneous linear equations has non-zero solutions if and only if the eigenvalue  $\alpha$  satisfies the *characteristic equation*

$$\det(A - \alpha I) = \det \begin{bmatrix} A_{11} - \alpha & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} - \alpha & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} - \alpha \end{bmatrix} = 0, \quad (\text{A.82})$$

where the symbol  $\det(\cdot)$  denotes the determinant of a square matrix. The solutions to the characteristic equation are the eigenvalues of the linear operator  $A$ . Since  $p(\alpha) \equiv \det(A - \alpha I)$  is a polynomial of degree  $n$ , a fundamental theorem of algebra tells us that the equation  $p(\alpha) = 0$  has  $n$  complex roots (eigenvalues)  $\alpha_1, \alpha_2, \dots, \alpha_n$ . This shows that Eq. (A.76) always has a solution. It is possible to prove that the characteristic equation depends only on the operator  $A$  and not on the specific matrix representation used for  $A$ . Therefore, the eigenvalues of a linear operator are independent of its matrix representation.

**Exercise A.6** Show that the eigenvectors of a linear operator  $A$  belonging to distinct eigenvalues are linearly independent.

### Spectral properties of Hermitian matrices

Any Hermitian matrix  $A$  presents a particular structure of its eigenvalues and eigenvectors. First of all, from the property that  $\langle \alpha | A\alpha \rangle$  is real, it can be easily shown that its eigenvalues are real. Indeed, from  $A|\alpha\rangle = \alpha|\alpha\rangle$  it follows that  $\langle \alpha | A\alpha \rangle = \alpha\langle \alpha | \alpha \rangle$  and this immediately shows that the eigenvalue  $\alpha$  has to be real.

The eigenvectors of a Hermitian operator form an orthonormal set in the Hilbert space  $\mathcal{H}$ . (It is assumed that the eigenvectors have unit norm; if not, they can be normalized by dividing them by their norms.) This property is easy to prove: assume that  $\alpha_1$  and  $\alpha_2$  are distinct eigenvalues corresponding to the eigenvectors  $|\alpha_1\rangle$  and  $|\alpha_2\rangle$ ; we have

$$\langle \alpha_j | A\alpha_i \rangle = \alpha_i \langle \alpha_j | \alpha_i \rangle, \quad (\text{A.83a})$$

$$\langle A\alpha_j | \alpha_i \rangle = \alpha_j \langle \alpha_j | \alpha_i \rangle. \quad (\text{A.83b})$$

Subtracting side-by-side (A.83b) from (A.83a), we obtain  $(\alpha_i - \alpha_j)\langle \alpha_j | \alpha_i \rangle = 0$  and, since  $\alpha_i \neq \alpha_j$ , we obtain  $\langle \alpha_j | \alpha_i \rangle = 0$ . Here we have assumed that the eigenvalues  $\alpha_i$  are not degenerate, that is,  $\alpha_i \neq \alpha_j$  for  $i \neq j$ . However, it can be shown that it is also possible to construct an orthonormal set of eigenvectors of the operator  $A$  in the degenerate case, in which there are linearly independent eigenvectors corresponding to the same eigenvalue. In summary, given an Hermitian operator  $A$ , it is always possible to construct an orthonormal basis of eigenvectors of  $A$ . Therefore, any vector in the Hilbert space  $\mathcal{H}$  can be expressed as a linear superposition of vectors of this basis. This property is called *completeness* and the basis of eigenvectors of  $A$  is said to be a *complete orthonormal set*.

With respect to this basis, the matrix representation of  $A$  is diagonal, and reads as follows:

$$A = \sum_{i=1}^n \lambda_i |i\rangle\langle i|, \quad (\text{A.84})$$

where  $\lambda_i$  are the eigenvalues of  $A$  and  $|i\rangle$  the corresponding eigenvectors. We call Eq. (A.84) the *spectral decomposition* of the Hermitian operator  $A$  (the ensemble of the eigenvalues of  $A$  constitutes its *spectrum*).

An Hermitian  $x \times n$  matrix  $A$  is said to be *positive definite* if

$$x^\dagger Ax > 0 \quad \forall x \in \mathbb{C}^n \quad (x \neq 0). \quad (\text{A.85})$$

It is immediate to see that condition (A.85) is equal to the requirement of having all the eigenvalues  $\lambda_i$  being positive. In a similar way one can define a notion of *negative definite*, positive semi-definite, and negative semi-definite Hermitian matrices, except that the previous expression is required to be always negative, non-negative, and non-positive, respectively.

#### A.1.6.1 Diagonalization and spectral decomposition

In general, an operator  $A$  is said to be *diagonalizable* if it has a diagonal representation. This amounts to say that it can be written as

$$D = P^{-1}AP, \quad (\text{A.86})$$

where  $D$  is a diagonal matrix, while  $P$  is an invertible matrix. In passing we note that this property can be used to compute powers of  $A$  very efficiently. Indeed, using the associativity of matrix product, we have:

$$A^k = (PDP^{-1})^k = (PDP^{-1}) \cdot (PDP^{-1}) \cdots (PDP^{-1}) = PD^k P^{-1}, \quad (\text{A.87})$$

which is easy to calculate since it only involves powers of a diagonal matrix. This approach enables to calculate complex matrix functions which are defined as power series, including the matrix exponential (A.57).

If the operator  $A$  is Hermitian, the decomposition (A.86) coincides exactly with the spectral decomposition of Eq. (A.84). In such case, the matrix  $P$  is unitary and contains the change-of-basis elements which enable one to pass into the diagonal basis of the eigenvectors, see Eq. (A.71). However there are operators that are not diagonalizable, such as the operator with matrix representation

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \quad (\text{A.88})$$

In this example there is only one eigenvalue,  $\lambda = 1$ , and the corresponding eigenvector

$$|u\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (\text{A.89})$$

spans a one-dimensional subspace and therefore cannot be a basis for the two-dimensional vector space on which the matrix (A.88) operates.

It is possible to show that both Hermitian and unitary operators are diagonalizable. Moreover, in such cases the invertible matrix  $P$  is also unitary. These operators belong to the larger class of *normal operators*, defined by the condition

$$AA^\dagger = A^\dagger A. \quad (\text{A.90})$$

We now state without proof the central theorem of the spectral theory:

**Theorem A.1** Spectral decomposition theorem. *A linear operator  $A$  is diagonalizable with an orthonormal basis of eigenvectors, that is, it exists a unitary matrix  $U$  such that*

$$A = UDU^\dagger, \quad (\text{A.91})$$

*if and only if it is normal.*

Example: Pauli matrices

An example of a diagonal representation is the Pauli matrix

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad (\text{A.92})$$

which is diagonal with respect to the eigenvector basis

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (\text{A.93})$$

where the eigenvectors  $|0\rangle$  and  $|1\rangle$  correspond to the eigenvalues +1 and -1, respectively. In the representation  $\{|0\rangle, |1\rangle\}$ , the Pauli matrix  $\sigma_x$  reads

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = |0\rangle\langle 1| + |1\rangle\langle 0|. \quad (\text{A.94})$$

The operator  $\sigma_x$  thus is diagonal in the basis

$$|+\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |-\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad (\text{A.95})$$

in which its matrix representation is given by

$$\sigma_x = |+\rangle\langle +| - |-\rangle\langle -. \quad (\text{A.96})$$

The new basis  $\{|+\rangle, |-\rangle\}$  is related to the original basis  $\{|0\rangle, |1\rangle\}$  by means of the unitary transformation

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (\text{A.97})$$

This follows directly from the fact that the Pauli matrices are Hermitian.

**Exercise A.7** Write down the Pauli matrices in the basis  $\{|+\rangle, |-\rangle\}$ .

### Commutator

We say that two operators  $A$  and  $B$  commute if they satisfy the following relation:

$$AB = BA. \quad (\text{A.98})$$

The *commutator* of two operators  $A$  and  $B$  is defined by

$$[A, B] = AB - BA. \quad (\text{A.99})$$

It is easy to check the following properties:

$$[A, A] = 0, \quad (\text{A.100a})$$

$$[A, B] = -[B, A], \quad (\text{A.100b})$$

$$[A, BC] = [A, B]C + B[A, C], \quad (\text{A.100c})$$

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0, \quad (\text{Jacobi identity}). \quad (\text{A.100d})$$

Another very useful identity involving nested commutators is the Baker-Campbell-Hausdorff expansion:

$$e^A B e^{-A} = \sum_{m=0}^{+\infty} \frac{1}{m!} B_m, \quad (\text{A.101})$$

where  $B_m = [A, B]_m$  is defined recursively as  $B_m = [A, [A, B]_{m-1}]$  and  $B_0 = B$ .

**Exercise A.8** Show that, if  $A$  and  $B$  are Hermitian,  $i[A, B]$  is Hermitian.

### Anti-commutator

The anti-commutator of two operators  $A$  and  $B$  is defined by

$$\{A, B\} = AB + BA. \quad (\text{A.102})$$

We say that two operators  $A$  and  $B$  *anti-commute* if  $\{A, B\} = 0$ . The following properties hold:

$$\{A, B\} = \{B, A\}, \quad (\text{A.103a})$$

$$\{A, BC\} = \{A, B\}C - B[A, C]. \quad (\text{A.103b})$$

It is easy to verify that the Pauli matrices anti-commute,

$$\{\sigma_i, \sigma_j\} = 0, \quad (i, j = x, y, z), \quad (\text{A.104})$$

while the following commutation relations hold:

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (\text{A.105})$$

Putting all these relations together, we obtain the following:

$$\sigma_j \sigma_k = i\epsilon_{jkl}\sigma_l + \delta_{jk}I, \quad (j, k, l = x, y, z). \quad (\text{A.106})$$

**Theorem A.2** Simultaneous diagonalization theorem: Two normal operators  $A$  and  $B$  commute if and only if there exists an orthonormal basis with respect to which both  $A$  and  $B$  are diagonal.

**Proof.** Assume that  $|i\rangle$  is an orthonormal basis for both  $A$  and  $B$ , that is,

$$A|i\rangle = \lambda_i|i\rangle, \quad B|i\rangle = \nu_i|i\rangle. \quad (\text{A.107})$$

Therefore,

$$AB|i\rangle = A\nu_i|i\rangle = \lambda_i\nu_i|i\rangle = \nu_i\lambda_i|i\rangle = BA|i\rangle. \quad (\text{A.108})$$

Thus,  $[A, B] = 0$ . To show the converse, let us denote by  $|i\rangle$  an orthonormal basis for the operator  $A$  with eigenvalues  $\lambda_i$ . Assume that the vectors  $|i\rangle$  are not eigenfunctions of the operator  $B$  and expand  $B|i\rangle$  on the  $|i\rangle$  basis:

$$B|i\rangle = \sum_{j=1}^n \langle j|B|i\rangle |j\rangle. \quad (\text{A.109})$$

Therefore,

$$[A, B]|i\rangle = AB|i\rangle - BA|i\rangle = \sum_{j=1}^n \langle j|B|i\rangle (\lambda_j - \lambda_i) |j\rangle = 0. \quad (\text{A.110})$$

Here we have  $[A, B]|i\rangle = 0$ , since we assume that  $A$  and  $B$  commute. If the eigenvalues are not degenerate, that is,  $\lambda_i \neq \lambda_j$  for  $i \neq j$ , then from Eq. (A.110) we obtain

$$\langle j|B|i\rangle = 0 \quad \text{for } i \neq j. \quad (\text{A.111})$$

If we denote  $\langle j|B|j\rangle = \nu_j$ , we have

$$\langle j|B|i\rangle = \nu_j \delta_{ij}. \quad (\text{A.112})$$

Inserting this relation into Eq. (A.109), we obtain

$$B|i\rangle = \nu_i|i\rangle. \quad (\text{A.113})$$

Therefore,  $|i\rangle$  is also an eigenvector of the operator  $B$ . Note that the proof can be extended to the case in which there are degeneracies in the eigenvalues  $\lambda_i$ .  $\square$

### Trace

The trace of a matrix  $A$  is defined as the sum of its diagonal elements:

$$\text{Tr}(A) = \sum_{i=1}^n A_{ii}. \quad (\text{A.114})$$

It is easy to check the following properties:

$$(i) \quad \text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B) \quad (\text{linearity}), \quad (\text{A.115a})$$

$$(ii) \quad \text{Tr}(cA) = c\text{Tr}(A) \quad (c \in \mathbb{C}), \quad (\text{A.115b})$$

$$(iii) \quad \text{Tr}(AB) = \text{Tr}(BA) \quad (\text{cyclic property}). \quad (\text{A.115c})$$

As a consequence of the cyclic property, for  $n$  operators  $A_1, A_2, \dots, A_n$ , we have

$$\text{Tr}(A_1 A_2 \cdots A_{n-1} A_n) = \text{Tr}(A_2 A_3 \cdots A_n A_1) = \dots = \text{Tr}(A_n A_1 \cdots A_{n-2} A_{n-1}). \quad (\text{A.116})$$

The trace of an operator  $A$  is defined as the trace of a matrix representation of  $A$ . It is easy to check that the trace is independent of the choice of representation. Indeed, from the relation  $\sum_j |j\rangle\langle j| = I$ , we obtain

$$\text{Tr}(A) = \sum_i \langle i | A | i \rangle = \sum_i \sum_j \sum_k \langle i | j \rangle \langle j | A | k \rangle \langle k | i \rangle = \sum_j \langle j | A | j \rangle. \quad (\text{A.117})$$

Alternatively, from property (iii) it follows that, for a unitary operator  $U$ ,

$$\text{Tr}(U^\dagger A U) = \text{Tr}(U U^\dagger A) = \text{Tr}(A), \quad (\text{A.118})$$

and therefore the trace is invariant under unitary transformations.

Another important property is the following: let  $|i\rangle$  be an orthonormal basis and  $|\alpha\rangle$  a generic vector, which we can expand over the  $|i\rangle$  basis as follows:  $|\alpha\rangle = \sum_i \langle i | \alpha \rangle |i\rangle$ . Then

$$\text{Tr}(A|\alpha\rangle\langle\alpha|) = \sum_i \langle i | A | \alpha \rangle \langle \alpha | i \rangle = \sum_i \langle \alpha | i \rangle \langle i | A | \alpha \rangle = \langle \alpha | A | \alpha \rangle. \quad (\text{A.119})$$

### A.1.6.2 Singular value decomposition

There are other matrix representations which reveal useful in different contexts. One of them can be thought as a generalization of the spectral decomposition (A.91) to the case of non-normal matrices. A generic  $m \times n$  complex matrix  $M$  can be always factorized according to the following:

$$M = W \Sigma V^\dagger, \quad (\text{A.120})$$

where  $W$  is a  $m \times m$  unitary matrix,  $\Sigma$  is a  $m \times n$  (rectangular) diagonal matrix (i.e., a matrix with only the entries  $\Sigma_{ii}$  on the diagonal possibly non-zero) with non-negative real numbers on the diagonal, while  $V^\dagger$  is the conjugate transpose of the  $n \times n$  unitary matrix  $V$ . The diagonal entries  $\sigma_j$  of  $\Sigma$ , with  $j = 1, \dots, \min(m, n)$ , denote the *singular values* of  $M$ . While  $\Sigma$  is uniquely determined by  $M$ , the unitary matrices  $W$  and  $V$  are not. The decomposition in Eq. (A.120) is called the *singular value decomposition* (often shortened as SVD) of the matrix  $M$ .

The SVD requires a weaker condition than the spectral decomposition, however the two decompositions are closely related. Since the matrices  $W$  and  $V^\dagger$  are unitary, their columns form a set of orthonormal vectors, which can be regarded as basis vectors. The same is true for their conjugate transposes. The difference with Eq. (A.91) is that here the relative left and right bases are different. Given the representation in Eq. (A.120), it is indeed easy to see that

$$M^\dagger M = V(\Sigma^\dagger \Sigma)V^\dagger \quad \text{and} \quad MM^\dagger = W(\Sigma\Sigma^\dagger)W^\dagger. \quad (\text{A.121})$$

These equalities show that the columns of  $V$  are eigenvectors of  $M^\dagger M$  (also called left singular vectors), while those of  $U$  are eigenvectors of  $MM^\dagger$  (also called right singular vectors). The non-zero elements of  $\Sigma$  are the square roots of the non-zero eigenvalues of  $M^\dagger M$  or of  $MM^\dagger$  (notice that these two matrices are Hermitian and positive semi-definite).

Let us now focus, for simplicity, on square matrices. In the case in which  $M$  is normal, its spectral decomposition states that it can be written as  $M = UDU^\dagger$ . This form coincides with the SVD, if  $M$  is positive semi-definite. However the diagonal decomposition and the SVD differ for all the other diagonalizable matrices. Indeed, the matrix  $P$  in Eq. (A.86) is not necessarily unitary and the eigenvalues are not necessarily non-negative. On the contrary, in (A.120)  $\sigma_j$  are non-negative, while  $W$  and  $V$  are unitary matrices not necessarily related, except through the matrix  $M$ .

#### A.1.6.3 Polar decomposition

A generic complex square matrix  $M$  can be always written in the form

$$M = UP, \quad (\text{A.122})$$

where  $U$  is a unitary matrix and  $P$  is, in general, a positive semi-definite Hermitian matrix, given by

$$P = \sqrt{M^\dagger M}. \quad (\text{A.123})$$

Sometimes the following notation is also used:  $P = |M|$ . The square root introduced here naturally extends the definition from complex numbers to matrices. (A given matrix  $B$  is said to be the square root of  $A$  if and only if  $BB = A$ .) This is always possible for positive semi-definite matrices.

**Exercise A.9** Show that the polar decomposition (A.122) is possible for any real matrix  $M$ .

Alternatively, the matrix  $M$  can be also decomposed as

$$M = P'U, \quad (\text{A.124})$$

where  $U$  is the same as before, while  $P' = UPU^{-1} = \sqrt{MM^\dagger}$ . This is known to be the left polar decomposition, while Eq. (A.122) is also called the right polar decomposition. We recall that, in general,  $P \neq P'$ , and  $|M| \neq \sqrt{MM^\dagger}$ . In the specific it can be shown that  $P = P'$  if and only if the matrix  $M$  is normal.

In terms of the SVD in Eq. (A.120), we can rewrite the polar decomposition such that  $P = V\Sigma V^\dagger$ ,  $P' = W\Sigma W^\dagger$  and  $U = WV^\dagger$ . This clearly shows that the existence of the SVD is equivalent to the existence of the polar decomposition.

#### A.1.6.4 Cholesky decomposition

A positive definite Hermitian matrix  $A$  can be decomposed into the product of a lower triangular matrix and its conjugate transpose, that is,

$$A = LL^\dagger, \quad (\text{A.125})$$

where  $L$  is a matrix in which all the entries above the main diagonal are zero, while the diagonal entries are real and positive. Conversely, if a given matrix  $A$  can be written as (A.125) for some given invertible  $L$ , then  $A$  is positive definite Hermitian. The Cholesky decomposition of a Hermitian positive definite matrix is unique.

If the matrix  $A$  is positive semi-definite, it is still possible to define an analogous decomposition, in which the diagonal entries of  $L$  are allowed to be zero. In such circumstance the decomposition needs not be unique.

#### A.1.6.5 QR decomposition

The QR decomposition applies to any complex  $m \times n$  ( $m \geq n$ ) matrix  $A$ , and states that

$$A = QR, \quad (\text{A.126})$$

where  $Q$  is a  $m \times m$  unitary matrix and  $R$  is an  $m \times n$  upper triangular matrix (that is, a matrix where the entries below the main diagonal –going from  $R_{11}$  to  $R_{nn}$ – are zero). If  $A$  is a positive definite square matrix, then  $R$  is equal to the upper triangular factor of the Cholesky decomposition of  $A^\dagger A$ .

In a similar way, it is possible to define QL, RQ, and LQ decompositions, with  $L$  being a lower triangular matrix.

#### A.1.7 Symplectic decompositions

A given  $2N \times 2N$  matrix  $S$  over the field  $\mathcal{F}$  is said to be *symplectic* if the following condition holds:

$$S \Omega S^T = \Omega, \quad (\text{A.127})$$

where  $\Omega$  is a  $2N \times 2N$  non-singular skew-symmetric block matrix of the form

$$\Omega = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix}. \quad (\text{A.128})$$

Here  $I_N$  denotes the  $N \times N$  identity over the field  $\mathcal{F}$  (for the sake of simplicity, in the following we will consider the field of reals  $\mathbb{R}$ ). From Eq. (A.127) it is straightforward to see that  $\det(S) = \pm 1$ . However, it can be further shown that every symplectic matrix  $S$  has unit determinant (Meyer *et al.*, 2009), and its inverse is given by

$$S^{-1} = \Omega^{-1} S^T \Omega. \quad (\text{A.129})$$

Another useful property of symplectic matrices is the following. Suppose  $S$  is written in a generic block-matrix form as

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad (\text{A.130})$$

with  $A, B, C, D$  being  $N \times N$  matrices. Then  $S$  is symplectic if and only if the following conditions hold:

$$AD^T - BC^T = I_N, \quad AB^T = BA^T, \quad CD^T = DC^T. \quad (\text{A.131})$$

Finally we remind that, since the product of two symplectic matrices  $T = S_1 S_2$  is again a symplectic matrix, they form a group. This is denoted by  $\mathrm{Sp}(2N)$ , and is called the *symplectic group*.

In the following we quote two theorems which turn out to be very useful in the manipulation and decomposition of symplectic matrices. The interested reader can find a more comprehensive treatment of this subject in Meyer *et al.* (2009).

### A.1.7.1 Euler's decomposition

A generic symplectic matrix  $S$  can be decomposed according to

$$S = O \begin{bmatrix} D & 0 \\ 0 & D^{-1} \end{bmatrix} O', \quad (\text{A.132})$$

where the matrices  $O$  and  $O'$  are orthogonal and symplectic themselves, while  $D$  is diagonal and positive definite. The decomposition (A.132) is closely related to the singular value decomposition of Sec. A.1.6.2, and is known as the *Euler decomposition* of a symplectic matrix.

### A.1.7.2 Williamson's theorem

Given an arbitrary positive definite  $2N \times 2N$  real matrix  $M$ , there always exists a symplectic matrix  $S$  which diagonalizes  $M$  according to

$$M = S \bigoplus_{j=1}^N \begin{bmatrix} \nu_j & 0 \\ 0 & \nu_j \end{bmatrix} S^T, \quad (\text{A.133})$$

where  $\nu_j > 0$  are called the *symplectic eigenvalues* of  $M$ . The symplectic spectrum can be computed as the modulus of the  $2N$  real eigenvalues of the matrix  $i\Omega M$ , where  $\Omega$  is given by Eq. (A.128).

## A.2 Infinite-dimensional vector spaces

### A.2.1 Discrete and continuous bases

Infinite-dimensional Hilbert spaces are used in quantum mechanics even in the simple example of a single particle moving along a line. The wave function  $\psi(x)$  describing such a system resides in the infinite-dimensional Hilbert space of the *square-integrable functions*  $\mathcal{L}^2(\mathbb{R})$ , namely in the space of the functions  $\psi(x)$  such that

$$\int_{-\infty}^{+\infty} dx |\psi(x)|^2 < +\infty, \quad x \in (-\infty, +\infty). \quad (\text{A.134})$$

In this space the inner product of two functions  $\psi_1(x)$  and  $\psi_2(x)$  is defined as  $\int_{-\infty}^{+\infty} dx \psi_1^*(x) \psi_2(x)$ . In Dirac's notation,  $|\psi\rangle$  denotes the ket (vector) associated to the square-integrable function  $\psi(x)$ . The inner product of  $|\psi_1\rangle$  and  $|\psi_2\rangle$  reads

$$\langle \psi_1 | \psi_2 \rangle = \int_{-\infty}^{+\infty} dx \psi_1^*(x) \psi_2(x), \quad (\text{A.135})$$

and the norm of a vector  $|\psi\rangle$  is given by  $\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle}$ .

The basic tools of linear algebra (such as linear operators, orthonormal bases, the completeness relation, tensor products,...) can be extended to infinite-dimensional spaces. However, it is convenient to introduce bases of vectors not belonging to the vector space, but in terms of which any vector residing in the vector space can nevertheless be expanded. To provide a concrete example, in  $\mathcal{L}^2(\mathbb{R})$  we can consider the Fourier transform  $\tilde{\psi}(k)$  of a square-integrable function  $\psi(x)$ :

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dk \tilde{\psi}(k) e^{ikx}, \quad (\text{A.136})$$

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx \psi(x) e^{-ikx}. \quad (\text{A.137})$$

By introducing the plane wave of wave vector  $k$ ,

$$v_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad (\text{A.138})$$

we can rewrite the above formulas for the Fourier transform as

$$|\psi\rangle = \int_{-\infty}^{+\infty} dk \tilde{\psi}(k) |v_k\rangle, \quad (\text{A.139})$$

$$\tilde{\psi}(k) = \langle v_k | \psi \rangle = \int_{-\infty}^{+\infty} dx v_k^*(x) \psi(x). \quad (\text{A.140})$$

That is, we can expand the vector  $|\psi\rangle$  over the set of the plane waves  $\{|v_k\rangle\}$ , where  $k \in (-\infty, +\infty)$  and the coefficients of the expansion are given by  $\tilde{\psi}(k) = \langle v_k | \psi \rangle$ . Eqs. (A.139) and (A.140) are similar to (A.18) and (A.19). However, the index  $k$  is continuous rather than discrete and the functions  $v_k(x)$  are not square-integrable,

since  $|v_k(x)|^2 = \frac{1}{2\pi}$  diverges when integrated over the whole  $x$  axis. Nevertheless, we say that the set  $\{|v_k\rangle\}$  is a basis (in a generalized sense) since we can expand in a unique way over it any square-integrable function. We also say that  $\{|v_k\rangle\}$  is a *continuous basis*, since  $k$  is a continuous index, while a basis  $\{|\alpha_i\rangle\}$  of vectors residing in the vector space is a *discrete basis*, since the index  $i$  is discrete. For a continuous basis, we express the coefficients of the expansion as  $\langle v_k | \psi \rangle$ , namely we extend the definition of scalar product to functions which are not square-integrable, provided  $\int_{-\infty}^{+\infty} dx v_k^*(x) \psi(x)$  converges.

### A.2.2 The Dirac delta function

In order to discuss orthonormality and the completeness relation in infinite-dimensional spaces, we need to introduce the Dirac delta “function”. We first define the function  $\delta^{(\epsilon)}(x)$ , given by  $\delta^{(\epsilon)}(x) = \frac{1}{\epsilon}$  for  $-\frac{\epsilon}{2} < x < \frac{\epsilon}{2}$ ,  $\delta^{(\epsilon)}(x) = 0$  otherwise, where  $\epsilon$  is a positive number. Given an arbitrary function  $f(x)$ , well defined for  $x = 0$ , for  $\epsilon$  sufficiently small to neglect the variation of  $f(x)$  over the interval  $[-\frac{\epsilon}{2}, \frac{\epsilon}{2}]$  we have

$$\int_{-\infty}^{+\infty} dx \delta^{(\epsilon)}(x) f(x) \approx f(0) \int_{-\infty}^{+\infty} dx \delta^{(\epsilon)}(x) = f(0). \quad (\text{A.141})$$

The approximation is better the smaller  $\epsilon$ . In the limit  $\epsilon \rightarrow 0$  we define the Dirac  $\delta$  function by the relation

$$\int_{-\infty}^{+\infty} dx \delta(x) f(x) = f(0). \quad (\text{A.142})$$

More generally,  $\delta(x - x_0)$  is defined by

$$\int_{-\infty}^{+\infty} dx \delta(x - x_0) f(x) = f(x_0). \quad (\text{A.143})$$

Of course, the above definition of the delta function is not mathematically rigorous, since  $\delta^{(\epsilon)}(x - x_0)$  diverges in  $x_0$  when  $\epsilon \rightarrow 0$ . The  $\delta$  function is actually not a function but a distribution. However, for the purposes of the present book, it will be sufficient to consider  $\delta$  as a function. Moreover, it is physically impossible to distinguish between the distribution  $\delta$  and a function  $\delta^{(\epsilon)}$  approximating  $\delta$ , provided  $\epsilon$  is much smaller than the scales available to experimental investigations.

For completeness, we report below the main properties of the  $\delta$  function, which can be easily proved from its definition (A.143):

$$\int_a^b dx \delta(x) f(x - x_0) = \begin{cases} f(x_0) & \text{if } x_0 \in [a, b], \\ 0 & \text{if } x_0 \notin [a, b], \end{cases} \quad (\text{A.144})$$

$$\delta(-x) = \delta(x), \quad (\text{A.145})$$

$$\delta(cx) = \frac{1}{|c|} \delta(x), \quad (\text{A.146})$$

$$\delta[g(x)] = \sum_j \frac{1}{|g'(x_j)|} \delta(x - x_j), \quad (\text{A.147})$$

provided  $g$  is a continuously differentiable function with its derivative  $g'(x)$  nowhere zero and the  $x_j$  are the zeros of the function  $g(x)$ . It is clear that (A.146) is a special instance of (A.147). Another property of the  $\delta$  function is

$$x\delta(x - x_0) = x_0\delta(x - x_0), \quad (\text{A.148})$$

and more generally

$$g(x)\delta(x - x_0) = g(x_0)\delta(x - x_0). \quad (\text{A.149})$$

We also have

$$\int_{-\infty}^{+\infty} dx \delta(x - y)\delta(x - z) = \delta(y - z). \quad (\text{A.150})$$

In three dimensions, the Dirac delta function is defined by

$$\int_{\mathbb{R}^3} d^3r \delta(\mathbf{r} - \mathbf{r}_0)f(\mathbf{r}) = f(\mathbf{r}_0), \quad (\text{A.151})$$

where  $d^3r = dx dy dz$ . We have

$$\delta(\mathbf{r} - \mathbf{r}_0) = \delta(x - x_0)\delta(y - y_0)\delta(z - z_0), \quad (\text{A.152})$$

where  $\mathbf{r} = (x, y, z)$  and  $\mathbf{r}_0 = (x_0, y_0, z_0)$ .

### A.2.3 Orthonormality and completeness relations

We can generalize the concept of an orthonormal basis to the continuous case. In the example of plane waves in  $\mathcal{L}^2(\mathbb{R})$  we have

$$\langle v_k | v_{k'} \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{i(k' - k)x} = \delta(k - k'), \quad (\text{A.153})$$

where  $\delta$  is the Dirac delta function. Instead of the Kroenecker delta,  $\delta_{ij}$ , as in an orthonormal discrete basis,  $\langle \alpha_i | \alpha_j \rangle = \delta_{ij}$ , in (A.153) we have two continuous indices  $k$  and  $k'$  and a Dirac delta function of the difference of the two indices,  $\delta(k - k')$ . Namely,  $\langle v_k | v_{k'} \rangle = 0$  if  $k \neq k'$ , while  $\langle v_k | v_k \rangle$  diverges, as expected since  $|v_k\rangle$  is not square-integrable. In spite of this divergence, it is customary to refer to (A.153) as orthonormality relation, or to say that the  $|v_k\rangle$  are orthonormalized in the Dirac sense.

The completeness relation for the basis  $\{|v_k\rangle\}$  reads

$$\int_{-\infty}^{+\infty} dk |v_k\rangle \langle v_k| = I. \quad (\text{A.154})$$

Indeed, for a generic  $|\psi\rangle \in \mathcal{L}^2(\mathbb{R})$  we obtain from Eqs. (A.139) and (A.140)

$$\left( \int_{-\infty}^{+\infty} dk |v_k\rangle \langle v_k| \right) |\psi\rangle = \int_{-\infty}^{+\infty} dk |v_k\rangle \langle v_k| \psi \rangle = \int_{-\infty}^{+\infty} dk \tilde{\psi}(k) |v_k\rangle = |\psi\rangle. \quad (\text{A.155})$$

More generally, we can consider the case of a mixed (discrete and continuous) basis  $\{|\alpha_i\rangle, |v_k\rangle\}$ , with  $i$  discrete index ( $i = 1, 2, \dots, i_{\max}$ , possibly with  $i_{\max} = \infty$ ) and  $k$  continuous index [ $k \in (-\infty, +\infty)$  or in some finite interval]. The orthonormalization relations are

$$\langle \alpha_i | \alpha_j \rangle = \delta_{ij}, \quad \langle v_k | v_{k'} \rangle = \delta(k - k'), \quad \langle \alpha_i | v_k \rangle = 0. \quad (\text{A.156})$$

The completeness relation reads

$$\sum_i |\alpha_i\rangle \langle \alpha_i| + \int dk |v_k\rangle \langle v_k| = I. \quad (\text{A.157})$$

In  $\mathcal{L}^2(\mathbb{R})$  the completeness relation can be written as

$$\sum_i \alpha_i(x) \alpha_i^*(x') + \int dk v_k(x) v_k^*(x') = \delta(x - x'). \quad (\text{A.158})$$

To prove this relation, let us consider a generic square-integrable vector  $|\psi\rangle$ . We have  $|\psi\rangle = I|\psi\rangle$ , and therefore

$$|\psi\rangle = \sum_i |\alpha_i\rangle \langle \alpha_i| \psi + \int dk |v_k\rangle \langle v_k| \psi. \quad (\text{A.159})$$

By inserting the explicit expressions of the inner products,

$$\langle \alpha_i | \psi \rangle = \int_{-\infty}^{+\infty} dx' \alpha_i^*(x') \psi(x'), \quad \langle v_k | \psi \rangle = \int_{-\infty}^{+\infty} dx' v_k^*(x') \psi(x'), \quad (\text{A.160})$$

we obtain

$$\psi(x) = \int_{-\infty}^{+\infty} dx' \left[ \sum_i \alpha_i^*(x') \alpha_i(x) + \int dk v_k^*(x') v_k(x) \right]. \quad (\text{A.161})$$

By comparing this latter equation with the definition (A.143) of the delta function we obtain the completeness relation (A.158). Note that this relation involves a delta function also in the case of a purely discrete basis  $\{|\alpha_i\rangle\}$ . This should not be a surprise if we keep in mind that we are considering infinite-dimensional spaces. Finally, we write the completeness relation also in the case of  $\mathcal{L}^2(\mathbb{R}^3)$ , namely of the square-integrable functions in the three-dimensional space. Given a generic mixed basis  $\{|\alpha_i\rangle, |v_k\rangle\}$ , we have

$$\sum_i \alpha_i(\mathbf{r}) \alpha_i^*(\mathbf{r}') + \int dk v_k(\mathbf{r}) v_k^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'). \quad (\text{A.162})$$

#### A.2.4 Position and momentum representations

We say that we choose a representation is we choose an orthonormal basis, either discrete or continuous, in the Hilbert space. Hereafter we focus on two representations, physically associated with the position and momentum of a particle in one dimension (the extension to two or three dimensions is straightforward).

Let us first consider the basis of plane waves (A.138):

$$v_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}, \quad (\text{A.163})$$

where  $p = \hbar k$  is interpreted as the momentum of the particle under consideration. We denote  $|p\rangle$  the ket associated to the function  $v_p(x)$ , and call the continuous basis  $\{|p\rangle\}$  [ $p \in (-\infty, \infty)$ ] *momentum basis*.

The *position basis* is composed of the delta “functions”

$$\xi_{x_0}(x) = \delta(x - x_0). \quad (\text{A.164})$$

The ket associated to  $\xi_{x_0}(x)$  is denoted by  $|x_0\rangle$ . Given a ket  $|\psi\rangle$ , we can write

$$\psi(x) = \int_{-\infty}^{+\infty} dx_0 \psi(x_0) \delta(x - x_0) = \int_{-\infty}^{+\infty} dx_0 \psi(x_0) \xi_{x_0}(x); \quad (\text{A.165})$$

in Dirac's notation

$$|\psi\rangle = \int_{-\infty}^{+\infty} dx_0 \psi(x_0) |x_0\rangle, \quad (\text{A.166})$$

$$\psi(x_0) = \langle x_0 | \psi \rangle = \int_{-\infty}^{+\infty} dx \xi_{x_0}^*(x) \psi(x). \quad (\text{A.167})$$

That is, we can expand  $|\psi\rangle$  over the basis  $|x_0\rangle$ , the coefficients of the expansion being the values  $\psi(x_0)$  of the function  $\psi(x)$  at the points  $x_0$ . The orthonormality condition for the basis  $\{|x_0\rangle\}$  follows from property (A.150) of the delta function:

$$\langle x_0 | x'_0 \rangle = \int_{-\infty}^{+\infty} dx \xi_{x_0}^*(x) \xi_{x'_0}(x) = \int_{-\infty}^{+\infty} dx \delta(x - x_0) \delta(x - x'_0) = \delta(x_0 - x'_0). \quad (\text{A.168})$$

The completeness relation reads

$$\int_{-\infty}^{+\infty} dx_0 |x_0\rangle \langle x_0| = I. \quad (\text{A.169})$$

From Eqs. (A.167) and (A.140) we can write the wave function in position and momentum representation by means of the components of the ket  $|\psi\rangle$  on the bases  $\{|x\rangle\}$  and  $\{|p\rangle\}$ , respectively. We have

$$\psi(x) = \langle x | \psi \rangle, \quad \tilde{\psi}(p) = \langle p | \psi \rangle. \quad (\text{A.170})$$

The change from the  $\{|x\rangle\}$  to the  $\{|p\rangle\}$  representation is obtained by means of the completeness relations for the two bases,

$$\int_{-\infty}^{+\infty} dx |x\rangle \langle x| = I, \quad \int_{-\infty}^{+\infty} dp |p\rangle \langle p| = I, \quad (\text{A.171})$$

and of the expression of the plane waves in the position basis:

$$\langle x | p \rangle = \langle p | x \rangle^* = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}. \quad (\text{A.172})$$

We recover the fact that  $\psi(x)$  and  $\tilde{\psi}(p)$  are related by means of the Fourier transform. Indeed

$$\langle x|\psi \rangle = \int_{-\infty}^{+\infty} dp \langle x|p\rangle \langle p|\psi \rangle, \quad (\text{A.173})$$

namely

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dp e^{ipx/\hbar} \tilde{\psi}(p); \quad (\text{A.174})$$

$$\langle p|\psi \rangle = \int_{-\infty}^{+\infty} dx \langle p|x\rangle \langle x|\psi \rangle, \quad (\text{A.175})$$

namely

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx e^{-ipx/\hbar} \psi(x). \quad (\text{A.176})$$

The inner product of two vectors  $|\psi_1\rangle$  and  $|\psi_2\rangle$  reads, in the position representation,

$$\langle \psi_1|\psi_2 \rangle = \int_{-\infty}^{+\infty} dx \langle \psi_1|x\rangle \langle x|\psi_2 \rangle = \int_{-\infty}^{+\infty} dx \psi_1^*(x) \psi_2(x), \quad (\text{A.177})$$

while in the momentum representation we have

$$\langle \psi_1|\psi_2 \rangle = \int_{-\infty}^{+\infty} dp \langle \psi_1|p\rangle \langle p|\psi_2 \rangle = \int_{-\infty}^{+\infty} dp \tilde{\psi}_1^*(p) \tilde{\psi}_2(p). \quad (\text{A.178})$$

In particular, if  $|\psi\rangle = |\psi_2\rangle \equiv |\psi\rangle$ , we recover the Parseval-Plancherel formula

$$\int_{-\infty}^{+\infty} dx |\psi_1(x)|^2 = \int_{-\infty}^{+\infty} dp |\tilde{\psi}_1(p)|^2, \quad (\text{A.179})$$

meaning that a function and its Fourier transform have the same norm.

We can change representation also for operators. If  $A(x',x) = \langle x'|A|x\rangle$  and  $A(p',p) = \langle p'|A|p\rangle$  denote the position and momentum representations of an operator  $A$ , by inserting twice the completeness relation  $\int_{-\infty}^{+\infty} dx |x\rangle \langle x| = I$  we obtain

$$\begin{aligned} A(p',p) &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' \langle p'|x'\rangle \langle x'|A|x\rangle \langle x|p\rangle \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' e^{i(px-p'x')/\hbar} A(x',x). \end{aligned} \quad (\text{A.180})$$

We derive similarly the formula to compute  $A(x',x)$  from  $A(p',p)$ :

$$A(x',x) = \int_{-\infty}^{+\infty} dp \int_{-\infty}^{+\infty} dp' \langle x'|p'\rangle \langle p'|A|p\rangle \langle p|x\rangle \quad (\text{A.181})$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dp \int_{-\infty}^{+\infty} dp' e^{-i(px-p'x')/\hbar} A(p',p). \quad (\text{A.182})$$

### A.2.5 Position and momentum operators

As discussed in Sec. 2.3, in quantum mechanics physical observables such as position, momentum, angular momentum and spin are represented by self-adjoint operators on a Hilbert space.<sup>1</sup> We discuss hereafter the position and the momentum operators.

The position operator  $X$  maps any ket  $|\psi\rangle$  into  $|\psi'\rangle = X|\psi\rangle$ , with

$$\langle x|\psi'\rangle = \langle x|X|\psi\rangle = x\langle x|\psi\rangle, \quad (\text{A.183})$$

namely in the  $\{|x\rangle\}$  representation the  $X$  operator coincides with the multiplication of the wave function by  $x$ :

$$\psi'(x) = x\psi(x). \quad (\text{A.184})$$

Quantities such as  $\langle\psi_1|X|\psi_2\rangle$  are conveniently computed in the position representation:

$$\langle\psi_1|X|\psi_2\rangle = \int_{-\infty}^{+\infty} dx \langle\psi_1|x\rangle \langle x|X|\psi_2\rangle = \int_{-\infty}^{+\infty} dx \psi_1^*(x) x \psi_2(x). \quad (\text{A.185})$$

Similarly, we define the momentum operator  $P$  by means of its action on the  $\{|p\rangle\}$  basis:

$$\langle p|P|\psi\rangle = p\langle p|\psi\rangle = p\tilde{\psi}(p). \quad (\text{A.186})$$

To determine the action of the operator  $P$  in the position representation, we compute

$$\langle x|P|\psi\rangle = \int_{-\infty}^{+\infty} dp \langle x|p\rangle \langle p|P|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dp e^{ipx/\hbar} p \tilde{\psi}(p). \quad (\text{A.187})$$

We can recognize from the right-hand side of this equation the Fourier transform of  $p\tilde{\psi}(p)$ , which is equal to  $-i\hbar \frac{d}{dx} \psi(x)$ . Therefore

$$\langle x|P|\psi\rangle = -i\hbar \frac{d}{dx} \langle x|\psi\rangle = -i\hbar \frac{d}{dx} \psi(x). \quad (\text{A.188})$$

Quantities such as  $\langle\psi_1|P|\psi_2\rangle$  are therefore given, in the position representation, by

$$\langle\psi_1|P|\psi_2\rangle = \int_{-\infty}^{+\infty} dx \langle\psi_1|x\rangle \langle x|P|\psi_2\rangle = \int_{-\infty}^{+\infty} dx \psi_1^*(x) \left( -i\hbar \frac{d}{dx} \right) \psi_2(x). \quad (\text{A.189})$$

We can compute in the position representation the commutator  $[X, P]$ :

$$\langle x|[X, P]|\psi\rangle = \langle x|(XP - PX)|\psi\rangle \quad (\text{A.190})$$

$$= x\langle x|P|\psi\rangle + i\hbar \frac{d}{dx} \langle x|X|\psi\rangle \quad (\text{A.191})$$

$$= x \left( -i\hbar \frac{d}{dx} \right) \langle x|\psi\rangle + i\hbar \frac{d}{dx} (x\langle x|\psi\rangle) = i\hbar \langle x|\psi\rangle. \quad (\text{A.192})$$

---

<sup>1</sup>We remark that for infinite-dimensional Hilbert spaces a Hermitian operator  $A$  is not necessarily self-adjoint, due to issues, not discussed here, related to the domains of the operators  $A$  and  $A^\dagger$ .

Since the above calculation is valid for any  $|\psi\rangle$  and  $|x\rangle$ , we can conclude that

$$[X, P] = i\hbar, \quad (\text{A.193})$$

and therefore the position and momentum operators do not commute.

The operators  $X$  and  $P$  are self-adjoint. The operator  $X$  is Hermitian since for any ket  $|\psi_1\rangle$  and  $|\psi_2\rangle$  we have

$$\langle\psi_1|X|\psi_2\rangle = \int_{-\infty}^{+\infty} dx \psi_1^*(x) x \psi_2(x) = \left[ \int_{-\infty}^{+\infty} dx \psi_2^*(x) x \psi_1(x) \right]^* = \langle\psi_2|X|\psi_1\rangle^*. \quad (\text{A.194})$$

A similar proof that  $P$  is Hermitian, that is,  $\langle\psi_1|P|\psi_2\rangle = \langle\psi_2|P|\psi_1\rangle^*$ , can be performed by using the momentum representation. Finally, we show that the operators  $X$  and  $P$  are also self-adjoint since for both operators we can find a (continuous) basis of eigenvectors. We have

$$\langle x|X|x_0\rangle = x \langle x|x_0\rangle = x \delta(x - x_0) = x_0 \delta(x - x_0) = x_0 \langle x|x_0\rangle, \quad (\text{A.195})$$

and therefore  $|x_0\rangle$  is an eigenvector of  $X$ , the corresponding eigenvalue being  $x_0$ :

$$X|x_0\rangle = x_0|x_0\rangle. \quad (\text{A.196})$$

We can conclude that  $\{|x_0\rangle\}$  [ $x_0 \in (-\infty, +\infty)$ ] is a basis of eigenvectors of  $X$ . Similarly we can prove that

$$P|p_0\rangle = p_0|p_0\rangle, \quad (\text{A.197})$$

and therefore the  $\{|p_0\rangle\}$  basis [ $p_0 \in (-\infty, +\infty)$ ] is a basis of eigenvectors of  $P$ .

# Appendix B

## Solutions to the exercises

### B.1 Chapter 1

#### Exercise 1.1

$$\begin{aligned}
 (a \uparrow b) \uparrow (a \uparrow b) &= \overline{(a \uparrow b) \wedge (a \uparrow b)} = 1 - (a \uparrow b)^2 \\
 &= 1 - (a \uparrow b) = 1 - (1 - ab) = ab = a \wedge b.
 \end{aligned} \tag{B.1a}$$

$$\begin{aligned}
 (a \uparrow a) \uparrow (b \uparrow b) &= (1 - a \wedge a) \uparrow (1 - b \wedge b) = (1 - a^2) \uparrow (1 - b^2) \\
 &= (1 - a) \uparrow (1 - b) = \overline{(1 - a) \wedge (1 - b)} \\
 &= 1 - (1 - a)(1 - b) = a + b - ab = a \vee b.
 \end{aligned} \tag{B.1b}$$

**Exercise 1.2** We first note that both  $f_1$  and  $f_2$  are composed of  $M = 4$  clauses containing  $N = 3$  variables. To see whether there exists a configuration of variables  $\hat{\mathbf{a}}$  satisfying the two functions, it is sufficient to draw the corresponding truth table:

Table B.1

$a_1$	$a_2$	$a_3$	$f_1$	$f_2$
0	0	0	1	0
0	0	1	0	0
0	1	0	0	0
0	1	1	1	0
1	0	0	0	0
1	0	1	0	0
1	1	0	1	0
1	1	1	0	0

From this we observe that the instances  $\hat{\mathbf{a}} = (0, 0, 0)$  or  $(0, 1, 1)$  or  $(1, 1, 0)$  satisfy  $f_1(\hat{\mathbf{a}})$ . On the other hand, the function  $f_2$  is not satisfiable. The latter can also be seen noting that the second clause in  $f_2$  forces  $a_2$  to be 0, while the fourth clause requires  $a_1 = 1$ . However the two requirements are obviously incompatible with those of the first clause.

**Exercise 1.3** Let us assign the Boolean variables A, B and C to the three accused. We associate being guilty with output 1, and being innocent with output 0. The whole

interrogation can be formalized by the following function (in non-CNF form):

$$f(A, B, C) = [B \wedge \bar{C}] \wedge [\bar{A} \vee C] \wedge [\bar{C} \wedge (A \vee B)], \quad (\text{B.2})$$

the statements  $s_j$  of the three people ( $j = A, B, C$ ) being represented respectively by the expressions in the square brackets. Let us write the truth table of  $f$  and of  $s_A, s_B, s_C$ :

Table B.2

$A$	$B$	$C$	$s_A$	$s_B$	$s_C$	$f$
0	0	0	0	1	0	0
0	0	1	0	1	0	0
0	1	0	1	1	1	1
0	1	1	0	1	0	0
1	0	0	0	0	1	0
1	0	1	0	1	0	0
1	1	0	1	0	1	0
1	1	1	0	1	0	0

One realizes that the three claims are not contradictory, since  $f$  can be satisfied by  $A = 0, B = 1, C = 0$ . If we assume that all of them are guilty ( $A = B = C = 1$ ), we realize that  $s_A = s_C = 0, s_B = 1$ , therefore A and C lied (i.e., their claims are not satisfied). On the contrary, assuming that nobody lied ( $s_A = s_B = s_C = 1$ ), we see that  $A = C = 0, B = 1$ , therefore B is guilty.

**Exercise 1.4** If we set as input  $a = b = 1$ , then we obtain as output  $c' = \bar{c}$ . If we set  $c = 0$ , then  $c' = a \wedge b$ . We can then obtain the OR gate as follows. From the NOT gate we obtain  $\bar{a}$  and  $\bar{b}$ ; we then perform a Toffoli gate on inputs  $\bar{a}, \bar{b}$  and  $c = 1$ , giving  $c' = a \vee b$ .

**Exercise 1.5** It is sufficient to show that from the Fredkin gate we can construct the universal set of gates AND, OR, NOT and FANOUT. If we set  $b = 0$  and  $c = 1$  as input of the Fredkin gate, then  $b' = a$  and  $c' = \bar{a}$ . Therefore, we obtain simultaneously the FANOUT and NOT gates. Setting  $c = 0$ , we obtain  $c' = a \wedge b$ . The OR gate can be constructed from the AND and NOT gates by means of a De Morgan identity:  $a \vee b = \bar{\bar{a}} \wedge \bar{b}$ .

**Exercise 1.6** The voltage  $V_C$  across the capacitor, the voltage drop  $V_R$  on the resistor and the current  $I$  flowing in the  $RC$  circuit are related by the equations

$$I = C \frac{dV_C}{dt}, \quad V_R = RI. \quad (\text{B.3})$$

For a time-independent current and under the condition that the capacitor is initially uncharged,  $V_C(t=0) = 0$ , from Eq. (B.3) we obtain the waveform of the driving voltage:

$$V(t) = V_R + V_C(t) = RI + \frac{I}{C} t. \quad (\text{B.4})$$

## B.2 Chapter 2

**Exercise 2.1** Any unitary operator  $U$  is normal and therefore diagonalizable. Hence, we can write its spectral decomposition as

$$U = \sum_j \lambda_j |j\rangle\langle j|. \quad (\text{B.5})$$

Since unitary operators preserve the inner product between vectors we have

$$\langle j|U^\dagger U|j\rangle = |\lambda_j|^2 \langle j|j\rangle, \quad (\text{B.6})$$

which implies  $|\lambda_j| = 1$ . Thus, Eq. (B.5) can be rewritten as follows:

$$U = \sum_j e^{i\alpha_j} |j\rangle\langle j|, \quad (\text{B.7})$$

where  $\alpha_j$  is a real number. We may now define the operator

$$A = \sum_j \alpha_j |j\rangle\langle j|. \quad (\text{B.8})$$

Since  $A$  is already written in its diagonal basis, it is straightforward to see that  $e^{iA} = U$ . Let us eventually prove that  $A$  is Hermitian. We have

$$U = \sum_{n=0}^{\infty} \frac{(iA)^n}{n!}, \quad (\text{B.9})$$

and therefore

$$U^\dagger = \sum_{n=0}^{\infty} \left[ \frac{(iA)^n}{n!} \right]^\dagger = \sum_{n=0}^{\infty} \frac{(-iA^\dagger)^n}{n!} = e^{-iA^\dagger}. \quad (\text{B.10})$$

Since  $U^{-1} = e^{-iA}$ , the condition  $U^\dagger = U^{-1}$  is satisfied when  $A^\dagger = A$ , namely, when  $A$  is a Hermitian operator.

**Exercise 2.2** The Heisenberg uncertainty principle tells us that

$$\Delta\sigma_x \Delta\sigma_y \geq \frac{1}{2} |\langle 0|[\sigma_x, \sigma_y]|0\rangle|. \quad (\text{B.11})$$

After evaluation of the commutator  $[\sigma_x, \sigma_y]$ , we obtain

$$\Delta\sigma_x \Delta\sigma_y \geq \frac{1}{2} \left| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2i & 0 \\ 0 & -2i \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right| = 1. \quad (\text{B.12})$$

**Exercise 2.3** The first apparatus prepares the state  $|+\rangle_z = |0\rangle$ . The state  $|0\rangle$  then enters the second apparatus, which prepares the state  $|+\rangle_y = \frac{1}{\sqrt{2}}(-i|0\rangle + |1\rangle)$ . Finally, the third apparatus analyzes the state  $|+\rangle_y$ : it measures  $\sigma_z$  and obtains the two possible outcomes  $|+\rangle_z = |0\rangle$  or  $|-\rangle_z = |1\rangle$  with equal probabilities  $p_0$  and  $p_1$ , since

$$p_0 = |\langle 0|+\rangle_y|^2 = \frac{1}{2}, \quad p_1 = |\langle 1|+\rangle_y|^2 = \frac{1}{2}. \quad (\text{B.13})$$

**Exercise 2.4** The solution to the Schrödinger equation (2.53) is given by

$$|\psi(t)\rangle = \begin{bmatrix} a(t) \\ b(t) \end{bmatrix} = U(t) \begin{bmatrix} a(0) \\ b(0) \end{bmatrix}, \quad (\text{B.14})$$

where the unitary time-evolution operator is

$$U(t) = \exp \left[ -\frac{i}{\hbar} Ht \right], \quad (\text{B.15})$$

and

$$H = -\mu \mathbf{H} \cdot \boldsymbol{\sigma} \quad (\text{B.16})$$

is the Hamiltonian of the system. Let us compute explicitly  $U(t)$ . We have

$$-\frac{i}{\hbar} Ht = \frac{i\mu t}{\hbar} (\mathbf{H} \cdot \boldsymbol{\sigma}) = i\alpha (\mathbf{n} \cdot \boldsymbol{\sigma}), \quad (\text{B.17})$$

where we have defined

$$\alpha = \frac{\mu t}{\hbar} \sqrt{H_x^2 + H_y^2 + H_z^2}, \quad \mathbf{n} = \frac{1}{\sqrt{H_x^2 + H_y^2 + H_z^2}} (H_x, H_y, H_z). \quad (\text{B.18})$$

Performing a Taylor expansion of the operator  $U(t)$  we obtain

$$\begin{aligned} U(t) &= e^{i\alpha \mathbf{n} \cdot \boldsymbol{\sigma}} \\ &= [I - \frac{1}{2!} \alpha^2 (\mathbf{n} \cdot \boldsymbol{\sigma})^2 + \dots] + i [\alpha (\mathbf{n} \cdot \boldsymbol{\sigma}) - \frac{1}{3!} \alpha^3 (\mathbf{n} \cdot \boldsymbol{\sigma})^3 + \dots] \\ &= \cos \alpha I + i \sin \alpha (\mathbf{n} \cdot \boldsymbol{\sigma}), \end{aligned} \quad (\text{B.19})$$

where the last equality follows since  $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = I$ . Hence, we have

$$U(t) = \begin{bmatrix} \cos \alpha + i \sin \alpha n_z & \sin \alpha (n_y + i n_x) \\ \sin \alpha (-n_y + i n_x) & \cos \alpha - i \sin \alpha n_z \end{bmatrix}, \quad (\text{B.20})$$

where  $n_x$ ,  $n_y$ , and  $n_z$  are the Cartesian components of the unit vector  $\mathbf{n}$ . The average values of the Pauli operators are computed as follows:

$$\langle \sigma_i \rangle = \langle \psi(t) | \sigma_i | \psi(t) \rangle = \langle \psi(0) | U^\dagger \sigma_i U | \psi(0) \rangle. \quad (\text{B.21})$$

To obtain the state  $|1\rangle$  starting from the initial state  $|0\rangle$ , we can choose a magnetic field directed along the  $x$  axis. This means

$$\mathbf{H} = (H_x, 0, 0) \quad \text{and} \quad \mathbf{n} = (1, 0, 0). \quad (\text{B.22})$$

We require that at time  $\tilde{t}$  the wave vector  $|\psi(\tilde{t})\rangle = U(\tilde{t})|0\rangle$  coincide with  $|1\rangle$ ; that is,

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = U \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\alpha(\tilde{t})) & i \sin(\alpha(\tilde{t})) \\ i \sin(\alpha(\tilde{t})) & \cos(\alpha(\tilde{t})) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (\text{B.23})$$

This condition is fulfilled (up to an overall phase factor of no physical significance) when  $\cos(\alpha(\tilde{t})) = 0$ , which is first satisfied after a time  $\tilde{t}$  such that

$$\alpha(\tilde{t}) = \frac{\mu |H_x| \tilde{t}}{\hbar} = \frac{\pi}{2}. \quad (\text{B.24})$$

**Exercise 2.5** A state  $|\psi\rangle$  of two spin- $\frac{1}{2}$  particles is separable if and only if we can write it as follows:

$$|\psi\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle), \quad (\text{B.25})$$

with  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  complex coefficients satisfying the normalization conditions  $|\alpha|^2 + |\beta|^2 = 1$  and  $|\gamma|^2 + |\delta|^2 = 1$ . If the state  $|\psi\rangle$  is given by (2.60), then the separability

condition (B.25) implies  $\alpha\gamma = \frac{1}{\sqrt{2}}$ ,  $\beta\delta = \frac{1}{\sqrt{2}}$ ,  $\alpha\delta = 0$  and  $\beta\gamma = 0$ . As these four relations cannot be satisfied simultaneously, the state must be entangled.

**Exercise 2.6** After insertion of the explicit expressions for  $|+\rangle_u$  and  $|-\rangle_u$ , given by Eq. (2.49), we obtain:

$$\begin{aligned} |+\rangle_u|-\rangle_u - |-\rangle_u|+\rangle_u &= \frac{1}{\sqrt{2}} \begin{bmatrix} \cos \frac{\theta}{2} e^{-i\phi/2} \\ \sin \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} -\sin \frac{\theta}{2} e^{-i\phi/2} \\ \cos \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} \\ &\quad - \frac{1}{\sqrt{2}} \begin{bmatrix} -\sin \frac{\theta}{2} e^{-i\phi/2} \\ \cos \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} \cos \frac{\theta}{2} e^{-i\phi/2} \\ \sin \frac{\theta}{2} e^{i\phi/2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}. \end{aligned} \quad (\text{B.26})$$

The final state is indeed equal to the singlet state  $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ , independently of the direction of  $\mathbf{u}$ .

**Exercise 2.7** (a) Measurement of  $\sigma_z$  for the first particle. It is convenient to rewrite the state (2.66) as

$$|\psi\rangle = \sqrt{|\alpha|^2 + |\beta|^2} |0\rangle \otimes \frac{\alpha|0\rangle + \beta|1\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}} + \sqrt{|\gamma|^2 + |\delta|^2} |1\rangle \otimes \frac{\gamma|0\rangle + \delta|1\rangle}{\sqrt{|\gamma|^2 + |\delta|^2}}. \quad (\text{B.27})$$

Therefore, a measurement of  $\sigma_z$  for the first particle results in outcome  $+1$  with probability ( $|\alpha|^2 + |\beta|^2$ ) and outcome  $-1$  with probability ( $|\gamma|^2 + |\delta|^2$ ). Furthermore, after a measurement with result  $\sigma_z = +1$ , the state of the second particle collapses onto the state

$$|\phi^{(0)}\rangle_2 = \frac{\alpha|0\rangle + \beta|1\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}}; \quad (\text{B.28a})$$

if instead the result is  $\sigma_z = -1$ , the state of the second particle collapses onto

$$|\phi^{(1)}\rangle_2 = \frac{\gamma|0\rangle + \delta|1\rangle}{\sqrt{|\gamma|^2 + |\delta|^2}}. \quad (\text{B.28b})$$

(b) Measurement of  $\sigma_x$  for the first particle. Taking into account that

$$|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle), \quad |1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle), \quad (\text{B.29})$$

we can rewrite the state (2.66) as

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)(\gamma|0\rangle + \delta|1\rangle) \\ &= \sqrt{\frac{|\alpha + \gamma|^2 + |\beta + \delta|^2}{2}} |+\rangle \otimes \frac{(\alpha + \gamma)|0\rangle + (\beta + \delta)|1\rangle}{\sqrt{|\alpha + \gamma|^2 + |\beta + \delta|^2}} \\ &\quad + \sqrt{\frac{|\alpha - \gamma|^2 + |\beta - \delta|^2}{2}} |-\rangle \otimes \frac{(\alpha - \gamma)|0\rangle + (\beta - \delta)|1\rangle}{\sqrt{|\alpha - \gamma|^2 + |\beta - \delta|^2}}. \end{aligned} \quad (\text{B.30})$$

Then, analogously to the previous case, we can compute the probability of obtaining  $\sigma_x = +1$  or  $\sigma_x = -1$  and the corresponding wave vectors onto which the state of the second particle collapses after the measurement.

**Exercise 2.8** It is easy to check that

$$\langle 0 | \boldsymbol{\sigma} \cdot \mathbf{r} | 0 \rangle = z, \quad \langle 1 | \boldsymbol{\sigma} \cdot \mathbf{r} | 1 \rangle = -z, \quad \langle 0 | \boldsymbol{\sigma} \cdot \mathbf{r} | 1 \rangle = x - iy, \quad \langle 1 | \boldsymbol{\sigma} \cdot \mathbf{r} | 0 \rangle = x + iy, \quad (\text{B.31})$$

where  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  and  $\mathbf{r} = (x, y, z)$ . Using equalities (B.31), we obtain

$$\begin{aligned} \langle \psi | (\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}) \otimes (\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}) | \psi \rangle &= \frac{1}{2} (\langle 01 | - \langle 10 |) (\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}) \otimes (\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}) (\langle 01 | - \langle 10 |) \\ &= \frac{1}{2} \langle 0 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 0 \rangle \langle 1 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 1 \rangle - \frac{1}{2} \langle 0 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 1 \rangle \langle 1 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 0 \rangle \\ &\quad - \frac{1}{2} \langle 1 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 0 \rangle \langle 0 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 1 \rangle + \frac{1}{2} \langle 1 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 1 \rangle \langle 0 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 0 \rangle \\ &= -\mathbf{a} \cdot \mathbf{b}, \end{aligned} \quad (\text{B.32})$$

where  $|\psi\rangle$  is the singlet state (2.63).

**Exercise 2.9** We can always choose the arbitrary phases multiplying the basis states for the two spin- $\frac{1}{2}$  particles so that  $\alpha$  and  $\beta$  are real and positive. Let us compute the correlator

$$\begin{aligned} C(\mathbf{a}, \mathbf{b}) &= \langle \psi | (\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}) (\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}) | \psi \rangle \\ &= (\alpha \langle 00 | + \beta \langle 11 |) (\boldsymbol{\sigma}^{(A)} \cdot \mathbf{a}) (\boldsymbol{\sigma}^{(B)} \cdot \mathbf{b}) (\alpha | 00 \rangle + \beta | 11 \rangle) \\ &= \alpha^2 \langle 0 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 0 \rangle \langle 0 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 0 \rangle + \beta^2 \langle 1 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 1 \rangle \langle 1 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 1 \rangle \\ &\quad + 2\alpha\beta \operatorname{Re} \left( \langle 0 | \boldsymbol{\sigma}^{(A)} \cdot \mathbf{a} | 1 \rangle \langle 0 | \boldsymbol{\sigma}^{(B)} \cdot \mathbf{b} | 1 \rangle \right) \\ &= z_a z_b + 2\alpha\beta(x_a x_b - y_a y_b), \end{aligned} \quad (\text{B.33})$$

where  $\mathbf{a} = (x_a, y_a, z_a)$  and  $\mathbf{b} = (x_b, y_b, z_b)$ . If we consider the set of directions  $\mathbf{a} = (1, 0, 0)$ ,  $\mathbf{a}' = (0, 0, 1)$ ,  $\mathbf{b} = (x_b, 0, z_b)$  and  $\mathbf{b}' = (-x_b, 0, z_b)$ , we obtain

$$|C(\mathbf{a}, \mathbf{b}) - C(\mathbf{a}, \mathbf{b}')| + |C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}', \mathbf{b}')| = 2(z_b + 2\alpha\beta x_b). \quad (\text{B.34})$$

Therefore, the CHSH inequality is violated if  $z_b + 2\alpha\beta x_b > 1$ . If  $\theta_b$  denotes the angle between the unit vector  $\mathbf{b}$  and the  $z$  axis, we obtain

$$z_b + 2\alpha\beta x_b = \cos \theta_b + 2\alpha\beta \sin \theta_b = 1 + 2\alpha\beta \theta_b + O(\theta_b^2), \quad (\text{B.35})$$

which, insofar as  $\alpha$  and  $\beta$  are both different from zero, is larger than 1, provided that  $\theta_b$  is positive and small enough. Therefore, the violation of Bell's inequalities is a generic feature of entangled states.

**Exercise 2.10** The reduced density matrix  $\rho_1$ , defined by Eq. (2.120), is Hermitian since

$$(\rho_1)_{ji}^* = \sum_{\alpha} \rho_{j\alpha; i\alpha}^* = \sum_{\alpha} \rho_{i\alpha; j\alpha} = (\rho_1)_{ij}. \quad (\text{B.36})$$

To show that  $\rho_1$  is non-negative, it is convenient to decompose the total density matrix  $\rho$  in the basis of its eigenvectors  $\{|u_k\rangle_1|v_\gamma\rangle_2\}$ :

$$\rho = \sum_{k,\gamma} \rho_{k\gamma;k\gamma} |u_k\rangle_1|v_\gamma\rangle_2 \langle u_k|_2 \langle v_\gamma|, \quad (\text{B.37})$$

with eigenvalues  $\rho_{k\gamma;k\gamma} \geq 0$ , as the density matrix  $\rho$  is non-negative. Then, for any  $|\phi\rangle_1 \in \mathcal{H}_1$ ,

$${}_1\langle\phi|\rho_1|\phi\rangle_1 = \sum_\gamma \sum_k \rho_{k\gamma;k\gamma} |{}_1\langle\phi|k\rangle_1|^2 \geq 0. \quad (\text{B.38})$$

Finally,  $\rho_1$  has unit trace since

$$\sum_i (\rho_1)_{ii} = \sum_{i,\alpha} \rho_{i\alpha;i\alpha} = 1. \quad (\text{B.39})$$

**Exercise 2.11** For a pure bipartite separable state, we can write

$$|\psi\rangle = |\alpha\rangle_1 \otimes |\beta\rangle_2. \quad (\text{B.40})$$

Hence, the corresponding density matrix is given by

$$\rho = |\psi\rangle\langle\psi| = |\alpha\rangle_1|\beta\rangle_2 {}_{1\langle}\alpha|_2\langle\beta| = |\alpha\rangle_1 {}_{1\langle}\alpha| \otimes |\beta\rangle_2 {}_{2\langle}\beta|. \quad (\text{B.41})$$

After partial tracing, we obtain

$$\begin{aligned} \rho_1 &= \text{Tr}_2(|\alpha\rangle_1 {}_{1\langle}\alpha| \otimes |\beta\rangle_2 {}_{2\langle}\beta|) = |\alpha\rangle_1 {}_{1\langle}\alpha|, \\ \rho_2 &= \text{Tr}_1(|\alpha\rangle_1 {}_{1\langle}\alpha| \otimes |\beta\rangle_2 {}_{2\langle}\beta|) = |\beta\rangle_2 {}_{2\langle}\beta|, \end{aligned} \quad (\text{B.42})$$

and therefore

$$\rho = \rho_1 \otimes \rho_2. \quad (\text{B.43})$$

**Exercise 2.12** If the state  $|\psi\rangle$  has Schmidt decomposition (2.141), then

$$\rho_1 = \sum_i p_i |i\rangle_1 {}_{1\langle} i|, \quad \rho_2 = \sum_i p_i |i'\rangle_2 {}_{2\langle} i'|, \quad \rho_3 = \sum_i p_i |i''\rangle_3 {}_{3\langle} i''. \quad (\text{B.44})$$

This means that the reduced density matrices  $\rho_1$ ,  $\rho_2$  and  $\rho_3$  should have the same spectrum. Therefore, to solve this exercise, it will be sufficient to provide a counterexample. For instance, the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|01\rangle_{12} + |10\rangle_{12}) \otimes |0\rangle_3 \quad (\text{B.45})$$

does not satisfy the requirement of equal spectra for  $\rho_1$ ,  $\rho_2$  and  $\rho_3$ . Indeed, this state is the tensor product of a Bell state for the first two spin- $\frac{1}{2}$  particles and a separable state for the third spin- $\frac{1}{2}$  particle, and therefore  $\rho_1 = \frac{1}{2} I_1$ ,  $\rho_2 = \frac{1}{2} I_2$ ,  $\rho_3 = |0\rangle_3 \langle 0|$ . This means that  $\rho_1$  and  $\rho_2$  have eigenvalues  $p_1 = p_2 = \frac{1}{2}$ , while  $\rho_3$  has a single eigenvalue equal to 1.

**Exercise 2.13** Let us consider the system of equations (2.146), for the case in which there is a system of two spin- $\frac{1}{2}$  particles and we add two ancillary spin- $\frac{1}{2}$  particles to purify it. The reduced density matrix  $\rho_1$  is a  $4 \times 4$  matrix. Taking into account that  $(\rho_1)_{ji} = (\rho_1)_{ij}^*$  ( $i, j = 1, \dots, 4$ ) and that  $\text{Tr } \rho_1 = 1$ ,  $\rho_1$  is determined by 15 independent

real parameters. The pure state of the extended 4-particle system depends on 16 complex coefficients  $c_{i\alpha}$  ( $i, \alpha = 1, \dots, 4$ ), namely on 32 real parameters, 30 of which are independent, taking into account the normalization condition and the existence of an arbitrary global phase. Therefore, we have the freedom to set  $30 - 15 = 15$  real parameters. We take

$$c_{12} = c_{13} = c_{14} = c_{23} = c_{24} = c_{34} = 0 \quad (\text{B.46})$$

and the coefficients  $c_{22}$ ,  $c_{33}$  and  $c_{44}$  real and positive. We now solve Eq. (2.146); that is, we determine the coefficients  $c_{i\alpha}$  as a function of the matrix elements  $(\rho_1)_{ij}$ . We obtain

$$(\rho_1)_{11} = \sum_{\alpha} c_{1\alpha} c_{1\alpha}^* = c_{11} c_{11}^*. \quad (\text{B.47})$$

We can exploit the existence of an arbitrary global phase to choose also  $c_{11}$  real and positive, so that

$$c_{11} = \sqrt{(\rho_1)_{11}}. \quad (\text{B.48})$$

Then we obtain

$$(\rho_1)_{12} = \sum_{\alpha} c_{1\alpha} c_{2\alpha}^* = c_{11} c_{21}^*, \quad (\text{B.49})$$

and therefore

$$c_{21} = \frac{(\rho_1)_{12}^*}{c_{11}} = \frac{(\rho_1)_{12}^*}{\sqrt{(\rho_1)_{11}}}. \quad (\text{B.50})$$

Similarly, we obtain

$$c_{31} = \frac{(\rho_1)_{13}^*}{\sqrt{(\rho_1)_{11}}}, \quad c_{41} = \frac{(\rho_1)_{14}^*}{\sqrt{(\rho_1)_{11}}}. \quad (\text{B.51})$$

Then we use

$$(\rho_1)_{22} = \sum_{\alpha} c_{2\alpha} c_{2\alpha}^* = \frac{|(\rho_1)_{12}|^2}{(\rho_1)_{11}} + |c_{22}|^2 \quad (\text{B.52})$$

to extract

$$c_{22} = \sqrt{(\rho_1)_{22} - \frac{|(\rho_1)_{12}|^2}{(\rho_1)_{11}}}. \quad (\text{B.53})$$

We can now derive  $c_{32}$  and  $c_{42}$  from the equations

$$\begin{aligned} (\rho_1)_{23} &= \frac{(\rho_1)_{12}^*(\rho_1)_{13}}{(\rho_1)_{11}} + \sqrt{(\rho_1)_{22} - \frac{|(\rho_1)_{12}|^2}{(\rho_1)_{11}}} c_{32}^* \\ (\rho_1)_{24} &= \frac{(\rho_1)_{12}^*(\rho_1)_{14}}{(\rho_1)_{11}} + \sqrt{(\rho_1)_{22} - \frac{|(\rho_1)_{12}|^2}{(\rho_1)_{11}}} c_{42}^*. \end{aligned} \quad (\text{B.54})$$

Finally, we obtain  $c_{33}$ ,  $c_{43}$  and  $c_{44}$  from the equations

$$\begin{aligned} (\rho_1)_{33} &= |c_{31}|^2 + |c_{32}|^2 + |c_{33}|^2, \quad (\rho_1)_{34} = c_{31} c_{41}^* + c_{32} c_{42}^* + c_{33} c_{43}^*, \\ (\rho_1)_{44} &= \sum_{\alpha=1}^4 |c_{4\alpha}|^2. \end{aligned} \quad (\text{B.55})$$

### B.3 Chapter 3

**Exercise 3.1** The most general  $2 \times 2$  Hermitian matrix  $A$  can be written as

$$A = \begin{bmatrix} a & b + ic \\ b - ic & d \end{bmatrix}, \quad (\text{B.56})$$

where  $a, b, c$  and  $d$  are real parameters. We have  $A = \alpha I + \beta \sigma_x + \gamma \sigma_y + \delta \sigma_z$ , provided  $\alpha = (a+d)/2$ ,  $\delta = (a-d)/2$ ,  $\beta = b$  and  $\gamma = -c$ . Therefore,  $\alpha, \beta, \gamma$  and  $\delta$  are all real.

**Exercise 3.2** We have  $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$  and  $\sigma_x \sigma_y = i \sigma_z$ ,  $\sigma_y \sigma_z = i \sigma_x$  and  $\sigma_z \sigma_x = i \sigma_y$ . A compact method to express these relations is

$$\sigma_j \sigma_k = \delta_{jk} I + i \sum_{l=1}^3 \epsilon_{jkl} \sigma_l, \quad (\text{B.57})$$

where  $\sigma_1 \equiv \sigma_x$ ,  $\sigma_2 \equiv \sigma_y$ ,  $\sigma_3 \equiv \sigma_z$  and  $\epsilon_{jkl}$  is the Ricci antisymmetric tensor, with  $\epsilon_{jkl} = 0$  if the three indices are not all different,  $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$  and  $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$ . As  $\text{Tr } I = 2$  and  $\text{Tr } \sigma_l = 0$ , we obtain from Eq. (B.57)  $\text{Tr}(\sigma_j \sigma_k) = 2\delta_{jk}$ .

**Exercise 3.3** Let  $\mathbf{r}_1 = (r_{11}, r_{12}, r_{13}) = (x_1, y_1, z_1)$  and  $\mathbf{r}_2 = (r_{21}, r_{22}, r_{23}) = (x_2, y_2, z_2)$  denote the Bloch vectors associated with the density matrices  $\rho_1$  and  $\rho_2$ , respectively. We obtain by direct computation

$$\begin{aligned} [\rho_1, \rho_2] &= \frac{1}{2} (I + \mathbf{r}_1 \cdot \boldsymbol{\sigma}) \frac{1}{2} (I + \mathbf{r}_2 \cdot \boldsymbol{\sigma}) - \frac{1}{2} (I + \mathbf{r}_2 \cdot \boldsymbol{\sigma}) \frac{1}{2} (I + \mathbf{r}_1 \cdot \boldsymbol{\sigma}) \\ &= \frac{1}{4} [(\mathbf{r}_1 \cdot \boldsymbol{\sigma})(\mathbf{r}_2 \cdot \boldsymbol{\sigma}) - (\mathbf{r}_2 \cdot \boldsymbol{\sigma})(\mathbf{r}_1 \cdot \boldsymbol{\sigma})] = \frac{1}{4} \sum_{jk} r_{1j} r_{2k} (\sigma_j \sigma_k - \sigma_k \sigma_j) \\ &= \frac{1}{4} \sum_{jk} r_{1j} r_{2k} [\sigma_j, \sigma_k] = \frac{1}{4} \sum_{jk} 2i \epsilon_{jkl} r_{1j} r_{2k} \sigma_l \\ &= \frac{1}{2} (x_1 y_2 - y_1 x_2) \sigma_z + \frac{1}{2} (y_1 z_2 - z_1 y_2) \sigma_x + \frac{1}{2} (z_1 x_2 - x_1 z_2) \sigma_y. \end{aligned} \quad (\text{B.58})$$

Thus,  $[\rho_1, \rho_2] = 0$  when the conditions  $x_1 y_2 - y_1 x_2 = 0$ ,  $y_1 z_2 - z_1 y_2 = 0$  and  $z_1 x_2 - x_1 z_2 = 0$  are simultaneously satisfied; that is, when  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are parallel.

**Exercise 3.4** It is convenient to perform a rigid rotation of the two states,  $|\psi_1\rangle \rightarrow |\psi'_1\rangle = R|\psi_1\rangle$  and  $|\psi_2\rangle \rightarrow |\psi'_2\rangle = R|\psi_2\rangle$ , with  $R$  a rotation matrix, in such a way that one of the rotated states, for example  $|\psi'_1\rangle$ , coincides with the north pole of the Bloch sphere. Then we have  $|\psi'_1\rangle = |0\rangle$  and  $|\psi'_2\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle$ , where the angle  $\theta$  and  $\phi$  parametrize the position of the state vector  $|\psi'_2\rangle$  on the Bloch sphere (see Eq. (3.4)). We note that  $\theta$  is also the angle between the Bloch vectors  $|\psi_1\rangle$  and  $|\psi_2\rangle$ . We have

$$f = |\langle \psi_1 | \psi_2 \rangle|^2 = |\langle \psi'_1 | \psi'_2 \rangle|^2 = \left| \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{bmatrix} \right|^2 = \cos^2 \frac{\theta}{2}. \quad (\text{B.59})$$

**Exercise 3.5** We start by noting that

$$H R_z(\alpha) H = e^{i\frac{\alpha}{2}} \begin{bmatrix} \cos \frac{\alpha}{2} & -i \sin \frac{\alpha}{2} \\ -i \sin \frac{\alpha}{2} & \cos \frac{\alpha}{2} \end{bmatrix} = e^{i\frac{\alpha}{2}} [\cos \frac{\alpha}{2} I - i \sin \frac{\alpha}{2} \sigma_x] \quad (\text{B.60})$$

performs a rotation through an angle  $-\alpha$  about the  $x$ -axis (the overall phase factor  $e^{i\frac{\alpha}{2}}$  is inessential). The chain of rotations written in Eq. (3.49) act as follows:

- (i)  $R_z(-\frac{\pi}{2} - \phi_1)$  is a rotation of the Bloch sphere through an angle  $-\frac{\pi}{2} - \phi_1$  about the  $z$  axis. Thus, the state parametrized on the Bloch sphere by  $(\theta_1, \phi_1)$  is moved into the state  $(\theta_1, -\frac{\pi}{2})$ . This state belongs to the plane  $(-y, z)$ .
- (ii) As shown in Eq. (B.60),  $H R_z(\theta_2 - \theta_1) H$  rotates the Bloch sphere through an angle  $\theta_1 - \theta_2$  about the  $x$  axis. Thus, the state vector becomes  $(\theta_2, -\frac{\pi}{2})$ .
- (iii)  $R_z(\frac{\pi}{2} + \phi_2)$  is a rotation through an angle  $\frac{\pi}{2} + \phi_2$  about the  $z$  axis and therefore leads to the final  $(\theta_2, \phi_2)$ .

**Exercise 3.6** After computation of  $|\psi'\rangle = R_x(\delta)|\psi\rangle$ , we obtain

$$\begin{cases} x' = x, \\ y' = y \cos \delta - z \sin \delta, \\ z' = y \sin \delta + z \cos \delta, \end{cases} \quad (\text{B.61})$$

where  $(x, y, z)$  and  $(x', y', z')$  denote the Cartesian coordinates of the vectors  $|\psi\rangle$  and  $|\psi'\rangle$ . It is evident that (B.61) represents a counterclockwise rotation through an angle  $\delta$  about the  $x$  axis. Likewise we can compute  $|\psi'\rangle = R_y(\delta)|\psi\rangle$ , obtaining

$$\begin{cases} x' = x \cos \delta + z \sin \delta, \\ y' = y, \\ z' = -x \sin \delta + z \cos \delta, \end{cases} \quad (\text{B.62})$$

which represent a counterclockwise rotation through an angle  $\delta$  about the  $y$  axis.

**Exercise 3.7** A simple comparison between the matrix  $U_1$ , Eq. (3.25), and the rotation matrix  $R_y(\delta)$ , Eq. (3.55), shows that  $U_1 = R_y(-\frac{\pi}{2})$ . Similarly, a comparison of  $U_2$ , Eq. (3.27), and  $R_x(\delta)$ , Eq. (3.54), shows that  $U_2 = R_x(-\frac{\pi}{2})$ .

**Exercise 3.8** A generic  $2 \times 2$  unitary matrix  $U$  can be seen as a rotation through an angle  $\delta$  about some axis of the Bloch sphere (this axis is directed along the unit vector  $\mathbf{n} = (n_x, n_y, n_z)$ ). Therefore, we can express  $U$  according to Eq. (3.60). It follows that

$$\sqrt{U} = \cos\left(\frac{\delta}{4}\right) I - i \sin\left(\frac{\delta}{4}\right) (\mathbf{n} \cdot \boldsymbol{\sigma}). \quad (\text{B.63})$$

**Exercise 3.9** We get by direct computation

$$\begin{aligned} (\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) &= (a_x \sigma_x + a_y \sigma_y + a_z \sigma_z)(b_x \sigma_x + b_y \sigma_y + b_z \sigma_z) \\ &= a_x b_x \sigma_x^2 + a_y b_y \sigma_y^2 + a_z b_z \sigma_z^2 + (a_x b_y - a_y b_x) \sigma_x \sigma_y \\ &\quad + (a_y b_z - a_z b_y) \sigma_y \sigma_z + (a_z b_x - a_x b_z) \sigma_z \sigma_x \\ &= a_x b_x I + a_y b_y I + a_z b_z I + (a_x b_y - a_y b_x) i \sigma_z \\ &\quad + (a_y b_z - a_z b_y) i \sigma_x + (a_z b_x - a_x b_z) i \sigma_y \\ &= (\mathbf{a} \cdot \mathbf{b}) I + i \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \end{aligned} \quad (\text{B.64})$$

Note that we have taken advantage of the following properties of the Pauli matrices: (i)  $\sigma_y \sigma_x = -\sigma_y \sigma_x$ ,  $\sigma_z \sigma_y = -\sigma_y \sigma_z$ , and  $\sigma_x \sigma_z = -\sigma_z \sigma_x$  (that is, the Pauli matrices anti-commute); (ii)  $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$ ; and (iii)  $\sigma_x \sigma_y = i \sigma_z$ ,  $\sigma_y \sigma_z = i \sigma_x$ , and  $\sigma_z \sigma_x = i \sigma_y$ .

**Exercise 3.10** The state vector (3.67) can be equivalently rewritten as

$$|\psi\rangle = a \left\{ |00\rangle + b_0 e^{i\phi_0} |01\rangle + b_1 e^{i\phi_1} |10\rangle + b_1 b_0 e^{i(\phi_1+\phi_0)} |11\rangle \right\}. \quad (\text{B.65})$$

The application of the CNOT gate to this state leads to

$$\text{CNOT } |\psi\rangle = a \left\{ |00\rangle + b_0 e^{i\phi_0} |01\rangle + b_1 e^{i\phi_1} |11\rangle + b_1 b_0 e^{i(\phi_1+\phi_0)} |10\rangle \right\}, \quad (\text{B.66})$$

which is separable if and only if  $b_0 e^{i\phi_0} = 1$ . Thus, CNOT generates an entangled state if and only if at least one of the following two conditions is fulfilled:

$$b_0 \neq 1, \quad \phi_0 \neq 0, \pi. \quad (\text{B.67})$$

**Exercise 3.11** The implementation of the generalized CNOT gate  $B$  (defined by Eq. (3.68)) by means of the standard CNOT and single-qubit (NOT) gates is shown in Fig. B.1. The first NOT gate flips the state of the control qubit and thus makes the standard CNOT gate act non-trivially only if the state of the control qubit was  $|0\rangle$  at the beginning (before the NOT gate). The second NOT gate restores the original state of the target qubit. With the same procedure we obtain the generalized CNOT gate  $D$  from  $C$ . Finally, we note that the NOT gates are implemented by the Pauli matrix  $\sigma_x$ .

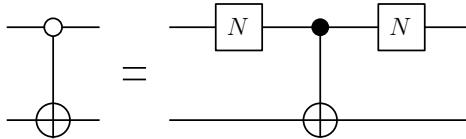


Fig. B.1 Decomposition of the generalized CNOT gate  $B$  into a standard CNOT gate and two NOT gates, denoted by  $N$ .

In order to prove the equality between the two circuits shown in Fig. 3.4, we must verify that

$$C = H^{\otimes 2} A H^{\otimes 2}, \quad (\text{B.68})$$

where  $A$  is the standard CNOT gate,  $C$  a generalized CNOT and  $H^{\otimes 2} \equiv H \otimes H$ . We have

$$H^{\otimes 2} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \quad (\text{B.69})$$

Then we can explicitly perform the matrix products in Eq. (B.68) and verify that

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = H^{\otimes 2} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} H^{\otimes 2}. \quad (\text{B.70})$$

**Exercise 3.12** It is sufficient to compute matrix products. For instance, we can check that the SWAP gate is implemented by the circuit of Fig. 3.5. Indeed, we have

$$A C A = \text{SWAP}, \quad (\text{B.71})$$

since

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (\text{B.72})$$

Equivalently, we can apply the sequence of generalized CNOT gates ( $ACA$ ) to the states of the computational basis and check that

$$\begin{aligned} ACA|00\rangle &= |00\rangle = \text{SWAP}|00\rangle, & ACA|01\rangle &= |10\rangle = \text{SWAP}|01\rangle, \\ ACA|10\rangle &= |01\rangle = \text{SWAP}|10\rangle, & ACA|11\rangle &= |11\rangle = \text{SWAP}|11\rangle. \end{aligned} \quad (\text{B.73})$$

Therefore,  $ACA = \text{SWAP}$  since both operators are linear and have the same action on a set of basis vectors.

The 24 possible permutation matrices are given by

$$\begin{aligned} P_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_2 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & P_3 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_4 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\ P_5 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & P_6 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_7 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_8 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ P_9 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_{10} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_{11} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_{12} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ P_{13} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_{14} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_{15} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, & P_{16} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \\ P_{17} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_{18} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_{19} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & P_{20} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\ P_{21} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & P_{22} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & P_{23} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & P_{24} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (\text{B.74})$$

Note that  $P_3 = \text{SWAP}$ . These permutation matrices are implemented by means of generalized CNOT gates, as shown in Fig. B.2.

**Exercise 3.13** It is easy to check by direct matrix multiplication that

$$\text{CMINUS} = (I \otimes H) \text{CNOT}(I \otimes H). \quad (\text{B.75})$$

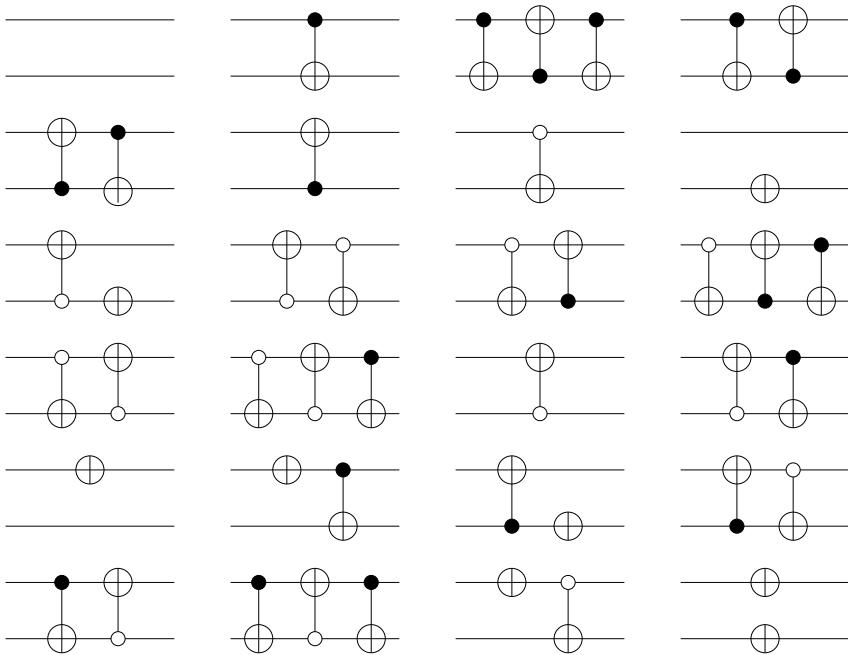


Fig. B.2 Quantum circuits implementing the permutation matrices (B.74). The  $\oplus$  symbol, without a control qubit, denotes a simple NOT gate (see, e.g., permutation  $P_8$ ). Permutations are ordered from top to bottom and from left to right: the top line gives quantum circuits from  $P_1$  (left) to  $P_4$  (right), the second line (from the top) from  $P_5$  (left) to  $P_8$  (right) and so on.

Then the relation

$$\text{CNOT} = (I \otimes H) \text{ CMINUS } (I \otimes H) \quad (\text{B.76})$$

follows immediately if we multiply both sides of Eq. (B.75) by  $(I \otimes H)$ , since  $(I \otimes H)^2 = I \otimes I$ .

**Exercise 3.14** Let us first consider a phase error acting on the target qubit. It transforms the state (3.73) into

$$(\alpha |0\rangle + \beta |1\rangle) \otimes \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}} (\alpha |00\rangle - \alpha |01\rangle + \beta |10\rangle - \beta |11\rangle). \quad (\text{B.77})$$

After application of the CNOT gate, this state becomes

$$\frac{1}{\sqrt{2}} (\alpha |00\rangle - \alpha |01\rangle + \beta |11\rangle - \beta |10\rangle) = \frac{1}{\sqrt{2}} (\alpha |0\rangle - \beta |1\rangle) \otimes (|0\rangle - |1\rangle). \quad (\text{B.78})$$

Therefore, this kind of error is particularly dangerous, since, even though initially it only affects the target qubit, after application of the CNOT gate it is also transferred to the control qubit.

Note that this is not the case for the other possible phase or amplitude errors. For instance, a phase error acting on the control qubit transforms the state (3.73) into

$$(\alpha |0\rangle - \beta |1\rangle) \otimes \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle),$$

and this state is unchanged by application of the CNOT gate. Therefore, the target qubit is not affected by this phase error. An amplitude error acting on the control qubit transforms (3.73) into

$$(\beta|0\rangle + \alpha|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$

which is not modified by the CNOT gate. Again, the target qubit is safe. Finally, an amplitude error acting on the target qubit does not affect the computation, that is, the state (3.73) stays the same. This is because the target qubit is symmetric under the amplitude error:  $|0\rangle + |1\rangle$  is not modified when  $|0\rangle \leftrightarrow |1\rangle$ .

**Exercise 3.15** By direct computation of the tensor products we obtain:

$$\begin{aligned} \sigma_1 \otimes \sigma_1 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_1 \otimes \sigma_2 = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 -i & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_1 \otimes \sigma_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 -1 & 0 & 0 & 0 \end{bmatrix}, \\ \sigma_2 \otimes \sigma_1 &= \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_2 \otimes \sigma_2 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_2 \otimes \sigma_3 = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 -i & 0 & 0 & 0 \end{bmatrix}, \\ \sigma_3 \otimes \sigma_1 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \sigma_3 \otimes \sigma_2 = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{bmatrix}, \quad \sigma_3 \otimes \sigma_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ \sigma_1 \otimes I &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \sigma_2 \otimes I = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix}, \quad \sigma_3 \otimes I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \\ I \otimes \sigma_1 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad I \otimes \sigma_2 = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad I \otimes \sigma_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \end{aligned} \tag{B.79}$$

and also  $I \otimes I = I$ .

**Exercise 3.16**

$$\langle \psi | \sigma_1 \otimes \sigma_1 | \psi \rangle = [c \ \alpha^* \ \beta^* \ \gamma^*] \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c \\ \alpha \\ \beta \\ \gamma \end{bmatrix} = c\gamma + c\gamma^* + \alpha\beta^* + \alpha^*\beta, \tag{B.80}$$

And similarly we have:

$$\begin{aligned}
 \langle \psi | \sigma_1 \otimes \sigma_2 | \psi \rangle &= i(-c\gamma + c\gamma^* - \alpha\beta^* + \alpha^*\beta), & \langle \psi | \sigma_1 \otimes \sigma_3 | \psi \rangle &= c\beta + c\beta^* - \alpha\gamma^* - \alpha^*\gamma, \\
 \langle \psi | \sigma_2 \otimes \sigma_1 | \psi \rangle &= i(-c\gamma + c\gamma^* + \alpha\beta^* - \alpha^*\beta), & \langle \psi | \sigma_2 \otimes \sigma_2 | \psi \rangle &= -c\gamma - c\gamma^* + \alpha\beta^* + \alpha^*\beta, \\
 \langle \psi | \sigma_2 \otimes \sigma_3 | \psi \rangle &= i(-c\beta + c\beta^* - \alpha\gamma^* + \alpha^*\gamma), & \langle \psi | \sigma_3 \otimes \sigma_1 | \psi \rangle &= c\alpha + c\alpha^* - \beta\gamma^* - \beta^*\gamma, \\
 \langle \psi | \sigma_3 \otimes \sigma_2 | \psi \rangle &= i(-c\alpha + c\alpha^* - \beta\gamma^* + \beta^*\gamma), & \langle \psi | \sigma_3 \otimes \sigma_3 | \psi \rangle &= c^2 - |\alpha|^2 - |\beta|^2 + |\gamma|^2, \\
 \langle \psi | \sigma_1 \otimes I | \psi \rangle &= c\beta + \alpha^*\gamma + \beta^*c + \gamma^*\alpha, & \langle \psi | \sigma_2 \otimes I | \psi \rangle &= i(-c\beta - \alpha^*\gamma + \beta^*c + \gamma^*\alpha), \\
 \langle \psi | \sigma_3 \otimes I | \psi \rangle &= c^2 + |\alpha|^2 - |\beta|^2 - |\gamma|^2, & \langle \psi | I \otimes \sigma_1 | \psi \rangle &= c\alpha + \alpha^*c + \beta^*\gamma + \gamma^*\beta, \\
 \langle \psi | I \otimes \sigma_2 | \psi \rangle &= i(-c\alpha + \alpha^*c - \beta^*\gamma + \gamma^*\beta), & \langle \psi | I \otimes \sigma_3 | \psi \rangle &= c^2 - |\alpha|^2 + |\beta|^2 - |\gamma|^2, \\
 \langle \psi | I \otimes I | \psi \rangle &= c^2 + |\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1. & & 
 \end{aligned} \tag{B.81}$$

**Exercise 3.17** If we write the state of the qubit at time  $t$  as

$$|\psi(t)\rangle = \alpha(t)|0\rangle + \beta(t)|1\rangle, \tag{B.82}$$

then equation (3.17) can be expressed as

$$i\hbar \frac{d}{dt} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = -\mu \begin{bmatrix} H_0 & H_1 e^{-i\omega t} \\ H_1 e^{i\omega t} & -H_0 \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix}. \tag{B.83}$$

If we define  $\omega_0 \equiv -2\mu H_0/\hbar$  and  $\omega_1 \equiv -2\mu H_1/\hbar$ , the we can write the Schrödinger equation (B.83) in the form

$$\begin{cases} i \frac{d}{dt} \alpha(t) = \frac{\omega_0}{2} \alpha(t) + \frac{\omega_1}{2} e^{-i\omega t} \beta(t), \\ i \frac{d}{dt} \beta(t) = \frac{\omega_1}{2} e^{i\omega t} \alpha(t) - \frac{\omega_0}{2} \beta(t). \end{cases} \tag{B.84}$$

Equations (B.84) constitute a linear homogeneous system with time-dependent coefficients. To solve this system, it is convenient to define new functions

$$a(t) \equiv \alpha(t) \exp(i\omega t/2) \quad \text{and} \quad b(t) \equiv \beta(t) \exp(-i\omega t/2). \tag{B.85}$$

If we introduce the vector

$$|\tilde{\psi}(t)\rangle = a(t)|0\rangle + b(t)|1\rangle, \tag{B.86}$$

we can easily see that

$$|\tilde{\psi}(t)\rangle = R_z(-\omega t)|\psi(t)\rangle, \tag{B.87}$$

where the rotation matrix  $R_z(-\omega t)$  was defined in Eq. (3.51) and represents a rotation of the Bloch sphere through an angle  $-\omega t$  about the axis  $z$ . Therefore, the transformation (B.85) corresponds to the change from a fixed reference frame to a frame rotating with the frequency  $\omega$  of the oscillating magnetic field. Substituting (B.85) into (B.84), we obtain

$$\begin{cases} i \frac{d}{dt} a(t) = \left(\frac{\omega_0 - \omega}{2}\right) a(t) + \frac{\omega_1}{2} b(t), \\ i \frac{d}{dt} b(t) = \frac{\omega_1}{2} a(t) - \left(\frac{\omega_0 - \omega}{2}\right) b(t). \end{cases} \tag{B.88}$$

Note that this new system of equations has constant coefficients. It corresponds to the Schrödinger equation in the rotating frame and can also be written as

$$i\hbar \frac{d}{dt} |\tilde{\psi}(t)\rangle = \tilde{H} |\tilde{\psi}(t)\rangle, \quad (\text{B.89})$$

where the Hamiltonian

$$\tilde{H} = \frac{\hbar}{2} \begin{bmatrix} \omega_0 - \omega & \omega_1 \\ \omega_1 & -(\omega_0 - \omega) \end{bmatrix} \quad (\text{B.90})$$

is time-independent. Thus, we obtain

$$|\tilde{\psi}(t)\rangle = \tilde{U} |\tilde{\psi}(0)\rangle = \tilde{U} |\psi(0)\rangle, \quad (\text{B.91})$$

where the unitary time-evolution operator  $\tilde{U}$  is given by

$$\tilde{U} = e^{-i\tilde{H}t/\hbar} = e^{-i[(\omega_0 - \omega)\sigma_z + \omega_1\sigma_x]t/2}. \quad (\text{B.92})$$

Finally, we obtain the formal solution to the Schrödinger equation (3.17):

$$|\psi(t)\rangle = R_z(\omega t) |\tilde{\psi}(t)\rangle = e^{-i\omega\sigma_z t/2} e^{-i[(\omega_0 - \omega)\sigma_z + \omega_1\sigma_x]t/2} |\psi(0)\rangle. \quad (\text{B.93})$$

Since the Hamiltonian  $\tilde{H}$  does not explicitly depend on time, we can write the solution of (3.17) in the form (2.29):

$$|\psi(t)\rangle = \sum_{n=1}^2 c_n(0) \exp\left(-\frac{i}{\hbar} E_n t\right) |n\rangle. \quad (\text{B.94})$$

Here  $E_1$  and  $E_2$  denote the eigenvalues of the Hamiltonian  $\tilde{H}$  and the coefficients  $c_n(0)$  ( $n = 1, 2$ ) are given by

$$c_n(0) = \langle \varphi_n | \tilde{\psi}(0) \rangle = \langle \varphi_n | \psi(0) \rangle, \quad (\text{B.95})$$

$|\varphi_1\rangle$  and  $|\varphi_2\rangle$  being the eigenvectors of  $\tilde{H}$ . It is easy to find that

$$E_1 = \frac{\hbar}{2} \sqrt{(\omega_0 - \omega)^2 + \omega_1^2}, \quad E_2 = -\frac{\hbar}{2} \sqrt{(\omega_0 - \omega)^2 + \omega_1^2}, \quad (\text{B.96})$$

$$|\varphi_1\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{bmatrix}, \quad |\varphi_2\rangle = \begin{bmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{bmatrix}, \quad (\text{B.97})$$

where the angle  $\theta$  is defined by

$$\tan \theta = \frac{\omega_1}{\omega_0 - \omega_1}. \quad (\text{B.98})$$

Let us assume that at time  $t = 0$  the system is in the state

$$|\psi(0)\rangle = |\tilde{\psi}(0)\rangle = |0\rangle. \quad (\text{B.99})$$

This corresponds to

$$c_1(0) = \langle \varphi_1 | 0 \rangle = \cos \frac{\theta}{2}, \quad c_2(0) = \langle \varphi_2 | 0 \rangle = -\sin \frac{\theta}{2}. \quad (\text{B.100})$$

After substitution of (B.100) into the general solution (B.94), we obtain

$$|\tilde{\psi}(t)\rangle = \cos \frac{\theta}{2} e^{-iE_1 t/\hbar} |\varphi_1\rangle - \sin \frac{\theta}{2} e^{-iE_2 t/\hbar} |\varphi_2\rangle. \quad (\text{B.101})$$

We can now compute the probability  $p_1(t)$  of finding the spin- $\frac{1}{2}$  particle in the state  $|1\rangle$  at time  $t$ . We obtain

$$\begin{aligned} p_1(t) &= |\langle 1 | \psi(t) \rangle|^2 = |\beta(t)|^2 = |b(t)|^2 = |\langle 1 | \tilde{\psi}(t) \rangle|^2 \\ &= \frac{\omega_1^2}{(\omega_0 - \omega)^2 + \omega_1^2} \sin^2 \left( \sqrt{(\omega_0 - \omega)^2 + \omega_1^2} \frac{t}{2} \right). \end{aligned} \quad (\text{B.102})$$

This probability is equal to 0 when  $t = 0$  and varies sinusoidally between 0 and  $\omega_1^2/[(\omega_0 - \omega)^2 + \omega_1^2]$ . These oscillations take place with frequency  $\Omega = \sqrt{(\omega_0 - \omega)^2 + \omega_1^2}$ . Formula (B.102) is known as *Rabi's formula* and  $\Omega$  is the Rabi frequency. The resonant case  $\omega = \omega_0$  is particularly important. In this case, the state of the particle oscillates with frequency  $\Omega = \omega_1$  between the states  $|0\rangle$  and  $|1\rangle$ . We have  $p_1(t) = 1$  at times  $t = (2n+1)\pi/\omega_1$ . Note that far from resonance the transition probability between the states  $|0\rangle$  and  $|1\rangle$  remains small, that is, the probability to measure the  $z$ -component of the spin and obtain  $\sigma_z = -1$  is small at all times.

**Exercise 3.18** To show that the  $C^2 - U$  gate can be decomposed as in Fig. 3.11, a tedious but straightforward method is to compute the products of the matrices associated with the gates that build up the decomposition. Since these gates act non-trivially only on two of the three qubits, it is necessary to embed them in the  $2^3$ -dimensional Hilbert space associated with the whole system. The simplest example of embedding is when a  $2 \times 2$  matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (\text{B.103})$$

is associated with a linear operator acting on a single qubit and we wish to extend it to the Hilbert space associated with two qubits. We have two possibilities, depending on whether the operator  $A$  acts on the less or more significant qubit:

$$I \otimes A = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}, \quad A \otimes I = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}. \quad (\text{B.104})$$

In the present exercise, we must embed  $4 \times 4$  matrices, associated with operators acting on two qubits, in the Hilbert space for three qubits. Given a generic  $4 \times 4$  matrix

$$M = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}, \quad (\text{B.105})$$

the embedding gives one of the following three matrices:

$$\begin{bmatrix} a & b & c & d & 0 & 0 & 0 & 0 \\ e & f & g & h & 0 & 0 & 0 & 0 \\ i & j & k & l & 0 & 0 & 0 & 0 \\ m & n & o & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & b & c & d \\ 0 & 0 & 0 & 0 & e & f & g & h \\ 0 & 0 & 0 & 0 & i & j & k & l \\ 0 & 0 & 0 & 0 & m & n & o & p \end{bmatrix}, \quad \begin{bmatrix} a & b & 0 & 0 & c & d & 0 & 0 \\ e & f & 0 & 0 & g & h & 0 & 0 \\ 0 & 0 & a & b & 0 & 0 & c & d \\ 0 & 0 & e & f & 0 & 0 & g & h \\ i & j & 0 & 0 & k & l & 0 & 0 \\ m & n & 0 & 0 & o & p & 0 & 0 \\ 0 & 0 & i & j & 0 & 0 & k & l \\ 0 & 0 & m & n & 0 & 0 & o & p \end{bmatrix}, \quad \begin{bmatrix} a & 0 & b & 0 & c & 0 & d & 0 \\ 0 & a & 0 & b & 0 & c & 0 & d \\ e & 0 & f & 0 & g & 0 & h & 0 \\ 0 & e & 0 & f & 0 & g & 0 & h \\ i & 0 & j & 0 & k & 0 & l & 0 \\ 0 & i & 0 & j & 0 & k & 0 & l \\ m & 0 & n & 0 & o & 0 & p & 0 \\ 0 & m & 0 & n & 0 & o & 0 & p \end{bmatrix}, \quad (B.106)$$

depending on whether the operator  $M$  acts trivially on the first (most significant), second, or third (least significant) qubit.

A much simpler way to solve this exercise is to compute the action of the five gates that build the circuit of Fig. 3.11 on the states  $|i_2 i_1 i_0\rangle$  of the computational basis and show that their composition is equivalent to the application of the  $C^2 - U$  gate. This way we have

$$\begin{aligned} |00i_0\rangle &\rightarrow |00i_0\rangle \rightarrow |00i_0\rangle \rightarrow |00i_0\rangle \rightarrow |00i_0\rangle \rightarrow |00i_0\rangle, \\ |01i_0\rangle &\rightarrow |01\rangle V |i_0\rangle \rightarrow |01\rangle V |i_0\rangle \rightarrow |01i_0\rangle \rightarrow |01i_0\rangle \rightarrow |01i_0\rangle, \\ |10i_0\rangle &\rightarrow |10i_0\rangle \rightarrow |11i_0\rangle \rightarrow |11\rangle V^\dagger |i_0\rangle \rightarrow |10\rangle V^\dagger |i_0\rangle \rightarrow |10i_0\rangle, \\ |11i_0\rangle &\rightarrow |11\rangle V |i_0\rangle \rightarrow |10\rangle V |i_0\rangle \rightarrow |10\rangle V |i_0\rangle \rightarrow |11\rangle V |i_0\rangle \rightarrow |11\rangle U |i_0\rangle, \end{aligned} \quad (B.107)$$

where we have used the relations  $V^2 = U$  and  $V^\dagger V = I = VV^\dagger$ . Thus, the circuit of Fig. 3.11 indeed implements the  $C^2 - U$  gate, since it acts non-trivially only when the two control qubits are set to 1 and in this case applies a  $U$  gate to the third qubit.

**Exercise 3.19** We must proceed as in exercise 3.11, but the NOT gates must be applied (before and after a standard  $C^{(n-1)}$ -NOT gate) to all the qubits that induce non-trivial action of the generalized  $C^{(n-1)}$ -NOT gate when they are in the state  $|0\rangle$ .

**Exercise 3.20** The matrix  $D$  for a  $4 \times 4$  matrix is given by

$$\begin{bmatrix} \cos \phi_1 & 0 & -\sin \phi_1 & 0 \\ 0 & \cos \phi_2 & 0 & -\sin \phi_2 \\ \sin \phi_1 & 0 & \cos \phi_1 & 0 \\ 0 & \sin \phi_2 & 0 & \cos \phi_2 \end{bmatrix}. \quad (B.108)$$

The circuit in Fig. 3.15 implements the transformation

$$(\text{CNOT}) (R_y(\theta_1) \otimes I) (\text{CNOT}) (R_y(\theta_0) \otimes I), \quad (B.109)$$

where

$$R_y(\theta) \otimes I = \begin{bmatrix} \cos \frac{\theta}{2} & 0 & -\sin \frac{\theta}{2} & 0 \\ 0 & \cos \frac{\theta}{2} & 0 & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ 0 & \sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} \end{bmatrix}, \quad (B.110)$$

and the generalized CNOT gate is given by

$$\overline{\text{CNOT}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (\text{B.111})$$

It is easy to check the equality between (B.108) and (B.109) by direct matrix multiplication, provided that  $\theta_0 = \phi_1 + \phi_2$  and  $\theta_1 = \phi_1 - \phi_2$ .

**Exercise 3.21** We have

$$U_f^2 |x\rangle|y\rangle = U_f|x\rangle|y \oplus f(x)\rangle = |x\rangle|y \oplus 2f(x)\rangle = |x\rangle|y\rangle, \quad (\text{B.112})$$

where  $|x\rangle \equiv |x_{n-1}, x_{n-2}, \dots, x_0\rangle$ . Hence,  $U_f^2 = I$ , that is,

$$U_f^{-1} = U_f. \quad (\text{B.113})$$

Let us show that  $U_f$  is Hermitian. The matrix elements of  $U_f$  in the computational basis are given by

$$U_f(x, y; x', y') = \langle x|\langle y|U_f|x'\rangle|y'\rangle = \langle x|x'\rangle\langle y|y' \oplus f(x')\rangle = \delta_{x,x'}\delta_{y,y' \oplus f(x)}. \quad (\text{B.114})$$

We now compute the matrix elements of the adjoint operator  $U_f^\dagger$ :

$$U_f^\dagger(x, y; x', y') = U_f^*(x', y'; x, y) = U_f(x', y'; x, y) = \delta_{x,x'}\delta_{y',y \oplus f(x)}. \quad (\text{B.115})$$

Since  $y = y' \oplus f(x)$  can be written equivalently as  $y \oplus f(x) = y' \oplus 2f(x) = y'$ , it follows that

$$U_f^\dagger(x, y; x', y') = U_f(x, y; x', y'). \quad (\text{B.116})$$

Equations (B.113) and (B.116) imply that

$$U_f^\dagger = U_f^{-1}, \quad (\text{B.117})$$

that is,  $U_f$  is unitary.

**Exercise 3.22** We have a Boolean function with two qubits ( $N = 2$ ), therefore we can start the AQC algorithm with the Hamiltonian  $H_I^{(2)}$  of Eq. (3.188). In order to write the final Hamiltonian, we have to build up the energy function  $h(a, b)$  for each clause composing  $f(a, b)$ . In this specific case we have

clause	$(a, b) = (0, 0)$	$(a, b) = (0, 1)$	$(a, b) = (1, 0)$	$(a, b) = (1, 1)$
$C_1 = \bar{a} \vee b$	$h_{C_1} = 0$	$h_{C_1} = 0$	$h_{C_1} = 1$	$h_{C_1} = 0$
$C_2 = a$	$h_{C_2} = 1$	$h_{C_2} = 1$	$h_{C_2} = 0$	$h_{C_2} = 0$
$C_3 = \bar{a} \vee \bar{b}$	$h_{C_3} = 0$	$h_{C_3} = 0$	$h_{C_3} = 0$	$h_{C_3} = 1$
$C_4 = \bar{b}$	$h_{C_4} = 0$	$h_{C_4} = 1$	$h_{C_4} = 0$	$h_{C_4} = 1$

Therefore the final Hamiltonian is given by:

$$H_F^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}. \quad (\text{B.118})$$

The adiabatic connection between the two Hamiltonians is provided by

$$H(t) = \left(1 - \frac{t}{T}\right) H_I^{(2)} + \left(\frac{t}{T}\right) H_F^{(2)} = \frac{1}{2} \begin{bmatrix} 2 & t-1 & t-1 & 0 \\ t-1 & 2(t+1) & 0 & t-1 \\ t-1 & 0 & 2 & t-1 \\ 0 & t-1 & t-1 & 2(t+1) \end{bmatrix} \quad (\text{B.119})$$

Its eigenenergies are:  $E_0^\pm = \frac{1}{2}(1 + 2x \pm \sqrt{1 - 2x + 2x^2})$  and  $E_1^\pm = \frac{1}{2}(3 \pm \sqrt{1 - 2x + 2x^2})$ , where  $x = t/T$ .

## B.4 Chapter 4

**Exercise 4.1** It is clear from Eq. (4.34) that the probability that Grover's algorithm fails is given by

$$p(x \neq x_0) = \cos^2[(2k+1)\theta], \quad (\text{B.120})$$

where  $\theta \approx 1/\sqrt{N}$  and

$$(2k+1)\theta = \frac{\pi}{2} + O(\theta) = \frac{\pi}{2} + O\left(\frac{1}{\sqrt{N}}\right). \quad (\text{B.121})$$

It follows that

$$p(x \neq x_0) = \cos^2\left[\frac{\pi}{2} + O\left(\frac{1}{\sqrt{N}}\right)\right] = O\left(\frac{1}{N}\right). \quad (\text{B.122})$$

**Exercise 4.2** One step of Grover's algorithm requires an oracle query,  $2n$  Hadamard gates and a reflection about the hyperplane orthogonal to  $|0\rangle$ . To operate this reflection, we need to implement a generalized  $C^{(n-1)}\text{-MINUS}$  gate, which puts a minus sign in front of the state vector  $|00\cdots 0\rangle$ . As we saw in Sec. 3.7 (see Fig. 3.12), this transformation can be decomposed into  $2(n-2)$  Toffoli gates plus a single CMINUS gate. The price to pay is that  $n-2$  ancillary qubits are required. Alternatively, it is possible to compute the  $C^{(n-1)}\text{-MINUS}$  gate without ancillary qubits in  $O(n^2)$  elementary gates (see Barenco *et al.*, 1995). Finally, we assume that the oracle answers the query instantaneously, that is, the time it takes to operate is not included in the complexity analysis. This is because the cost of the oracle call depends upon the specific application. In conclusion, one step of Grover's algorithm takes a time (measured in the number of elementary gates) of the order of  $n$  (with ancillary qubits) or of the order of  $n^2$  (without ancillaries). Since Grover's algorithm requires  $O(\sqrt{N})$  steps, in the first case we need  $O(\sqrt{N} \log N)$  elementary gates, in the second  $O(\sqrt{N}(\log N)^2)$  elementary gates.

**Exercise 4.3** Let us call  $|\psi\rangle$  the exact wave function at the end of the quantum Fourier transform and  $|\tilde{\psi}\rangle$  the actual wave function when unitary errors take place. We know from Sec. 3.8 that, if errors are uniformly bound at each step by some constant  $\delta$ , then

$$\||\tilde{\psi}\rangle - |\psi\rangle\| < n_g \delta, \quad (\text{B.123})$$

where  $n_g = O(n^2)$  is the number of elementary quantum gates required to implement the Fourier transform. If the desired accuracy in the output state is  $\epsilon$ , it is sufficient to implement the single quantum gates with precision  $\delta$  such that  $n_g\delta < \epsilon$ . Therefore,  $\delta = O(1/n^2)$ , namely, it drops only polynomially with the number of qubits. An interesting consequence is the following (Coppersmith, 1994). If a final accuracy  $\epsilon$  is required, we can simply skip from the quantum Fourier transform algorithm the controlled-phase shift gates  $R_k$  of angles  $2\pi i/2^k < \epsilon/n_g$ . This is just because these  $R_k$  gates differ from the identity by less than  $\epsilon/n_g$ . This observation ensures that it is not necessary to perform controlled-phase shift gates with exponentially small phases; a polynomial control of the phases is sufficient.

**Exercise 4.4** Since  $F^{-1}F = I$ , it is sufficient to run the circuit in Fig. 4.5 from right to left.

**Exercise 4.5** We have  $3^6 = 729 = 1 \bmod 91$ , and therefore the order  $r = 6$  is even. Moreover,  $3^{r/2} = 27 \neq \pm 1 \bmod 91$ , so that  $\gcd(27 - 1, 91) = 13$  and  $\gcd(27 + 1, 91) = 7$  are factors of  $N = 91$ .

## B.5 Chapter 5

**Exercise 5.1** Let  $f_A$  and  $f_B$  denote the public encryption keys whose corresponding secret decryption keys  $f_A^{-1}$  and  $f_B^{-1}$  are possessed by Alice alone and by Bob alone, respectively. Alice encrypts her signature  $S_A$  by means of her secret decryption key  $f_A^{-1}$  into  $f_A^{-1}(S_A)$ . She then encrypts the plain text  $P$  plus her signature using Bob's public encryption key. This way she produces the cypher text  $C = f_B(P + f_A^{-1}(S_A))$ , which is sent over a public channel to Bob. Bob then decrypts the cypher text  $C$  by means of his secret key  $f_B^{-1}$  as follows:  $f_B^{-1}(C) = P + f_A^{-1}(S_A)$ . After this he uses the public key  $f_A$  to verify Alice's signature  $S_A = f_A(f_A^{-1}(S_A))$ . We underline that the above authentication procedure is valid because only Alice knows her secret key  $f_A^{-1}$ : nobody else could have produced  $f_A^{-1}(S_A)$ .

**Exercise 5.2** The average fidelity is given by

$$\bar{f} = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta |\langle \psi_1 | \psi_2 \rangle|^2, \quad (\text{B.124})$$

where the spherical coordinates  $\theta$  and  $\phi$  single out the state  $|\psi_1\rangle$  on the Bloch sphere (see Sec. 2.1). To compute the integral (B.124) it is convenient to choose the  $z$ -axis along the polarization direction of  $|\psi_2\rangle$ . This choice corresponds to  $|\psi_2\rangle = |0\rangle$  and therefore we have

$$\bar{f} = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \cos^2 \frac{\theta}{2} = \frac{1}{2}. \quad (\text{B.125})$$

**Exercise 5.3** Alice and Bob share the EPR state  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . If Alice measures  $\sigma_x$ , she obtains outcomes  $\pm 1$  with equal probabilities  $p_+^{(A)} = p_-^{(A)} = \frac{1}{2}$ . After Alice's measurement, the state of Bob's qubit is

$$|\psi_\pm\rangle_B = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle), \quad (\text{B.126})$$

where the  $\pm$  sign corresponds to the  $\pm 1$  result of Alice's measurement. If Bob performs a spin measurement along the  $\mathbf{u}$  axis singled out by the spherical coordinates  $\theta$  and  $\phi$ , he

obtains  $\sigma_u = +1$  with probability

$$|\mathbf{u}\langle +|\psi_{\pm}\rangle_B|^2 = \frac{1}{2}[1 \pm \cos\phi \sin\theta], \quad (\text{B.127})$$

where  $|+\rangle_u$  is the eigenvector of the operator  $\sigma_u$  corresponding to the eigenvalue +1 (the explicit expressions for  $\sigma_u$  and  $|+\rangle_u$  are given by Eq. (2.48) and Eq. (2.49), respectively). The eigenvector  $|-\rangle_u$  of  $\sigma_u$  corresponding to the eigenvalue -1 is obtained from  $|+\rangle_u$  via the transformation  $\phi \rightarrow \phi + \pi$  and  $\theta \rightarrow \pi - \theta$ . Hence, Bob obtains  $\sigma_u = -1$  with probability

$$|\mathbf{u}\langle -|\psi_{\pm}\rangle_B|^2 = \frac{1}{2}[1 \pm \cos(\phi + \pi) \sin(\pi - \theta)] = \frac{1}{2}[1 \mp \cos\phi \sin\theta]. \quad (\text{B.128})$$

Since Bob receives the states  $|\psi_{\pm}\rangle_B$  with probabilities  $\frac{1}{2}$ , then he obtains  $\sigma_u = +1$  with probability

$$p_+^{(B)} = \frac{1}{2}|\mathbf{u}\langle +|\psi_+\rangle_B|^2 + \frac{1}{2}|\mathbf{u}\langle +|\psi_-\rangle_B|^2 = \frac{1}{2}, \quad (\text{B.129a})$$

and  $\sigma_u = -1$  with probability

$$p_-^{(B)} = \frac{1}{2}|\mathbf{u}\langle -|\psi_+\rangle_B|^2 + \frac{1}{2}|\mathbf{u}\langle -|\psi_-\rangle_B|^2 = \frac{1}{2}. \quad (\text{B.129b})$$

If instead Alice measures  $\sigma_z$ , then the state of Bob's qubit collapses onto  $|\phi_+\rangle_B = |0\rangle$  or  $|\phi_-\rangle_B = |1\rangle$  with probabilities  $\frac{1}{2}$ . Since

$$\begin{aligned} |\mathbf{u}\langle +|\phi_+\rangle_B|^2 &= \cos^2 \frac{\theta}{2}, & |\mathbf{u}\langle +|\phi_-\rangle_B|^2 &= \sin^2 \frac{\theta}{2}, \\ |\mathbf{u}\langle -|\phi_+\rangle_B|^2 &= \sin^2 \frac{\theta}{2}, & |\mathbf{u}\langle -|\phi_-\rangle_B|^2 &= \cos^2 \frac{\theta}{2}, \end{aligned} \quad (\text{B.130})$$

then Bob's spin measurement along the  $\mathbf{u}$  axis gives outcomes  $\pm 1$  with probabilities

$$p_+^{(B)} = \frac{1}{2}, \quad p_-^{(B)} = \frac{1}{2}. \quad (\text{B.131})$$

This result is trivially independent of the chosen axis  $\mathbf{u}$ . It is important to note that the outcomes  $p_+^{(B)}$  and  $p_-^{(B)}$  always have the same probability. This implies that the EPR phenomenon cannot be used for faster than light communication. Whatever axis Alice and Bob choose for their measurements, Bob always obtains randomly +1 or -1. Therefore, no information has been transmitted from Alice to Bob.

**Exercise 5.4** Let us first consider the preparation of the two bottom qubits in Fig. 5.2, initially prepared in the state  $|00\rangle$ . To simplify writing, we adopt  $C_i \equiv \cos\theta_i$  and  $S_i \equiv \sin\theta_i$ . After application of the first rotation matrix we obtain the state

$$C_1|00\rangle + S_1|10\rangle. \quad (\text{B.132})$$

The first CNOT gate leads to

$$C_1|00\rangle + S_1|11\rangle, \quad (\text{B.133})$$

the second rotation to

$$C_1C_2|00\rangle + C_1S_2|01\rangle - S_1S_2|10\rangle + S_1C_2|11\rangle, \quad (\text{B.134})$$

the second CNOT to

$$C_1 C_2 |00\rangle + C_1 S_2 |11\rangle - S_1 S_2 |10\rangle + S_1 C_2 |01\rangle, \quad (\text{B.135})$$

and the third rotation to

$$\begin{aligned} & C_1 C_2 C_3 |00\rangle + C_1 C_2 S_3 |10\rangle - C_1 S_2 S_3 |01\rangle + C_1 S_2 C_3 |11\rangle \\ & + S_1 S_2 S_3 |00\rangle - S_1 S_2 C_3 |10\rangle + S_1 C_2 C_3 |01\rangle + S_1 C_2 S_3 |11\rangle. \end{aligned} \quad (\text{B.136})$$

From (5.22) we have  $C_1 = \sqrt{\frac{\sqrt{5}+1}{2\sqrt{5}}}$ ,  $C_2 = \sqrt{\frac{3+\sqrt{5}}{6}}$ ,  $C_3 = \sqrt{\frac{\sqrt{5}+2}{2\sqrt{5}}}$ ,  $S_1 = \sqrt{\frac{\sqrt{5}-1}{2\sqrt{5}}}$ ,  $S_2 = \sqrt{\frac{3-\sqrt{5}}{6}}$ ,  $S_3 = \sqrt{\frac{\sqrt{5}-2}{2\sqrt{5}}}$ . Inserting these numerical values in (B.136) we obtain the state vector (5.23).

We now discuss the copying part of the circuit in Fig. 5.2. The four CNOT gates map the state

$$(\alpha|0\rangle + \beta|1\rangle) \frac{1}{\sqrt{6}} (2|00\rangle + |01\rangle + |11\rangle) \quad (\text{B.137})$$

into (5.24). Tracing over the bottom qubit in Fig. 5.2 we obtain the two-qubit density matrix

$$\begin{aligned} \rho_{12} = & \left[ \alpha \sqrt{\frac{2}{3}} |00\rangle + \beta \sqrt{\frac{1}{6}} (|10\rangle + |01\rangle) \right] \left[ \alpha^* \sqrt{\frac{2}{3}} \langle 00| + \beta^* \sqrt{\frac{1}{6}} (\langle 10| + \langle 01|) \right] \\ & + \left[ \beta \sqrt{\frac{2}{3}} |11\rangle + \alpha \sqrt{\frac{1}{6}} (|10\rangle + |01\rangle) \right] \left[ \beta^* \sqrt{\frac{2}{3}} \langle 11| + \alpha^* \sqrt{\frac{1}{6}} (\langle 10| + \langle 01|) \right]. \end{aligned} \quad (\text{B.138})$$

Hence

$$\begin{aligned} \rho_1 &= \text{Tr}_2 \rho_{12} \\ &= \frac{2}{3} |\alpha|^2 |0\rangle \langle 0| + \frac{1}{3} \alpha \beta^* |0\rangle \langle 1| + \frac{1}{6} |\beta|^2 |0\rangle \langle 0| + \frac{1}{3} \beta \alpha^* |1\rangle \langle 0| + \frac{1}{6} |\beta|^2 |1\rangle \langle 1| \\ &+ \frac{2}{3} |\beta|^2 |1\rangle \langle 1| + \frac{1}{3} \beta \alpha^* |1\rangle \langle 0| + \frac{1}{6} |\alpha|^2 |0\rangle \langle 0| + \frac{1}{3} \alpha \beta^* |0\rangle \langle 1| + \frac{1}{6} |\alpha|^2 |1\rangle \langle 1|. \end{aligned} \quad (\text{B.139})$$

It is easy to see that this expression is equal to (5.27). Since the two-qubit density matrix  $\rho_{12}$  is symmetric under exchange of the two qubits, we have  $\rho_2 = \rho_1$ .

**Exercise 5.5** We must consider only the cases in which Alice and Bob used the same alphabet, since the raw key shared by Alice and Bob comes from here. We have the following four cases:

Case	1	2	3	4
Alice's data bits	0	0	1	1
Alphabet	$x$	$z$	$x$	$z$
Transmitted qubits	$ +\rangle$	$ 0\rangle$	$ -\rangle$	$ 1\rangle$

Eve measures the spin polarization  $\sigma_u$  along an arbitrary direction singled out by the spherical coordinates  $\theta$  and  $\phi$ . For the sake of simplicity, we consider the case  $\phi = 0$  (generalization to arbitrary  $\phi$  is straightforward). She obtains one bit of information as follows: if the outcome of her measurement is  $\sigma_u = +1$ , she decides that the bit value is 0; if the outcome is  $\sigma_u = -1$ , the bit is 1. The probabilities of these two outcomes are given by

$$p_0^{(i)} = |\mathbf{u} \langle + | \psi^{(i)} \rangle|^2, \quad p_1^{(i)} = |\mathbf{u} \langle - | \psi^{(i)} \rangle|^2, \quad (\text{B.140})$$

where  $|+\rangle_u$  and  $|-\rangle_u$  are the eigenstates of  $\sigma_u$  corresponding to the eigenvalues +1 and -1 (their explicit expressions are given by Eq. (2.49)), and the index  $(i)$  denotes one of the four possible transmitted qubits, that is,  $|\psi^{(1)}\rangle = |+\rangle$ ,  $|\psi^{(2)}\rangle = |0\rangle$ ,  $|\psi^{(3)}\rangle = |-\rangle$  and  $|\psi^{(4)}\rangle = |1\rangle$ . We have eight possibilities ( $p_0^{(i)}$  and  $p_1^{(i)}$ , with  $i = 1, \dots, 4$ ), which take place with the following probabilities:

$$\begin{aligned} p_0^{(1)} &= \frac{1}{2}(1 + \sin \theta \cos \phi) = p_1^{(3)}, & p_1^{(1)} &= \frac{1}{2}(1 - \sin \theta \cos \phi) = p_0^{(3)}, \\ p_0^{(2)} &= \cos^2 \frac{\theta}{2} = p_1^{(4)}, & p_1^{(2)} &= \sin^2 \frac{\theta}{2} = p_0^{(4)}. \end{aligned} \quad (\text{B.141})$$

After her measurement, Eve resends the state  $|+\rangle_u$  or the state  $|-\rangle_u$  with the probabilities  $p_0^{(i)}$  and  $p_1^{(i)}$ , respectively. Bob measures this state in the same basis as Alice's original basis (remember that we are interested in the bits that constitute the raw key). There are sixteen possible cases, corresponding to the state obtained by Bob, provided that Alice sent a given state and Eve resent another state. The error rate is obtained by adding all the cases in which the bit obtained by Bob differs from the original Alice's bit. Using this procedure, it is easy to check that the error rate is  $\frac{1}{4}$ .

**Exercise 5.6** In order to measure the state  $|\psi_1\rangle$  without disturbing it, we must measure an observable such that  $|\psi_1\rangle$  is an eigenstate of the Hermitian operator associated with the observable. As we wish to gain information about which state was sent by Alice without disturbing the state, we require that  $|\psi_1\rangle$  and  $|\psi_2\rangle$  both be eigenstates of the same Hermitian operator, corresponding to two different eigenvalues. However, these requirements cannot be fulfilled, because we have assumed that the states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  are not orthogonal. Therefore,  $|\psi_1\rangle$  and  $|\psi_2\rangle$  cannot be eigenstates of the same operator and any measurement necessarily disturbs at least one of the two states.

**Exercise 5.7** We start from the initial state  $|\psi\rangle|0\rangle|0\rangle$ , where  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , and compute the quantum gates represented in Fig. 5.9. Let us write down explicitly the first few steps and the final result:

$$\begin{aligned} |\psi\rangle|0\rangle|0\rangle &\rightarrow |\psi\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle \rightarrow \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle)(|00\rangle + |11\rangle) \\ &\rightarrow \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle) \\ &\rightarrow \dots \\ &\rightarrow \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)(\alpha|0\rangle + \beta|1\rangle). \end{aligned} \quad (\text{B.142})$$

Given the final state of (B.142), it is also possible to apply two Hadamard gates to the first two qubits and end up with the state  $|0\rangle|0\rangle|\psi\rangle$ , which coincides with the initial state, except for a permutation of the qubit states.

**Exercise 5.8** Similarly, to the previous exercise, we write down the action of the first few quantum gates of the circuit drawn in Fig. 5.10 and the final result:

$$\begin{aligned}
 & \frac{1}{\sqrt{2}} (\alpha |01\rangle + \beta |10\rangle) (|000\rangle + |111\rangle) \\
 & \rightarrow \frac{1}{2} (\alpha |00\rangle - \alpha |01\rangle + \beta |10\rangle + \beta |11\rangle) (|000\rangle + |111\rangle) \\
 & \rightarrow \frac{1}{2} (\alpha |00000\rangle + \alpha |00111\rangle - \alpha |01100\rangle - \alpha |01011\rangle \\
 & \quad + \beta |10000\rangle + \beta |10111\rangle + \beta |11100\rangle + \beta |11011\rangle) \\
 & \rightarrow \dots \\
 & \rightarrow \frac{1}{2} |1\rangle (|00\rangle + |01\rangle + |10\rangle + |11\rangle) (\alpha |01\rangle + \beta |10\rangle). \tag{B.143}
 \end{aligned}$$

**Exercise 5.9** The wave function (5.72) is a solution of the Schrödinger equation for a one-dimensional free particle since

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) = -\frac{\hbar^2}{2m} \left\{ \left[ -\frac{(x - x_0 - p_0 \frac{t}{m})}{\delta^2 (1 + i \frac{\hbar t}{m \delta^2})} + i \frac{p_0}{\hbar} \right]^2 - \frac{1}{\delta^2 (1 + i \frac{\hbar t}{m \delta^2})} \right\}. \tag{B.144}$$

We know from the postulates of quantum mechanics that the average value of any observable  $O$  is given by  $\langle O \rangle = \langle \psi | O | \psi \rangle$ . In particular, we have

$$\langle X(t) \rangle = \langle \psi(t) | X | \psi(t) \rangle = \int_{-\infty}^{+\infty} dx x |\psi(x, t)|^2. \tag{B.145}$$

If we define  $x' = x - x_0 - \frac{p_0}{m}t$ , we obtain

$$\langle X(t) \rangle = \int_{-\infty}^{+\infty} dx' \left( x' + x_0 + \frac{p_0}{m}t \right) f(x'), \tag{B.146}$$

where

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{B.147}$$

with  $\sigma = \frac{\delta}{\sqrt{2}} \sqrt{1 + \frac{\hbar^2 t^2}{m^2 \delta^4}}$ . Using

$$\int_{-\infty}^{+\infty} dx f(x) = 1, \quad \int_{-\infty}^{+\infty} dx x f(x) = 0, \tag{B.148}$$

we have

$$\langle X(t) \rangle = x_0 + \frac{p_0}{m}t. \tag{B.149}$$

Similarly, we obtain

$$\langle P(t) \rangle = \langle \psi(t) | P | \psi(t) \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x, t) \left( -i\hbar \frac{\partial}{\partial x} \right) \psi(x, t) = p_0. \tag{B.150}$$

The variances  $(\Delta X)^2 = \langle (X - \langle X \rangle)^2 \rangle$  and  $(\Delta P)^2 = \langle (P - \langle P \rangle)^2 \rangle$  are computed in the same manner: the result of Eq. (5.75) is obtained using (B.148) and

$$\int_{-\infty}^{+\infty} dx x^2 f(x) = \sigma^2. \tag{B.151}$$

**Exercise 5.10** From the properties of the Hermite polynomials, given in the formulæ in the footnote on page 222, we derive that

$$H_{n+1}(\xi) = \left(2\xi - \frac{d}{d\xi}\right) H_n(\xi). \quad (\text{B.152})$$

We have

$$a = \frac{1}{\sqrt{2}} \left( \xi + \frac{d}{d\xi} \right), \quad a^\dagger = \frac{1}{\sqrt{2}} \left( \xi - \frac{d}{d\xi} \right), \quad (\text{B.153})$$

where  $\xi = \sqrt{\frac{m\omega}{\hbar}} x$ . It follows immediately that

$$a^\dagger \phi_n(\xi) = \sqrt{n+1} \phi_{n+1}(\xi), \quad a \phi_n(\xi) = \sqrt{n} \phi_{n-1}(\xi). \quad (\text{B.154})$$

Note that the iterative application of the first of these relations leads to

$$\phi_n(\xi) = \frac{a^\dagger}{\sqrt{n!}} \phi_0(\xi). \quad (\text{B.155})$$

**Exercise 5.11** (i) We have

$$\langle n|a|\alpha\rangle = \alpha \langle n|\alpha\rangle = \sqrt{n+1} \langle n+1|\alpha\rangle. \quad (\text{B.156})$$

This relation can be iterated, thus obtaining

$$\langle n|\alpha\rangle = \frac{\alpha^n}{\sqrt{n!}} \langle 0|\alpha\rangle, \quad (\text{B.157})$$

which implies

$$|\alpha\rangle = \sum_{n=0}^{\infty} |n\rangle \langle n|\alpha\rangle = \langle 0|\alpha\rangle \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (\text{B.158})$$

Normalization of this state ( $\langle \alpha|\alpha\rangle = 1$ ) leads to  $|\langle 0|\alpha\rangle|^2 = e^{-|\alpha|^2}$ , so that state (5.85) is obtained.

(ii) We have

$$p(n) \equiv |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2}. \quad (\text{B.159})$$

Therefore, the probability  $p(n)$  of finding  $n$  photons in  $|\alpha\rangle$  is given by a Poisson distribution and the mean photon number

$$\bar{n} = \langle \alpha|N|\alpha\rangle = \sum_{n=0}^{\infty} np(n) = e^{-|\alpha|^2} \sum_{n=0}^{\infty} n \frac{(|\alpha|^2)^n}{n!} = e^{-|\alpha|^2} |\alpha|^2 e^{|\alpha|^2} = |\alpha|^2, \quad (\text{B.160})$$

where  $N = a^\dagger a$  is the number operator. We also obtain

$$\Delta n = \sqrt{\langle \alpha|N^2|\alpha\rangle - \langle \alpha|N|\alpha\rangle} = \sqrt{\bar{n}}. \quad (\text{B.161})$$

(iii) We obtain

$$(\Delta X)^2 = \frac{\hbar}{2m\omega}, \quad (\Delta P)^2 = \frac{m\hbar\omega}{2}, \quad (\text{B.162})$$

so that  $\Delta X \Delta P = \frac{\hbar}{2}$ , independently of  $\alpha$ .

The temporal evolution is governed by the Schrödinger equation and we have

$$\begin{aligned} |\psi(t)\rangle &= e^{-\frac{i}{\hbar}Ht}|\alpha\rangle = e^{-i\omega(N+\frac{1}{2})t}e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \\ &= e^{-\frac{i}{2}\omega t}e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha e^{-i\omega t})^n}{\sqrt{n!}}|n\rangle = e^{-\frac{i}{2}\omega t}|\alpha(t)\rangle, \end{aligned} \quad (\text{B.163})$$

where  $\alpha(t) = e^{-i\omega t}\alpha$ . Therefore, the wave packet remains a minimum uncertainty coherent state at all times, and its centre in phase space follows the classical path in time. This is made more explicitly by computing the expectation values of  $X$  and  $P$ :

$$\begin{aligned} \langle X(t)\rangle &= \sqrt{\frac{2\hbar}{m\omega}}\operatorname{Re}(\alpha(t)) = \langle X(0)\rangle \cos(\omega t) + \frac{\langle P(0)\rangle}{m\omega} \sin(\omega t), \\ \langle P(t)\rangle &= \sqrt{2m\hbar\omega}\operatorname{Im}(\alpha(t)) = \langle P(0)\rangle \cos(\omega t) - m\omega\langle X(0)\rangle \sin(\omega t). \end{aligned} \quad (\text{B.164})$$

(iv) Equation (5.87) is readily derived by using the representation (5.85) of coherent states in the Fock basis. To derive the closure relation (5.88), it is convenient to use again (5.85) and then write the left-hand side of (5.88) in polar coordinates in the  $\alpha$  plane (setting  $\alpha = \rho e^{i\phi}$ ), thus obtaining

$$\frac{1}{\pi}\int_0^\infty d\rho \rho \int_0^{2\pi} d\phi e^{-\rho^2} \sum_{n,m} e^{i(n-m)\phi} \frac{\rho^{n+m}}{\sqrt{n!m!}} |n\rangle\langle m| = \sum_n I_n \frac{1}{n!} |n\rangle\langle n| = I, \quad (\text{B.165})$$

where we have used

$$\int_0^{2\pi} d\phi e^{i(n-m)\phi} = 2\pi\delta_{nm}, \quad (\text{B.166})$$

$$I_n \equiv 2\int_0^\infty d\rho \rho e^{-\rho^2} \rho^{2n} = \int_0^\infty du e^{-u} u^n = nI_{n-1} = n!I_0 = n!. \quad (\text{B.167})$$

**Exercise 5.12** We obtain, to the first order in  $\delta\alpha$ ,

$$D^\dagger(\delta\alpha) a D(\delta\alpha) = a + [a, \delta\alpha a^\dagger - \delta\alpha^* a] = a + \delta\alpha I. \quad (\text{B.168})$$

After writing a generic displacement operator  $D(\alpha)$  as the product of the infinitesimal displacements  $D(\delta\alpha)$ , with  $\alpha = \sum \delta\alpha$ , we obtain Eq. (5.94). Finally, we have

$$a D(-\alpha) |\alpha\rangle = D(-\alpha) D^\dagger(-\alpha) a D(-\alpha) |\alpha\rangle = D(-\alpha) (a - \alpha I) |\alpha\rangle = 0, \quad (\text{B.169})$$

where the last equality follows from the definition of a coherent state, Eq. (5.84). The fact that  $a D(-\alpha) |\alpha\rangle = 0$  implies that  $D(-\alpha) |\alpha\rangle$  is the vacuum state. We can conclude that coherent states are displaced vacuums,  $|\alpha\rangle = D(\alpha) |0\rangle$ .

**Exercise 5.13** For a coherent state,

$$(\Delta X_1)^2 = \langle\alpha|X_1^2|\alpha\rangle - \langle\alpha|X_1|\alpha\rangle^2 = \frac{1}{4}\langle\alpha|[a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2]|\alpha\rangle - \frac{1}{4}\langle\alpha|(a + a^\dagger)|\alpha\rangle^2 = \frac{1}{4}. \quad (\text{B.170})$$

Similarly,  $(\Delta X_2)^2 = \frac{1}{4}$ , so that  $\Delta X_1 \Delta X_2 = \frac{1}{4}$ .

For a Fock state, an analogous calculation gives

$$(\Delta X_1)^2 = \langle n|X_1^2|n\rangle - \langle n|X_1|n\rangle^2 = \frac{1}{4}(2n+1) = (\Delta X_2)^2. \quad (\text{B.171})$$

Therefore, the Fock states are not minimum uncertainty states, with the exception of the vacuum state  $|0\rangle$ , which is also a coherent state.

**Exercise 5.14** From the first terms of the Baker-Campbell-Hausdorff expansion (A.101) we have

$$\tilde{a} = S(z)aS^\dagger(z) = B + [A, B] + \frac{1}{2!} [A, [A, B]] + \dots, \quad (\text{B.172})$$

By repeatedly using the commutation relation  $[a, a^\dagger] = 1$ , we obtain

$$[A, B] = \left( \frac{1}{2} z^* a^2 - \frac{1}{2} z a^{\dagger 2} \right) a - a \left( \frac{1}{2} z^* a^2 - \frac{1}{2} z a^{\dagger 2} \right) = z a^\dagger, \quad (\text{B.173})$$

and therefore

$$\tilde{a} = a + z a^\dagger + \frac{1}{2} z z^* a + \dots = a \left( 1 + \frac{1}{2} r^2 \right) + e^{i\phi} a^\dagger r + \dots. \quad (\text{B.174})$$

We recognize here the first terms in the expansion of  $\cosh(r)$  and  $\sinh(r)$ . By considering all terms in the Baker-Campbell-Hausdorff expansion, we derive Eq. (5.121).

**Exercise 5.15**

$$S(z)|0\rangle = \exp \left( -\frac{1}{2} a^{\dagger 2} e^{i\phi} \tanh r \right) \exp \left[ -\frac{1}{2} \ln(\cosh r) \right] |0\rangle, \quad (\text{B.175})$$

which after Taylor expansion of the first exponent can be seen to be equal to (5.122).

**Exercise 5.16**

$$\int_{-\infty}^{\infty} dp W(x, p) = \int_{-\infty}^{\infty} dy \left\langle x + \frac{y}{2} \middle| \rho \middle| x - \frac{y}{2} \right\rangle \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{-\frac{i}{\hbar} py}, \quad (\text{B.176})$$

where we have assumed that the integrals over  $y$  and  $p$  can be interchanged. Using relation (A.153), we obtain

$$\int_{-\infty}^{\infty} dp W(x, p) = \int_{-\infty}^{\infty} dy \left\langle x + \frac{y}{2} \middle| \rho \middle| x - \frac{y}{2} \right\rangle \delta(y) = \langle x | \rho | x \rangle. \quad (\text{B.177})$$

Analogously, we obtain

$$\int_{-\infty}^{\infty} dx W(x, p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx'' e^{-\frac{i}{\hbar} p(x'' - x')} \langle x'' | \rho | x' \rangle = \langle p | \rho | p \rangle, \quad (\text{B.178})$$

where we have introduced the new integration variables  $x' = x - \frac{1}{2}y$  and  $x'' = x + \frac{1}{2}y$  and the last equality follows from Eq. (A.180) connecting the position and momentum representations of an operator.

**Exercise 5.17** The  $g_F = 0$  curve is an ellipse as in Eq. (5.142), with

$$\bar{x}(t) = x - x_0 - \frac{p_0 t}{m}, \quad a = \frac{1}{\delta^2}, \quad b = \frac{\delta^2}{\hbar^2} \left( 1 + \frac{\hbar^2 t^2}{m^2 \delta^4} \right), \quad c = -\frac{t}{m \delta^2}. \quad (\text{B.179})$$

As shown in Fig. B.3, the shape of the ellipse changes over time. However, its area

$$A = \frac{\pi}{\sqrt{ab - c^2}} = \pi\hbar \quad (\text{B.180})$$

remains constant.

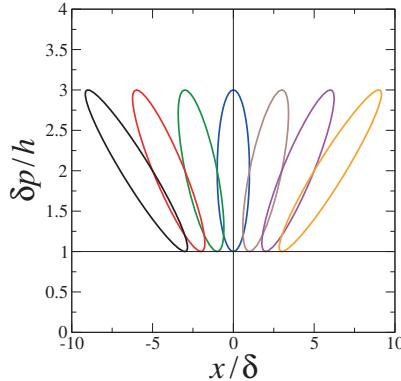


Fig. B.3 Phase space curves  $g_F = 0$  at different times (from left to right,  $\tau = \frac{\hbar t}{m\delta^2} = -3, -2, -1, 0, 1, 2, 3$ ) for a free particle, whose state at  $t = 0$  is a minimum uncertainty Gaussian wave packet centered at  $(\frac{x_0}{\delta} = 0, \frac{\delta p_0}{\hbar} = 2)$ .

**Exercise 5.18** Let us start by imposing the commutation relation  $[b_i, b_j^\dagger] = \delta_{ij}$ . Using the definition of  $\{b_j, b_j^\dagger\}_{j=1, \dots, N}$  given in Eq. (5.163), we get:

$$\begin{aligned} [b_i, b_j^\dagger] &= \left[ \sum_k (F_{ik}a_k + G_{ik}a_k^\dagger) + \alpha_i, \sum_l (F_{jl}^*a_l^\dagger + G_{jl}^*a_l) + \alpha_j^* \right] \\ &= \sum_{kl} \left( F_{ik}F_{jl}^* [a_k, a_l^\dagger] + F_{ik}G_{jl}^* [a_k, a_l] + G_{ik}F_{jl}^* [a_k^\dagger, a_l^\dagger] + G_{ik}G_{jl}^* [a_k^\dagger, a_l] \right) \\ &= \sum_{kl} F_{ik}F_{jl}^* \delta_{kl} + G_{ik}G_{jl}^* (-\delta_{kl}) = \sum_k F_{ik}F_{jk}^* - G_{ik}G_{jk}^* = [FF^\dagger - GG^\dagger]_{ij}, \end{aligned} \quad (\text{B.181})$$

from which it follows that  $FF^\dagger - GG^\dagger = I$ , where  $I$  is the  $N \times N$  identity matrix. Notice that in the third equality, we used the commutation relations for the bosonic operators  $a_j^{(\dagger)}$ :  $[a_i, a_j^\dagger] = \delta_{ij}$  and  $[a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0$ .

The second constraint can be obtained from the commutation relation  $[b_i, b_j] = 0$ :

$$\begin{aligned} [b_i, b_j] &= \left[ \sum_k (F_{ik}a_k + G_{ik}a_k^\dagger) + \alpha_i, \sum_l (F_{jl}a_l + G_{jl}a_l^\dagger) + \alpha_j \right] \\ &= \sum_{kl} \left( F_{ik}F_{jl} [a_k, a_l] + F_{ik}G_{jl} [a_k, a_l^\dagger] + G_{ik}F_{jl} [a_k^\dagger, a_l] + G_{ik}G_{jl} [a_k^\dagger, a_l^\dagger] \right) \\ &= \sum_{kl} F_{ik}G_{jl} \delta_{kl} + G_{ik}F_{jl} (-\delta_{kl}) = \sum_k F_{ik}G_{jk} - G_{ik}F_{jk} = [FG^T - GF^T]_{ij}, \end{aligned} \quad (\text{B.182})$$

where again in the third equality, we used the commutation relations for the  $a_j$  bosons.

## B.6 Chapter 6

**Exercise 6.1** For a separable state, the decomposition (6.2) holds. Therefore,

$$\langle(\Sigma_i)^2\rangle = \text{Tr}\left[\sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)} (\Sigma_i)^2\right] = \sum_k p_k \text{Tr}\left[\rho_A^{(k)} \otimes \rho_B^{(k)} (\Sigma_i)^2\right] = \sum_k p_k \langle(\Sigma_i)^2\rangle_k, \quad (\text{B.183})$$

where  $\langle \dots \rangle_k$  denotes the average over the density operator  $\rho_A^{(k)} \otimes \rho_B^{(k)}$ . We then obtain

$$\begin{aligned} \langle(\Delta\Sigma_i)^2\rangle &= \sum_k p_k \langle(\Sigma_i)^2\rangle_k - \langle\Sigma_i\rangle^2 \\ &= \sum_k p_k \left[ \langle(\sigma_A^i)^2\rangle_k + 2\langle\sigma_A^i\rangle_k \langle\sigma_B^i\rangle_k + \langle(\sigma_B^i)^2\rangle_k \right] - \langle\Sigma_i\rangle^2 \\ &= \sum_k p_k \left[ \langle(\Delta\sigma_A^i)^2\rangle_k + \langle(\Delta\sigma_B^i)^2\rangle_k + \langle\Sigma_i\rangle_k^2 \right] - \left( \sum_k p_k \langle\Sigma_i\rangle_k \right)^2. \end{aligned} \quad (\text{B.184})$$

Finally, we apply the Cauchy–Schwarz inequality

$$\sum_l p_l \sum_k p_k \langle\Sigma_i\rangle_k^2 \geq \left( \sum_k p_k \langle\Sigma_i\rangle_k \right)^2 \quad (\text{B.185})$$

to see that the last two terms in (B.184) are bounded from below by zero. Hence,

$$\langle(\Delta\Sigma_i)^2\rangle \geq \sum_k p_k \left[ \langle(\Delta\sigma_A^i)^2\rangle_k + \langle(\Delta\sigma_B^i)^2\rangle_k \right]. \quad (\text{B.186})$$

As this latter inequality is valid for  $i = x, y, z$ , we obtain

$$\langle(\Delta\Sigma_x)^2\rangle + \langle(\Delta\Sigma_y)^2\rangle + \langle(\Delta\Sigma_z)^2\rangle \geq. \quad (\text{B.187})$$

$$\sum_k p_k \left[ \langle(\Delta\sigma_A^x)^2\rangle_k + \langle(\Delta\sigma_A^y)^2\rangle_k + \langle(\Delta\sigma_A^z)^2\rangle_k + \langle(\Delta\sigma_B^x)^2\rangle_k + \langle(\Delta\sigma_B^y)^2\rangle_k + \langle(\Delta\sigma_B^z)^2\rangle_k \right].$$

Note that we have bounded from below a sum of variances of a two-qubit state by a sum of variances of single-qubit states, which are quantities much easier to compute experimentally. We now compute, for a single qubit,  $\langle(\Delta\sigma_x)^2\rangle_k + \langle(\Delta\sigma_y)^2\rangle_k + \langle(\Delta\sigma_z)^2\rangle_k$ . After writing the single-qubit density matrix as in (3.11), it is easy to check that

$$\langle(\Delta\sigma_x)^2\rangle_k + \langle(\Delta\sigma_y)^2\rangle_k + \langle(\Delta\sigma_z)^2\rangle_k = 3 - (x^2 + y^2 + z^2) \geq 2. \quad (\text{B.188})$$

Finally, we substitute this inequality into (B.187) and obtain (6.10). If a given state  $\rho_{AB}$  does not satisfy the inequality (6.10), we can therefore conclude that  $\rho_{AB}$  is entangled.

In the case of the Werner state (6.7) we obtain by direct computation

$$\langle\Sigma_i\rangle = \text{Tr}[(\rho_W)_{AB} \Sigma_i] = 0, \quad \langle(\Sigma_i)^2\rangle = \text{Tr}[(\rho_W)_{AB} (\Sigma_i)^2] = 2 - 2q, \quad (\text{B.189})$$

with  $i = x, y, z$ . Therefore,

$$\langle(\Delta\Sigma_x)^2\rangle + \langle(\Delta\Sigma_y)^2\rangle + \langle(\Delta\Sigma_z)^2\rangle = 6 - 6q, \quad (\text{B.190})$$

so that the inequality (6.10) is satisfied when  $q \leq \frac{1}{3}$ . We can therefore conclude that the Werner state is entangled for  $\frac{1}{3} < q \leq 1$ . Note that the separability criterion (6.10) is much easier to use in experiments than the Peres criterion since it only requires that the single particle polarizations are measured. The Peres criterion can instead be applied only after that the entire two-qubit density matrix is reconstructed. Finally, we point out that a separability criterion for continuous variable systems can be derived following the same procedure used in this exercise (see Duan *et al.*, 2000).

**Exercise 6.2** To find the maximum Shannon entropy, defined by Eq. (6.25), we must solve the system

$$\begin{cases} \frac{\partial}{\partial p_j} \left[ -\sum_{i=1}^k p_i \log p_i - \lambda \left( \sum_{i=1}^k p_i - 1 \right) \right] = 0, & (j = 1, \dots, k), \\ \sum_{i=1}^k p_i = 1, \end{cases} \quad (\text{B.191})$$

where the Lagrange multiplier  $\lambda$  is introduced to take into account the constraint  $\sum_i p_i = 1$ . The system of equations (B.191) is solved for  $p_1 = \dots = p_k = 1/k$ .

**Exercise 6.3** We first write the relative entropy as

$$D(p||q) = -\sum_x p(x) \log \frac{q(x)}{p(x)}. \quad (\text{B.192})$$

Hence

$$D(p||q) \geq \frac{1}{\ln 2} \sum_x p(x) \left( 1 - \frac{q(x)}{p(x)} \right) = \frac{1}{\ln 2} \sum_x [p(x) - q(x)] = 0 \quad (\text{B.193})$$

since  $\sum_x p(x) = \sum_x q(x) = 1$ . Note that the equality occurs if and only if  $p(x) = q(x)$  for all  $x$ .

**Exercise 6.4** From Eq. (6.29) we can see that  $H(Y|X) \geq 0$  since  $\log p(y|x) \leq 0$ , and therefore  $H(X, Y) = X(X) + H(Y|X) \geq H(X)$ .

**Exercise 6.5** We use the inequality  $\log x \leq (x - 1)/\ln 2$ . We have

$$\begin{aligned} H(X) + X(Y) - X(X, Y) &= -\sum_{x,y} p(x, y) \log \frac{p(x)p(y)}{p(x, y)} \\ &\geq \frac{1}{\ln 2} \sum_{x,y} p(x, y) \left( 1 - \frac{p(x)p(y)}{p(x, y)} \right) = \frac{1}{\ln 2} [\sum_{x,y} p(x, y) - \sum_{x,y} p(x)p(y)] = 0. \end{aligned} \quad (\text{B.194})$$

From this proof we can see that we have the equality  $H(X, Y) = H(X) + H(Y)$  if and only if  $p(x, y) = p(x)p(y)$  for all  $x, y$ , namely if the probability distributions  $p(x)$  and  $p(y)$  are independent.

**Exercise 6.6** Using the suggested spectral decomposition, we obtain

$$\begin{aligned} D(\rho||\sigma) &= \sum_i p_i \log p_i - \sum_i \langle i | \rho \log \sigma | i \rangle \\ &= \sum_i p_i \log p_i - \sum_i p_i \langle i | \log \sigma | i \rangle = \sum_i p_i \left( \log p_i - \sum_{\alpha} A_{i\alpha} \log q_{\alpha} \right), \end{aligned} \quad (\text{B.195})$$

where  $A_{i\alpha} = \langle i | \alpha \rangle \langle \alpha | i \rangle = |\langle i | \alpha \rangle|^2 \geq 0$  and  $\sum_i A_{i\alpha} = \sum_{\alpha} A_{i\alpha} = 1$  readily follows from the completeness relation for the bases  $\{|i\rangle\}$  and  $\{|\alpha\rangle\}$ , respectively. Due to the concavity of the logarithm,  $\sum_{\alpha} A_{i\alpha} \log q_{\alpha} \leq \log r_i$ , with  $r_i = \sum_{\alpha} A_{i\alpha} q_{\alpha}$  and the equality holds if and only if  $A_{i\alpha} = 1$  for some  $\alpha$ . Therefore

$$D(\rho||\sigma) \geq \sum_i [p_i (\log p_i - \log r_i)] = D(p|r) \geq 0, \quad (\text{B.196})$$

where  $D(p|r)$  is the (classical) relative entropy of the probability distribution  $\{p_i\}$  relative to  $\{r_i\}$ . We know from exercise 6.3 that  $D(p|r) = 0$  if and only if  $p_i = r_i$  for all  $i$ . We can conclude that  $D(\rho||\sigma) = 0$  if and only if  $A_{ij}$  is a permutation matrix. In this case, after relabeling the eigenstates of  $\sigma$ ,  $A$  becomes the identity matrix, namely  $\rho$  and  $\sigma$  are diagonal in the same basis. From the condition  $p_i = r_i$  we then conclude that the spectral decompositions of  $\rho$  and  $\sigma$  are identical and therefore  $\rho = \sigma$ .

**Exercise 6.7** The quantum relative entropy

$$\begin{aligned} D(\rho_{AB}||\rho_A \otimes \rho_B) &= \text{Tr}[\rho_{AB} (\log \rho_{AB} - \log(\rho_A \otimes \rho_B))] \\ &= -S(\rho_{AB}) - \text{Tr}[\rho_{AB} \log(\rho_A \otimes \rho_B)] \geq 0, \end{aligned} \quad (\text{B.197})$$

with equality if and only if  $\rho_{AB} = \rho_A \otimes \rho_B$ . Moreover,

$$-\text{Tr}[\rho_{AB} \log(\rho_A \otimes \rho_B)] = -\text{Tr}(\rho_A \log \rho_A) - \text{Tr}(\rho_B \log \rho_B) = S(\rho_A) + S(\rho_B). \quad (\text{B.198})$$

We can therefore conclude that  $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ , with equality if and only if  $\rho_{AB} = \rho_A \otimes \rho_B$ .

**Exercise 6.8** We have

$$S(\rho_A) = S\left(\sum_i p_i \rho_i\right), \quad S(\rho_B) = S\left(\sum_i p_i |i\rangle\langle i|\right) = H(p_1, p_2, \dots). \quad (\text{B.199})$$

Let  $\lambda_i^{\alpha}$  and  $|u_i^{\alpha}\rangle$  be the eigenvalues and the corresponding eigenvectors of  $\rho_i$ . Since the states  $\rho_i |i\rangle\langle i|$  have support on orthogonal subspaces,  $p_i \lambda_i^{\alpha}$  and  $|u_i^{\alpha}\rangle$  are the eigenvalues and eigenvectors of  $\rho_{AB}$ . Hence

$$\begin{aligned} S(\rho_{AB}) &= -\sum_{i,\alpha} p_i \lambda_i^{\alpha} \log(p_i \lambda_i^{\alpha}) \\ &= -\sum_i p_i \log p_i - \sum_i p_i \sum_{\alpha} \lambda_i^{\alpha} \log \lambda_i^{\alpha} = H(p_1, p_2, \dots) + \sum_i p_i S(\rho_i). \end{aligned} \quad (\text{B.200})$$

Using the subadditivity inequality  $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ , we finally obtain  $\sum_i p_i S(\rho_i) \leq S(\sum_i p_i \rho_i)$ .

**Exercise 6.9** The solution follows steps already used for exercise 6.8. Let  $\lambda_i^\alpha$  and  $|u_i^\alpha\rangle$  be the eigenvalues and the corresponding eigenvectors of  $\rho_i$ . Since the states  $\rho_i$  have support on orthogonal subspaces, then  $p_i\lambda_i^\alpha$  and  $|u_i^\alpha\rangle$  are the eigenvalues and eigenvectors of  $\rho$  and therefore

$$S(\rho) = - \sum_{i,\alpha} p_i \lambda_i^\alpha \log(p_i \lambda_i^\alpha) = H(p_1, p_2, \dots) + \sum_i p_i S(\rho_i). \quad (\text{B.201})$$

**Exercise 6.10** Since  $|\psi\rangle_{ABC}$  is a pure state,  $S(\rho_A) = S(\rho_{BC})$  and  $S(\rho_C) = S(\rho_{AB})$ . Applying subadditivity to subsystem  $BC$ , we obtain

$$S(\rho_A) = S(\rho_{BC}) \leq S(\rho_B) + S(\rho_C) = S(\rho_B) + S(\rho_{AB}), \quad (\text{B.202})$$

and therefore  $S(\rho_{AB}) \geq S(\rho_A) - S(\rho_B)$ . The same reasoning with  $A$  and  $B$  interchanged leads to  $S(\rho_{AB}) \geq S(\rho_B) - S(\rho_A)$ .

**Exercise 6.11** Given a bipartition of the Hilbert space of the system into two parts,  $A$  and  $B$ , with dimensions  $N_A$  and  $N_B$ , the purity reads

$$P = \sum_{j,j'=0}^{N_A-1} \sum_{l,l'}^{N_B-1} r_{jl} r_{j'l} r_{j'l'} r_{jl'} \exp[i(\phi_{jl} - \phi_{j'l} + \phi_{j'l'} - \phi_{jl'})], \quad (\text{B.203})$$

where  $|k\rangle = |jl\rangle = |j\rangle_A \otimes |l\rangle_B$ . It is convenient to split  $P$  in two parts:

$$P = X + M, \quad (\text{B.204})$$

where

$$X = \sum'_{j,j'} \sum'_{l,l'} r_{jl} r_{j'l} r_{j'l'} r_{jl'} \exp[i(\phi_{jl} - \phi_{j'l} + \phi_{j'l'} - \phi_{jl'})], \quad (\text{B.205})$$

$$M = \sum'_{j,j'} \sum_l r_{jl}^2 r_{j'l}^2 + \sum_j \sum'_{l,l'} r_{jl}^2 r_{j'l'}^2 + \sum_{j,l} r_{jl}^4, \quad (\text{B.206})$$

where  $\sum'$  means that equal indexes are banned in the sum. Since  $\langle e^{i\phi_k} \rangle = 0$  we obtain  $\langle X \rangle = 0$ . Therefore,

$$P = \langle M \rangle = N(N_A + N_B - 2)\langle r_0^2 r_1^2 \rangle + N\langle r_0^4 \rangle, \quad (\text{B.207})$$

where we have used  $\langle r_k^4 \rangle = \langle r_0^4 \rangle$  for all  $k$  and  $\langle r_k^2 r_{k'}^2 \rangle = \langle r_0^2 r_1^2 \rangle$  for all  $k, k'$  with  $k \neq k'$ .

We now evaluate the marginal distribution

$$\begin{aligned} p(r_0, \dots, r_{m-1}) &= C_N r_0 \cdots r_{m-1} \int_0^1 dr_m r_m \int_0^1 dr_{m+1} r_{m+1} \cdots \int_0^1 dr_{N-1} r_{N-1} \delta(r^2 - 1) \\ &= \frac{C_N}{2} r_0 \cdots r_{m-1} \int_0^{\sqrt{1-r_0^2-\dots-r_{m-1}^2}} dr_m r_m \cdots \int_0^{\sqrt{1-r_0^2-\dots-r_{N-3}^2}} dr_{N-2} r_{N-2} \\ &= \frac{C_N}{2^{N-m}} \frac{1}{(N-m-1)!} r_0 \cdots r_{m-1} \left(1 - \sum_{j=0}^{m-1} r_j^2\right)^{N-m-1}. \end{aligned} \quad (\text{B.208})$$

In particular,

$$p(r_0) = \frac{C_N}{2^{N-1}} \frac{1}{(N-2)!} r_0 (1 - r_0^2)^{N-2}. \quad (\text{B.209})$$

The normalization condition  $\int_0^1 dr_0 p(r_0) = 1$  allows us to determine  $C_N = 2^N(N-1)!$ . Thus, we obtain

$$p(r_0) = 2(N-1)r_0(1 - r_0^2)^{N-2}, \quad (\text{B.210})$$

$$\langle r_0^4 \rangle = \int_0^1 dr_0 r_0^4 p(r_0) = \frac{2}{N(N+1)}, \quad (\text{B.211})$$

$$p(r_0, r_1) = 4(N-1)(N-2)r_0 r_1 (1 - r_0^2 - r_1^2)^{N-3}, \quad (\text{B.212})$$

$$\langle r_0^2 r_1^2 \rangle = \int_0^1 dr_0 r_0^2 \int_0^1 dr_1 r_1^2 p(r_0, r_1) = \frac{1}{N(N+1)}. \quad (\text{B.213})$$

After substitution of Eqs. (B.211) and (B.213) into (B.207) we readily obtain Lubkin's formula (6.77).

The variance  $\sigma_P^2$  can be computed with the same technique as above. However, to obtain the variance (6.78) for large  $N$  it is sufficient to replace in Eqs. (B.205) and (B.206)  $r_k$  with its mean value  $1/\sqrt{N}$ :

$$\sigma_P^2 = \langle P^2 \rangle - P_L^2 = \langle X^2 \rangle + \langle M^2 \rangle - P_L^2 \approx \langle X^2 \rangle \approx \frac{2}{N^2}. \quad (\text{B.214})$$

We can see from Eqs. (B.205) and (B.206) that  $X$  and  $M$  are sums of  $O(N^2)$  terms of order  $1/N^2$ . Therefore, the central limit theorem implies that, for large  $N$ , the purity tends to a Gaussian distribution with mean  $P_L$  and variance  $\sigma_P$ .

**Exercise 6.12** We start by writing the state  $\rho_{AB}(\xi)$  in the computational basis for the two qubits:

$$\rho_{AB} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{\xi}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\xi}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad (\text{B.215})$$

which has eigenvalues  $\lambda_1 = \lambda_2 = 0$  and  $\lambda_{3,4} = \frac{1}{2}(1 \pm \xi)$ . The reduced density matrix  $\rho_A = \rho_B = \frac{1}{2}I$ , therefore the mutual information of  $\rho_{AB}$  can be immediately calculated:

$$\mathcal{I}(A:B) = 1 + 1 + \left(\frac{1}{2} - \xi\right) \log \left(\frac{1}{2} - \xi\right) + \left(\frac{1}{2} + \xi\right) \log \left(\frac{1}{2} + \xi\right). \quad (\text{B.216})$$

In order to evaluate the conditional entropy, let us define the two projectors  $\Pi_1$  and  $\Pi_2$  starting from two orthogonal states on the Bloch sphere, which depend on the angles  $\{\theta, \phi\}$ . Using the compact notation  $C = \cos \frac{\theta}{2}$  and  $S = \sin \frac{\theta}{2}$ , we have:

$$|1\rangle = \begin{bmatrix} C \\ e^{i\phi} S \end{bmatrix}, \quad |2\rangle = \begin{bmatrix} -e^{-i\phi} S \\ C \end{bmatrix}. \quad (\text{B.217})$$

These states identify the corresponding projectors

$$\Pi_1 = \begin{bmatrix} C^2 & e^{-i\phi} CS \\ e^{i\phi} CS & S^2 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} S^2 & -e^{-i\phi} CS \\ -e^{i\phi} CS & C^2 \end{bmatrix}. \quad (\text{B.218})$$

It is easily seen that  $\Pi_1 + \Pi_2 = I$ . From these expressions one can compute the following:

$$\{I \otimes \Pi_1\} \rho_{AB} \{I \otimes \Pi_1\} = \begin{bmatrix} C^4 & e^{-i\phi} C^3 S & e^{i\phi} C^3 S \xi & C^2 S^2 \xi \\ e^{i\phi} C^3 S & C^2 S^2 & e^{2i\phi} C^2 S^2 \xi & e^{i\phi} C S^3 \xi \\ e^{-i\phi} C^3 S \xi & e^{-2i\phi} C^2 S^2 \xi & C^2 S^2 & e^{-i\phi} C S^3 \\ C^2 S^2 \xi & e^{-i\phi} C S^3 \xi & e^{i\phi} C S^3 & S^4 \end{bmatrix}, \quad (\text{B.219a})$$

$$\{I \otimes \Pi_2\} \rho_{AB} \{I \otimes \Pi_2\} = \begin{bmatrix} S^4 & -e^{-i\phi} C S^3 & -e^{i\phi} C S^3 \xi & C^2 S^2 \xi \\ -e^{i\phi} C S^3 & C^2 S^2 & e^{2i\phi} C^2 S^2 \xi & -e^{i\phi} C^3 S \xi \\ -e^{-i\phi} C S^3 \xi & e^{-2i\phi} C^2 S^2 \xi & C^2 S^2 & -e^{-i\phi} C^3 S \\ C^2 S^2 \xi & -e^{-i\phi} C^3 S \xi & -e^{i\phi} C^3 S & C^4 \end{bmatrix}. \quad (\text{B.219b})$$

Calculating the traces, it is easy to find that  $p_1 = p_2 = \text{Tr}[\{I \otimes \Pi_1\} \rho_{AB} \{I \otimes \Pi_1\}] = \frac{1}{2}$ . The conditional states are thus obtained by performing the partial traces over  $B$  of the above matrices:

$$\rho_{A|\Pi_1} = \begin{bmatrix} C^2 & e^{i\phi} C S \xi \\ e^{-i\phi} C S \xi & S^2 \end{bmatrix}, \quad \rho_{A|\Pi_2} = \begin{bmatrix} S^2 & -e^{i\phi} C S \xi \\ -e^{-i\phi} C S \xi & C^2 \end{bmatrix}. \quad (\text{B.220})$$

Notice that the eigenvalues of such matrices do not depend of  $\phi$ , therefore we can fix  $\phi = 0$ . We can then obtain the von Neumann entropy of each of the above density matrices numerically, and use the result to calculate the classical correlation  $\mathcal{J}(A:B)$  in Eq. (6.117), from which eventually the quantum discord measure  $\mathcal{D}(A:B) = \mathcal{I}(A:B) - \mathcal{J}(A:B)$  follows.

The final result of this calculation is plotted in Fig. B.4, where it is possible to see that the optimal measurement basis which maximizes the classical correlations (that is, it minimizes the quantum discord), is obtained by placing  $\theta = 0$  in Eq. (B.217). Note also that the discord becomes zero only when both parameters  $\theta$  and  $\xi$  are zero. This means that there exists a preferred basis of measurement, for which the combined system exhibits a classical behaviour.

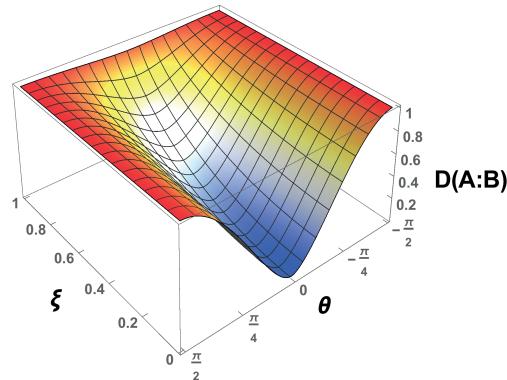


Fig. B.4 Quantum discord for the state  $\rho_{AB}(\xi)$ , as a function of  $\xi$  and of the angle  $\theta$  of the projective measurement basis  $\{|1\rangle, |2\rangle\}$  defined in Eq. (B.217).

**Exercise 6.13** Using the representation of Eq. (6.136) applied to the Werner state (6.7), it is not difficult to recognize that it can be cast in the following simple form:

$$\rho_W = \frac{1}{4} \left[ I \otimes I - q(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z) \right], \quad (\text{B.221})$$

therefore the only non-zero entries of matrix  $T_{\mu\nu}$  are  $T_{00} = 1$  and  $T_{11} = T_{22} = T_{33} = -q$ . Using now the formula for the geometric discord in Eq. (6.137), we obtain that  $\mathcal{D}_G = \frac{1}{4}(T_{11}^2 + T_{22}^2 + T_{33}^2 - \lambda_{\max})$ , where  $\lambda_{\max}$  is the maximum eigenvalue of the  $3 \times 3$  diagonal matrix  $L = \vec{T}\vec{T}^T = \text{diag}(T_{11}, T_{22}, T_{33})$ . Since all the three diagonal entries of  $\vec{T}$  are equal to  $-q$ , we have  $\lambda_{\max} = q^2$ , and thus  $\mathcal{D}_G = \frac{1}{4}(3q^2 - q^2) = q^2/2$ .

## B.7 Chapter 7

**Exercise 7.1**  $\mathbb{T}$  is positive since  $\rho'_1 = \mathbb{T}(\rho_1) = \rho_1^T$  has the same eigenvalues as  $\rho_1$ , and is therefore non-negative as is  $\rho_1$ . Let us now show that  $\mathbb{T}$  is not completely positive. The density operator  $\rho$  corresponding to the state (7.14) is given by

$$\begin{aligned} \rho_{1E} = & \frac{1}{2} \left( |0\rangle_1 \langle 0| \otimes |1\rangle_{EE} \langle 1| + |1\rangle_1 \langle 1| \otimes |0\rangle_{EE} \langle 0| \right. \\ & \left. + |0\rangle_1 \langle 1| \otimes |1\rangle_{EE} \langle 0| + |1\rangle_1 \langle 0| \otimes |0\rangle_{EE} \langle 1| \right). \end{aligned} \quad (\text{B.222})$$

Therefore,

$$\begin{aligned} \rho'_{1E} \equiv (\mathbb{T} \otimes \mathbb{I}_E)(\rho_{1E}) = & \frac{1}{2} \left( |0\rangle_1 \langle 0| \otimes |1\rangle_{EE} \langle 1| + |1\rangle_1 \langle 1| \otimes |0\rangle_{EE} \langle 0| \right. \\ & \left. + |1\rangle_1 \langle 0| \otimes |1\rangle_{EE} \langle 0| + |0\rangle_1 \langle 1| \otimes |0\rangle_{EE} \langle 1| \right). \end{aligned} \quad (\text{B.223})$$

The matrix representation of  $\rho'_{1E}$  in the computational basis is given by

$$\rho'_{1E} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad (\text{B.224})$$

whose eigenvalues are  $\lambda_0 = -\frac{1}{2}$  and  $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{2}$ . Therefore,  $\mathbb{T} \otimes \mathbb{I}_E$  is not positive, implying that  $\mathbb{T}$  is not completely positive.

**Exercise 7.2** The condition  $\sum_k E_k^\dagger E_k = I$  implies that

$$\sum_k (\gamma_k^* I + \mathbf{a}_k^* \cdot \boldsymbol{\sigma})(\gamma_k I + \mathbf{a}_k \cdot \boldsymbol{\sigma}) = I. \quad (\text{B.225})$$

Using the relation (see exercise 3.9)

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = (\mathbf{a} \cdot \mathbf{b})I + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}), \quad (\text{B.226})$$

we obtain

$$\begin{aligned} & \sum_k |\gamma_k|^2 I + \sum_k [\gamma_k^* \mathbf{a}_k + \gamma_k \mathbf{a}_k^*] \cdot \boldsymbol{\sigma} + \sum_k (\mathbf{a}_k^* \cdot \boldsymbol{\sigma})(\mathbf{a}_k \cdot \boldsymbol{\sigma}) \\ &= \sum_k |\gamma_k|^2 I + \sum_k [\gamma_k^* \mathbf{a}_k + \gamma_k \mathbf{a}_k^*] \cdot \boldsymbol{\sigma} + \sum_k (\mathbf{a}_k^* \cdot \mathbf{a}_k) I + i \sum_k (\mathbf{a}_k^* \times \mathbf{a}_k) \cdot \boldsymbol{\sigma} = I. \end{aligned} \quad (\text{B.227})$$

This implies that

$$\sum_k |\gamma_k|^2 + \sum_k (\mathbf{a}_k^* \cdot \mathbf{a}_k) = 1, \quad \sum_k (\gamma_k^* \mathbf{a}_k + \gamma_k \mathbf{a}_k^*) = i \sum_k \mathbf{a}_k \times \mathbf{a}_k^*. \quad (\text{B.228})$$

It is convenient to employ the Bloch sphere representation (7.27) for  $\rho$  and  $\rho' = \sum_k E_k \rho E_k^\dagger$ . We have

$$\begin{aligned} I + \mathbf{r}' \cdot \boldsymbol{\sigma} &= \sum_k (\gamma_k I + \mathbf{a}_k \cdot \boldsymbol{\sigma})(I + \mathbf{r} \cdot \boldsymbol{\sigma})(\gamma_k^* I + \mathbf{a}_k^* \cdot \boldsymbol{\sigma}) \\ &= \sum_k [|\gamma_k|^2 I + \gamma_k \mathbf{a}_k^* \cdot \boldsymbol{\sigma} + |\gamma_k|^2 (\mathbf{r} \cdot \boldsymbol{\sigma}) + \gamma_k (\mathbf{r} \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma}) + \gamma_k^* \mathbf{a}_k \cdot \boldsymbol{\sigma} \\ &\quad + (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma}) + \gamma_k^* (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \boldsymbol{\sigma}) + (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma})]. \end{aligned} \quad (\text{B.229})$$

Let us examine the eight terms of the right-hand side of this equation separately. The sum of the fourth and the seventh term simplifies as follows:

$$\begin{aligned} & \sum_k [\gamma_k (\mathbf{r} \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma}) + \gamma_k^* (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \boldsymbol{\sigma})] \\ &= \sum_k [\gamma_k (\mathbf{r} \cdot \mathbf{a}_k^*) I + i \gamma_k \boldsymbol{\sigma} \cdot \mathbf{r} \times \mathbf{a}_k^* + \gamma_k^* (\mathbf{r} \cdot \mathbf{a}_k) I + i \gamma_k^* \boldsymbol{\sigma} \cdot \mathbf{a}_k \times \mathbf{r}] \\ &= \sum_k [(\gamma_k \mathbf{a}_k^* + \gamma_k^* \mathbf{a}_k) \cdot \mathbf{r} I + i(\gamma_k^* \mathbf{a}_k - \gamma_k \mathbf{a}_k^*) \times \mathbf{r} \cdot \boldsymbol{\sigma}]. \end{aligned} \quad (\text{B.230})$$

The sixth term can be written as

$$\sum_k (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma}) = \sum_k [(\mathbf{a}_k \cdot \mathbf{a}_k^*) I + i \boldsymbol{\sigma} \cdot \mathbf{a}_k \times \mathbf{a}_k^*]. \quad (\text{B.231})$$

Finally, for the eighth term we have

$$\begin{aligned} \sum_k (\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \boldsymbol{\sigma})(\mathbf{a}_k^* \cdot \boldsymbol{\sigma}) &= \sum_k \{(\mathbf{a}_k \cdot \boldsymbol{\sigma})[(\mathbf{r} \cdot \mathbf{a}_k^*) I + i \boldsymbol{\sigma} \cdot \mathbf{r} \times \mathbf{a}_k^*]\} \\ &= \sum_k [(\mathbf{r} \cdot \mathbf{a}_k^*)(\mathbf{a}_k \cdot \boldsymbol{\sigma}) + i(\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \times \mathbf{a}_k^* \cdot \boldsymbol{\sigma})] \\ &= \sum_k \{(\mathbf{r} \cdot \mathbf{a}_k^*)(\mathbf{a}_k \cdot \boldsymbol{\sigma}) + i[(\mathbf{a}_k \cdot \mathbf{r} \times \mathbf{a}_k^*) I + i \boldsymbol{\sigma} \cdot \mathbf{a}_k \times (\mathbf{r} \times \mathbf{a}_k^*)]\} \\ &= \sum_k \{(\mathbf{r} \cdot \mathbf{a}_k^*)(\mathbf{a}_k \cdot \boldsymbol{\sigma}) + i(\mathbf{a}_k \cdot \mathbf{r} \times \mathbf{a}_k^*) I - \boldsymbol{\sigma} \cdot [(\mathbf{a}_k \cdot \mathbf{a}_k^*) \mathbf{r} - (\mathbf{a}_k \cdot \mathbf{r}) \mathbf{a}_k^*]\}. \end{aligned} \quad (\text{B.232})$$

Note that we have used the relation

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}, \quad (\text{B.233})$$

with  $\mathbf{a} \rightarrow \mathbf{a}_k$ ,  $\mathbf{b} \rightarrow \mathbf{r}$ , and  $\mathbf{c} \rightarrow \mathbf{a}_k^*$ . After insertion of Eqs. (B.230)–(B.232) into Eq. (B.229), we obtain

$$\begin{aligned} I + \mathbf{r}' \cdot \boldsymbol{\sigma} = & \sum_k \left\{ |\gamma_k|^2 I + (\gamma_k \mathbf{a}_k^* + \gamma_k^* \mathbf{a}_k) \cdot \boldsymbol{\sigma} + |\gamma_k|^2 (\mathbf{r} \cdot \boldsymbol{\sigma}) + (\gamma_k \mathbf{a}_k^* + \gamma_k^* \mathbf{a}_k) \cdot \mathbf{r} I \right. \\ & + i(\gamma_k^* \mathbf{a}_k - \gamma_k \mathbf{a}_k^*) \times \mathbf{r} \cdot \boldsymbol{\sigma} + (\mathbf{a}_k \cdot \mathbf{a}_k^*) I + i\boldsymbol{\sigma} \cdot \mathbf{a}_k \times \mathbf{a}_k^* + (\mathbf{r} \cdot \mathbf{a}_k^*)(\mathbf{a}_k \cdot \boldsymbol{\sigma}) \\ & \left. + i(\mathbf{r} \cdot \mathbf{a}_k^* \times \mathbf{a}_k) I - \boldsymbol{\sigma} \cdot [(\mathbf{a}_k \cdot \mathbf{a}_k^*) \mathbf{r} - (\mathbf{a}_k \cdot \mathbf{r}) \mathbf{a}_k^*] \right\}. \quad (B.234) \end{aligned}$$

We can simplify Eq. (B.234) taking advantage of the expressions (B.228). We obtain

$$\begin{aligned} \mathbf{r}' \cdot \boldsymbol{\sigma} = & \sum_k \left\{ 2i(\mathbf{a}_k \times \mathbf{a}_k^*) \cdot \boldsymbol{\sigma} + [| \gamma_k |^2 - (\mathbf{a}_k \cdot \mathbf{a}_k^*)] (\mathbf{r} \cdot \boldsymbol{\sigma}) + i(\gamma_k^* \mathbf{a}_k - \gamma_k \mathbf{a}_k^*) \times \mathbf{r} \cdot \boldsymbol{\sigma} \right. \\ & \left. + [(\mathbf{a}_k \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \mathbf{a}_k^*) + (\mathbf{a}_k^* \cdot \boldsymbol{\sigma})(\mathbf{r} \cdot \mathbf{a}_k)] \right\}. \quad (B.235) \end{aligned}$$

Therefore,  $\mathbf{r}' = M\mathbf{r} + \mathbf{c}$ , with

$$\mathbf{c} = 2i \sum_k (\mathbf{a}_k \times \mathbf{a}_k^*), \quad (B.236)$$

that is

$$c_j = 2i \sum_{klm} \epsilon_{jlm} a_{kl} a_{km}^*. \quad (B.237)$$

Similarly, one can see that  $M$  is given by Eq. (7.30).

**Exercise 7.3** Let  $A$ ,  $B$ ,  $C$  and  $B$  denote the quantum gates of Fig. 7.7, from left to right. We have

$$A = \begin{bmatrix} C & 0 & -S & 0 \\ 0 & C & 0 & -S \\ S & 0 & C & 0 \\ 0 & S & 0 & C \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C & 0 & S & 0 \\ 0 & C & 0 & S \\ -S & 0 & C & 0 \\ 0 & -S & 0 & C \end{bmatrix}, \quad (B.238)$$

where  $C \equiv \cos \frac{\theta}{2}$  and  $S \equiv \sin \frac{\theta}{2}$ . The action of the circuit in Fig. 7.7 is described by the unitary operator

$$U = BCBA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & C^2 - S^2 & 0 & -2CS \\ 0 & 0 & 1 & 0 \\ 0 & 2CS & 0 & C^2 - S^2 \end{bmatrix}. \quad (B.239)$$

We have  $\rho_{\text{fin}}^{(\text{tot})} = U\rho_{\text{in}}^{(\text{tot})}U^\dagger$ , with

$$\rho_{\text{in}}^{(\text{tot})} = \begin{bmatrix} \rho & 0 \\ 0 & 0 \end{bmatrix}. \quad (B.240)$$

The Kraus operator  $F_k$  is defined as  $F_k = {}_e\langle k|U|0\rangle_e$ , where the subscript  $e$  refers to the environmental qubit. Since  $(F_k)_{ij} = \langle ki|U|0j\rangle$ , it is easy to see that  $U$  is represented as

the block matrix

$$U = \begin{bmatrix} F_0 & .. \\ F_1 & .. \end{bmatrix}, \quad (\text{B.241})$$

where

$$F_0 = \begin{bmatrix} 1 & 0 \\ 0 & C^2 - S^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \cos \theta \end{bmatrix}, \quad F_1 = \begin{bmatrix} 0 & 0 \\ 0 & 2CS \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \sin \theta \end{bmatrix}. \quad (\text{B.242})$$

**Exercise 7.4** Following the same procedure of Sec. 7.2.3, we obtain

$$\rho' = |\alpha|^2 U \rho U^\dagger + |\beta|^2 \rho. \quad (\text{B.243})$$

Therefore, the Kraus operators are given by  $E_0 = |\beta|I$  and  $E_1 = |\alpha|U$ .

**Exercise 7.5** Let  $A$ ,  $B$ ,  $C$ , and  $D$  denote the quantum gates of Fig. 7.10, from left to right. We have

$$A = \begin{bmatrix} C & 0 & -S & 0 \\ 0 & C & 0 & -S \\ S & 0 & C & 0 \\ 0 & S & 0 & C \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad (\text{B.244})$$

$$C = \begin{bmatrix} C & 0 & S & 0 \\ 0 & C & 0 & S \\ -S & 0 & C & 0 \\ 0 & -S & 0 & C \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

where  $C \equiv \cos \frac{\theta}{2}$  and  $S \equiv \sin \frac{\theta}{2}$ . The action of the circuit in Fig. 7.10 is described by the unitary operator

$$U = DCBA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2CS & 0 & C^2 - S^2 \\ 0 & C^2 - S^2 & 0 & -2CS \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (\text{B.245})$$

As discussed in the solution of exercise 7.3, the Kraus operators are read from the first two columns of the matrix  $U$ ; that is,

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & 2CS \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sin \theta \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 & C^2 - S^2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \cos \theta \\ 0 & 0 \end{bmatrix}. \quad (\text{B.246})$$

It is instructive to derive the transformation of the Bloch sphere. We have

$$\begin{aligned}\rho' &= \frac{1}{2} \begin{bmatrix} 1+z' & x'-iy' \\ x'+iy' & 1-z' \end{bmatrix} = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \sin \theta \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sin \theta \end{bmatrix} + \begin{bmatrix} 0 & \cos \theta \\ 0 & 0 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \cos \theta & 0 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1+\cos^2 \theta+z \sin^2 \theta & (x-iy) \sin \theta \\ (x+iy) \sin \theta & 1-(\cos^2 \theta+z \sin^2 \theta) \end{bmatrix}. \end{aligned} \quad (\text{B.247})$$

Therefore,

$$x' = x \sin \theta, \quad y' = y \sin \theta, \quad z' = \cos^2 \theta + z \sin^2 \theta. \quad (\text{B.248})$$

These equations show that the Bloch sphere  $x^2 + y^2 + z^2$  is deformed into the ellipsoid

$$\frac{x'^2 + y'^2}{\sin^2 \theta} + \frac{(z' - \cos^2 \theta)^2}{\sin^4 \theta}. \quad (\text{B.249})$$

This ellipsoid has  $z$  as symmetry axis and centre  $(0, 0, \cos^2 \theta)$ . A displacement of the centre of the Bloch sphere necessarily demands a deformation of the sphere, if we wish  $\rho'$  to still represent a density matrix. Note that Eq. (B.249) corresponds to the minimum deformation required to the Bloch sphere in order to displace its centre along the  $z$ -axis by  $\cos^2 \theta$ . Indeed, the ellipsoid (B.249) and the Bloch sphere have a higher order tangency (see Fig. B.5).

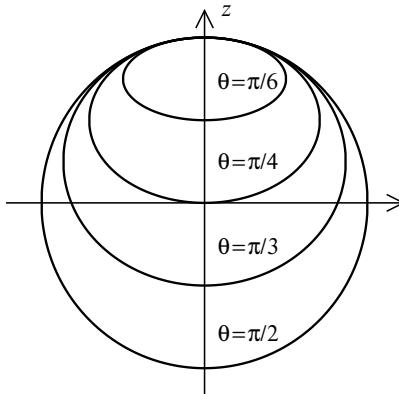


Fig. B.5 A visualization of the minimum deformation required to displace the centre of the Bloch sphere along the  $z$ -axis. The horizontal axis may be any axis in the  $(x, y)$  plane.

**Exercise 7.6** By direct computation we obtain

$$R_z \rho R_z^\dagger = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix} \begin{bmatrix} p & \alpha \\ \alpha^* & 1-p \end{bmatrix} \begin{bmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{bmatrix} = \begin{bmatrix} p & \alpha e^{-i\theta} \\ \alpha^* e^{i\theta} & 1-p \end{bmatrix}. \quad (\text{B.250})$$

It is now sufficient to use

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\theta} e^{-\frac{\theta^2}{2\sigma^2}} d\theta = e^{-\frac{\sigma^2}{2}}, \quad (\text{B.251})$$

with  $\sigma = \sqrt{2\lambda}$ , to obtain

$$\rho' = \begin{bmatrix} p & \alpha e^{-\lambda} \\ \alpha^* e^{-\lambda} & 1-p \end{bmatrix}. \quad (\text{B.252})$$

As  $p = \frac{1}{2}(1+z') = \frac{1}{2}(1+z)$  and  $\alpha e^{-\lambda} = \frac{1}{2}(x'-iy') = e^{-\lambda}\frac{1}{2}(x-iy)$ , Eq. (7.74) immediately follows.

**Exercise 7.7** Let  $\rho_1$  denote the density matrix describing the less significant qubit in Fig. 7.11 after the action of the two-qubit unitary transformation  $D$ . We have

$$\rho_1 = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger, \quad (\text{B.253})$$

where the Kraus operators  $E_0$  and  $E_1$  are read directly from the first two columns of  $D$ :

$$E_0 = \begin{bmatrix} C_0 & 0 \\ 0 & C_1 \end{bmatrix}, \quad E_1 = \begin{bmatrix} S_0 & 0 \\ 0 & S_1 \end{bmatrix}. \quad (\text{B.254})$$

Therefore, we obtain from (B.253) that

$$\begin{aligned} \rho_1 &= \frac{1}{2} \begin{bmatrix} 1+z_1 & x_1-iy_1 \\ x_1+iy_1 & 1-z_1 \end{bmatrix} \\ &= \begin{bmatrix} C_0 & 0 \\ 0 & C_1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} C_0 & 0 \\ 0 & C_1 \end{bmatrix} + \begin{bmatrix} S_0 & 0 \\ 0 & S_1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \begin{bmatrix} S_0 & 0 \\ 0 & S_1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1+z & (x-iy)(C_0C_1+S_0S_1) \\ (x+iy)(C_0C_1+S_0S_1) & 1-z \end{bmatrix}. \end{aligned} \quad (\text{B.255})$$

Thus,

$$x_1 = (C_0C_1 + S_0S_1)x = \cos(\theta_0 - \theta_1)x, \quad y_1 = \cos(\theta_0 - \theta_1)y, \quad z_1 = z. \quad (\text{B.256})$$

Finally,  $\rho'$  is obtained from  $\rho_1$  after a rotation of the Bloch sphere through an angle  $\xi$  about the axis directed along the unit vector  $\mathbf{n}$  (see Sec. 3.4.1).

**Exercise 7.8** In exercise 7.5 we studied a one-parameter map of the Bloch sphere having the north pole as fixed point. Let us call  $U_1(\theta_1)$  the two-qubit unitary transformation that realizes such single-qubit map. If  $R$  denotes a Bloch sphere rotation that maps  $z \rightarrow -z$ , then  $U_2(\theta_2) \equiv (I \otimes R) U_1(\theta_1) (I \otimes R^\dagger)$  induces a one-parameter map in the Bloch-sphere coordinates such as the south pole is the fixed point. If we combine  $U_1(\theta_1)$  and  $U_2(\theta_2)$  we move both the north and the south pole. We then consider the transformations  $U_3(\theta_3) = (I \otimes P) U_1(\theta_3) (I \otimes P^\dagger)$  and  $U_4(\theta_4) = (I \otimes P) U_2(\theta_4) (I \otimes P^\dagger)$ , where  $P$  is the matrix rotating the  $z$ -axis of the Bloch sphere to the  $x$ -axis. Similarly, we consider two other transformations,  $U_5(\theta_5) = (I \otimes Q) U_1(\theta_5) (I \otimes Q^\dagger)$  and  $U_6(\theta_6) = (I \otimes Q) U_2(\theta_6) (I \otimes Q^\dagger)$ , where the matrix  $Q$  rotates the  $z$ -axis to the  $y$ -axis. We have thus generated a generic 6-parameter  $(\theta_1, \dots, \theta_6)$  map  $\rho \rightarrow \rho_E$  of the Bloch sphere into an ellipsoid whose centre is in general located away from the centre of the Bloch sphere but whose axes are parallel to the

axes of the Bloch sphere. If we add two generic three-parameter rotations  $W_1(\theta_7, \theta_8, \theta_9)$  and  $W_2(\theta_{10}, \theta_{11}, \theta_{12})$ , we obtain a generic 12-parameter affine shift of the Bloch sphere:  $\rho \rightarrow \rho' = W_2 \rho_E W_1$ .

**Exercise 7.9** As the unitary operator  $V$  acts only on the environmental qubit, it does not modify the system density matrix. Indeed, we have

$$\begin{aligned}\rho' &= \text{Tr}_{\text{env}}[(V \otimes I)U(|0\rangle\langle 0| \otimes \rho)U^\dagger(V^\dagger \otimes I)] \\ &= \text{Tr}_{\text{env}}[U(|0\rangle\langle 0| \otimes \rho)U^\dagger(V^\dagger \otimes I)(V \otimes I)] \\ &= \text{Tr}_{\text{env}}[U(|0\rangle\langle 0| \otimes \rho)U^\dagger].\end{aligned}\quad (\text{B.257})$$

**Exercise 7.10** To solve this exercise, it is useful to remember that any  $2 \times 2$  unitary matrix  $U$  can be seen (up to an overall phase factor) as a rotation through an angle  $\delta$  about some axis of the Bloch sphere (see Sec. 3.4.1). Hence, we have

$$U = \cos \frac{\delta}{2} I - i \sin \frac{\delta}{2} (\mathbf{n} \cdot \boldsymbol{\sigma}), \quad (\text{B.258})$$

where  $\mathbf{n}$  is the unit vector directed along the rotation axis.

(i) For  $U = \sigma_x$ ,  $\delta = -\pi$  and  $\mathbf{n} = (1, 0, 0)$ . Thus,  $\sigma_x$  maps a vector  $\mathbf{r} = (x, y, z)$  of the Bloch sphere into  $\mathbf{r}_1 = (x_1, y_1, z_1) = (x, -y, -z)$ . As discussed in exercise 7.5, the next four quantum gates map  $\mathbf{r}_1$  into  $\mathbf{r}_2 = (x_2, y_2, z_2) = (x_1 \sin \theta, y_1 \sin \theta, \cos^2 \theta + z_1 \sin^2 \theta)$ . Finally,  $U^\dagger$  maps  $\mathbf{r}_2$  into  $\mathbf{r}' = (x', y', z') = (x_2, -y_2, -z_2)$ . The composition of the above three unitary transformations leads to  $\mathbf{r}' = (x \sin \theta, y \sin \theta, -\cos^2 \theta + z \sin^2 \theta)$ . The fixed point of the transformation  $\mathbf{r} \rightarrow \mathbf{r}'$  is the south pole of the Bloch sphere; that is,  $\mathbf{r} = (0, 0, -1)$ . Note that, if the transformations  $U$  and  $U^\dagger$  had not been applied, the fixed point would have been the north pole of the Bloch sphere,  $\mathbf{r} = (0, 0, 1)$ .

(ii) We can see from (B.258) that  $U = \frac{1}{\sqrt{2}}(I \pm i\sigma_j)$  induces a rotation through an angle  $\mp\pi/2$  about the  $j$ -axis of the Bloch sphere. Let us consider the three cases separately.

- a) For  $U = \frac{1}{\sqrt{2}}(I \pm i\sigma_x)$  we have  $\mathbf{r}_1 = (x, \pm z, \mp y)$ ,  $\mathbf{r}_2 = (x_1 \sin \theta, y_1 \sin \theta, \cos^2 \theta + z_1 \sin^2 \theta)$ ,  $\mathbf{r}' = (x_2, \mp z_2, \pm y_2)$ . Therefore,  $\mathbf{r}' = (x \sin \theta, \mp \cos^2 \theta + y \sin^2 \theta, z \sin \theta)$ .
- b) For  $U = \frac{1}{\sqrt{2}}(I \pm i\sigma_y)$  we have  $\mathbf{r}_1 = (\mp z, y, \pm x)$ ,  $\mathbf{r}_2 = (x_1 \sin \theta, y_1 \sin \theta, \cos^2 \theta + z_1 \sin^2 \theta)$ ,  $\mathbf{r}' = (\pm z_2, y_2, \mp x_2)$ . Therefore,  $\mathbf{r}' = (\pm \cos^2 \theta + x \sin^2 \theta, y \sin \theta, z \sin \theta)$ .
- c) For  $U = \frac{1}{\sqrt{2}}(I \pm i\sigma_z)$  we have  $\mathbf{r}_1 = (\pm y, \mp x, z)$ ,  $\mathbf{r}_2 = (x_1 \sin \theta, y_1 \sin \theta, \cos^2 \theta + z_1 \sin^2 \theta)$ ,  $\mathbf{r}' = (\mp y_2, \pm x_2, z_2)$ . Therefore,  $\mathbf{r}' = (x \sin \theta, y \sin \theta, \cos^2 \theta + z \sin^2 \theta)$ .

**Exercise 7.11** Let us first discuss the teleportation protocol (see Fig. B.6). The state  $|\psi\rangle$  to be teleported and the imperfect Bell states  $\rho_{\text{Bell}}^{(\text{imp})}$  are given by  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  and

$$\rho_{\text{Bell}}^{(\text{imp})} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & C & 0 \\ 0 & C & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (\text{B.259})$$

where  $C = \cos \theta$ . The matrix representation of the Bell measurement  $B$  is given by Eq. (5.48). We have

$$\rho_B = (B \otimes I) \rho_A (B^\dagger \otimes I), \quad (\text{B.260})$$

where

$$\rho_A = |\psi\rangle\langle\psi| \otimes \rho_{\text{Bell}}^{(\text{imp})} \quad (\text{B.261})$$

and

$$B \otimes I = \frac{1}{\sqrt{2}} \begin{bmatrix} I & 0 & 0 & I \\ 0 & I & I & 0 \\ I & 0 & 0 & -I \\ 0 & I & -I & 0 \end{bmatrix}. \quad (\text{B.262})$$

We obtain

$$\rho_B = \frac{1}{4} \begin{bmatrix} D & .. & .. & .. \\ .. & E & .. & .. \\ .. & .. & F & .. \\ .. & .. & .. & G \end{bmatrix}, \quad (\text{B.263})$$

where

$$\begin{aligned} D &= \begin{bmatrix} |\beta|^2 & \alpha^* \beta C \\ \alpha \beta^* C & |\alpha|^2 \end{bmatrix}, & E &= \begin{bmatrix} |\alpha|^2 & \alpha \beta^* C \\ \alpha^* \beta C & |\beta|^2 \end{bmatrix}, \\ F &= \begin{bmatrix} |\beta|^2 & -\alpha^* \beta C \\ -\alpha \beta^* C & |\alpha|^2 \end{bmatrix}, & G &= \begin{bmatrix} |\alpha|^2 & -\alpha \beta^* C \\ -\alpha^* \beta C & |\beta|^2 \end{bmatrix}, \end{aligned} \quad (\text{B.264})$$

and we have denoted by .. the  $2 \times 2$  matrix blocks whose expressions are not needed in our subsequent calculations. The outcome of the measurement performed on the first two qubits (by means of the detectors  $D_0$  and  $D_1$ ) determines the state of the third qubit ( $D$  if the outcome is 00,  $E$  if the outcome is 01,  $F$  if the outcome is 10, and  $G$  if the outcome is 11). As discussed in Sec. 5.5, in all cases the unitary operator  $U$  recovers the state  $\rho_f = E$ . The teleportation fidelity is

$$f = \langle \psi | \rho_f | \psi \rangle = |\alpha|^4 + |\beta|^4 + 2C|\alpha|^2|\beta|^2 = 1 - 2(1-C)(|\alpha|^2 - |\alpha|^4). \quad (\text{B.265})$$

Note that teleportation is perfect only for  $C = 1$ , otherwise  $f < 1$ .

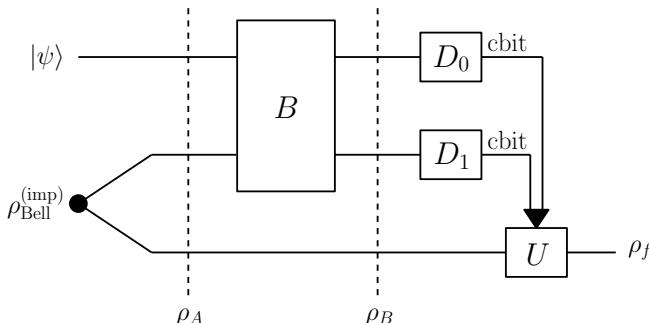


Fig. B.6 A schematic drawing of the teleportation protocol with an imperfect Bell state.

Let us now discuss the dense-coding protocol (see Fig. B.7). As we saw in Sec. 5.4, Alice applies the unitary transformation  $U \in \{I, \sigma_x, \sigma_y, \sigma_z\}$  to her half of the (imperfect)

Bell state. We have

$$\rho_A = (I \otimes U)\rho_{\text{Bell}}^{(\text{imp})}(I \otimes U^\dagger) = \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}^{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & C & 0 \\ 0 & C & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U^\dagger & 0 \\ 0 & U^\dagger \end{bmatrix}. \quad (\text{B.266})$$

The final density matrix is then given by  $\rho_{\text{fin}} = B\rho_A B^\dagger$ , where  $B$  is defined in Eq. (5.48). If Alice applies  $U = I$  we obtain

$$\frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1+C & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-C \end{bmatrix}. \quad (\text{B.267})$$

If instead  $U = \sigma_x$ , then

$$\frac{1}{2} \begin{bmatrix} 1+C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1-C & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (\text{B.268})$$

If  $U = \sigma_y$ , then

$$\frac{1}{2} \begin{bmatrix} 1-C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1+C & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (\text{B.269})$$

Finally, if Alice applies  $U = \sigma_z$ , we obtain

$$\frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1-C & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+C \end{bmatrix}. \quad (\text{B.270})$$

In all cases, Bob correctly recovers the two classical bits transmitted by Alice with probability  $p = \frac{1}{2}(1 + C)$ . Only when  $C = 1$  the error probability is zero; that is,  $p = 1$ .

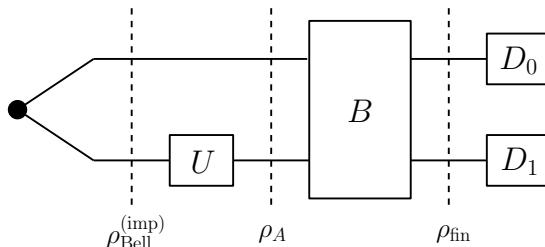


Fig. B.7 A schematic drawing of the dense-coding protocol with an imperfect Bell state.

**Exercise 7.12** We can read the relation between the Bloch coordinates and the elements of the density matrix for a qubit from Eq. (3.11). We then obtain

$$\mathcal{F} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathcal{F}^{-1} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & -i & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}. \quad (\text{B.271})$$

For  $n = 2$  qubits we have

$$\mathcal{F} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i & 0 & 0 & -i & 0 & 0 & i & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & i & 0 & 0 & i & 0 & 0 & -i & 0 & 0 & -i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 & 0 & 0 & 0 & -i & -i & 0 & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & i & 0 & 0 & 0 & 0 & i & -i & 0 & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & , \\ 0 & i & 0 & 0 & -i & 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 & 0 & i & 0 & , \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & , \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & , \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & , \\ 0 & i & 0 & 0 & -i & 0 & 0 & 0 & 0 & 0 & i & 0 & 0 & 0 & -i & 0 & , \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & , \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & \end{bmatrix}, \quad (\text{B.272})$$

$$\mathcal{F}^{-1} = \frac{1}{4} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -i & 0 & 0 & 1 & -i & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & -i & -i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -i & 0 & 0 & -i & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & i & 0 & 0 & 1 & i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\ 1 & i & 0 & 0 & -i & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & i & -i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & i & i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -i & 0 & 0 & i & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & i & 0 & 0 & 1 & -i & 0 & 0 \\ 1 & i & 0 & 0 & i & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & -i & i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -i & 0 & 0 & 1 & i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix}. \quad (\text{B.273})$$

**Exercise 7.13** We obtain

$$\chi_F = \begin{bmatrix} \sqrt{1-p} & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1-p & p \end{bmatrix}, \quad (\text{B.274})$$

corresponding to the deformation of the Bloch ball into an ellipsoid, with its center displaced along the  $z$ -axis.

**Exercise 7.14** We obtain

$$\mathcal{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad (\text{B.275})$$

$$\mathcal{R}^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 \\ -2 & -2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & -2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & -2 & -2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (\text{B.276})$$

Note that the 16 columns of  $\mathcal{R}$  correspond, from left to right, to the states  $|0\rangle \otimes |0\rangle$ ,  $|0\rangle \otimes |1\rangle$ ,  $\dots$ ,  $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ,  $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ , that is, states are ordered with the first qubit being the most significant one.

**Exercise 7.15** The quantum process matrix for the uncorrelated dephasing channel is given by

$$\chi_F = \begin{bmatrix} g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & g^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{B.277})$$

where  $g \equiv e^{-\lambda}$ . For the correlated dephasing channel instead we obtain

$$\chi_F = \begin{bmatrix} h & 0 & 0 & 0 & 0 & k & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & -k & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -k & 0 & 0 & h & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ k & 0 & 0 & 0 & 0 & h & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{B.278})$$

where  $h \equiv \frac{1}{2}(1 + g^4)$ ,  $k \equiv \frac{1}{2}(1 - g^4)$ .

**Exercise 7.16** It is convenient to write the master equation (7.170) in the  $\{|g\rangle, |e\rangle\}$  basis. If we write the density matrix  $\rho$  as

$$\rho = \begin{bmatrix} \rho_{gg} & \rho_{ge} \\ \rho_{eg} & \rho_{ee} \end{bmatrix}, \quad (\text{B.279})$$

we obtain

$$-\frac{i}{\hbar} [H, \rho] = -i\omega_0 \begin{bmatrix} 0 & \rho_{ge} \\ -\rho_{eg} & 0 \end{bmatrix}, \quad (\text{B.280})$$

$$\sigma_- \rho \sigma_+ - \frac{1}{2} \sigma_+ \sigma_- \rho - \frac{1}{2} \rho \sigma_+ \sigma_- = \frac{1}{2} \begin{bmatrix} 2\rho_{ee} & -\rho_{ge} \\ -\rho_{eg} & -2\rho_{ee} \end{bmatrix}, \quad (\text{B.281})$$

$$\sigma_+ \rho \sigma_- - \frac{1}{2} \sigma_- \sigma_+ \rho - \frac{1}{2} \rho \sigma_- \sigma_+ = \frac{1}{2} \begin{bmatrix} -2\rho_{gg} & -\rho_{ge} \\ -\rho_{eg} & 2\rho_{gg} \end{bmatrix}. \quad (\text{B.282})$$

After insertion of (B.280)–(B.281) into (7.170) we obtain the following equations:

$$\begin{aligned} \dot{\rho}_{gg} &= \gamma(\bar{n} + 1)\rho_{ee} - \gamma\bar{n}\rho_{gg}, \\ \dot{\rho}_{ee} &= -\gamma(\bar{n} + 1)\rho_{ee} + \gamma\bar{n}\rho_{gg}, \\ \dot{\rho}_{ge} &= -[\frac{1}{2}\gamma(2\bar{n} + 1) - i\omega_0]\rho_{ge}, \\ \dot{\rho}_{eg} &= -[\frac{1}{2}\gamma(2\bar{n} + 1) + i\omega_0]\rho_{eg}. \end{aligned} \quad (\text{B.283})$$

Their solution is

$$\begin{aligned} \rho_{gg}(t) &= B_1 - B_2 \exp[-\gamma(2\bar{n} + 1)t], \\ \rho_{ee}(t) &= \frac{\bar{n}}{\bar{n} + 1} B_1 + B_2 \exp[-\gamma(2\bar{n} + 1)t], \\ \rho_{ge}(t) &= B_3 \exp\left\{-[\frac{1}{2}\gamma(2\bar{n} + 1) - i\omega_0]t\right\}, \\ \rho_{eg}(t) &= B_4 \exp\left\{-[\frac{1}{2}\gamma(2\bar{n} + 1) + i\omega_0]t\right\}. \end{aligned} \quad (\text{B.284})$$

From the initial conditions and from the relation  $\text{Tr } \rho(0) = \rho_{gg}(0) + \rho_{ee}(0) = 1$  we obtain

$$\begin{aligned} B_1 &= \frac{\bar{n} + 1}{2\bar{n} + 1}, & B_2 &= \rho_{ee}(0) - \frac{\bar{n}}{2\bar{n} + 1}, \\ B_3 &= \rho_{ge}(0), & B_4 &= \rho_{eg}(0). \end{aligned} \quad (\text{B.285})$$

It is clear from Eq. (B.284) that the asymptotic density matrix is given by

$$\rho_s \equiv \lim_{t \rightarrow \infty} \rho(t) = \begin{bmatrix} \frac{\bar{n} + 1}{2\bar{n} + 1} & 0 \\ 0 & \frac{\bar{n}}{2\bar{n} + 1} \end{bmatrix}. \quad (\text{B.286})$$

We note that the diagonal terms approach equilibrium on a time scale  $\tau_d = [\gamma(2\bar{n} + 1)]^{-1}$  while the off-diagonal terms require a time scale  $\tau_{nd} = 2\tau_d$ . The stationary density matrix  $\rho_s$  corresponds to the Bloch vector  $r_s = (0, 0, \frac{1}{2\bar{n} + 1})$ .

**Exercise 7.17** We expand the commutators in Eq. (7.172) obtaining

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho] + \frac{1}{2} \sum_{i,j=1}^{N^2-1} A_{ij} \{2\sigma_i \rho \sigma_j^\dagger - \rho \sigma_j^\dagger \sigma_i - \sigma_j^\dagger \sigma_i \rho\}. \quad (\text{B.287})$$

After substitution of Eq. (7.175), namely of the expression

$$A_{ij} = \sum_{k=1}^{N^2-1} S_{ki}^* \tilde{A}_{kk} S_{kj}, \quad (\text{B.288})$$

into (B.287), we have

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho] + \sum_{i,j,k=1}^{N^2-1} S_{ki}^* \tilde{A}_{kk} S_{kj} \left\{ \sigma_i \rho \sigma_j^\dagger - \frac{1}{2} \rho \sigma_j^\dagger \sigma_i - \frac{1}{2} \sigma_j^\dagger \sigma_i \rho \right\}. \quad (\text{B.289})$$

Finally, we define

$$L_k = \sqrt{\tilde{A}_{kk}} \sum_i S_{ki}^* \sigma_i, \quad (\text{B.290})$$

which implies that

$$L_k^\dagger = \sqrt{\tilde{A}_{kk}} \sum_j S_{kj} \sigma_j^\dagger. \quad (\text{B.291})$$

Substitution of (B.290) and (B.291) into (B.289) finally leads to the GKLS master equation (7.169).

**Exercise 7.18** Equation (7.187) can be written as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & -\omega_0 & \Delta' \\ \omega_0 & 0 & -\Delta \\ -\Delta' & \Delta & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (\text{B.292})$$

This matrix is antisymmetric and therefore corresponds to a rotation. Indeed, we can write

$$\dot{\mathbf{r}} = \boldsymbol{\Omega} \times \mathbf{r}, \quad (\text{B.293})$$

with

$$\boldsymbol{\Omega} = (\Delta, \Delta', \omega_0). \quad (\text{B.294})$$

Therefore, we immediately obtain the rotation axis (7.188) and the rotation frequency  $\Omega = \sqrt{\Delta^2 + \Delta'^2 + \omega_0^2}$ .

**Exercise 7.19** Using the Bloch-Fano representation, we write two generic single-qubit states as

$$\rho_1 = \frac{1}{2}(I + \mathbf{r}_1 \cdot \boldsymbol{\sigma}), \quad \rho_2 = \frac{1}{2}(I + \mathbf{r}_2 \cdot \boldsymbol{\sigma}). \quad (\text{B.295})$$

The trace distance between  $\rho_1$  and  $\rho_2$  is then easily computed:

$$D(\rho_1, \rho_2) = \frac{1}{2} \text{Tr} |\rho_1 - \rho_2| = \frac{1}{4} \text{Tr} |(\mathbf{r}_1 - \mathbf{r}_2) \cdot \boldsymbol{\sigma}| = \frac{1}{4} |\mathbf{r}_1 - \mathbf{r}_2|. \quad (\text{B.296})$$

That is, the trace distance between two single-qubit states is equal to half of the Euclidean distance between their Bloch vectors.

**Exercise 7.20** The state of the system after having obtained  $n$  times the outcome 0 from the generalized measurement (7.240) is given by

$$|\psi^{(n)}\rangle = \alpha^{(n)}|0\rangle + \beta^{(n)}|1\rangle \approx \alpha \left(1 + \frac{1}{2}|\beta|^2\theta^2\right)^n |0\rangle + \beta \left(1 - \frac{1}{2}|\alpha|^2\theta^2\right)^n |1\rangle. \quad (\text{B.297})$$

We therefore obtain

$$\frac{\beta^{(n)}}{\alpha^{(n)}} = \frac{\left(1 - \frac{1}{2}|\alpha|^2\theta^2\right)^n \beta}{\left(1 + \frac{1}{2}|\beta|^2\theta^2\right)^n \alpha} \approx \exp\left(-\frac{1}{2}n\theta^2\right) \frac{\beta}{\alpha} \quad (\text{B.298})$$

and, since  $|\alpha^{(n)}|^2 + |\beta^{(n)}|^2 = 1$ ,

$$|\alpha^{(n)}|^2 \approx \frac{|\alpha|^2}{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)}, \quad |\beta^{(n)}|^2 \approx \frac{|\beta|^2 \exp(-n\theta^2)}{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)}. \quad (\text{B.299})$$

The probability of obtaining outcome 0 at the  $n$ -th weak measurement, provided the same outcome 0 was obtained in all previous measurements, is

$$p_0^{(n)} = |\alpha^{(n)}|^2 + |\beta^{(n)}|^2 \cos^2 \theta \approx 1 - |\beta^{(n)}|^2 \theta^2. \quad (\text{B.300})$$

The probability that the outcome 0 is always obtained is

$$\begin{aligned} p_0 &= \prod_{n=0}^{+\infty} p_0^{(n)} \approx \prod_{n=0}^{+\infty} \left( 1 - |\beta^{(n)}|^2 \theta^2 \right) \approx \prod_{n=0}^{+\infty} \left( 1 - \frac{|\beta|^2 \exp(-n\theta^2)}{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)} \theta^2 \right) \\ &= \prod_{n=0}^{+\infty} \left( \frac{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)(1-\theta^2)}{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)} \right) \approx \prod_{n=0}^{+\infty} \frac{|\alpha|^2 + |\beta|^2 \exp[-(n+1)\theta^2]}{|\alpha|^2 + |\beta|^2 \exp(-n\theta^2)} \\ &= \left( \frac{|\alpha|^2 + |\beta|^2 \exp(-\theta^2)}{|\alpha|^2 + |\beta|^2} \right) \left( \frac{|\alpha|^2 + |\beta|^2 \exp(-2\theta^2)}{|\alpha|^2 + |\beta|^2 \exp(-\theta^2)} \right) \dots \\ &= \lim_{n \rightarrow +\infty} \frac{|\alpha|^2 + |\beta|^2 \exp[-(n+1)\theta^2]}{|\alpha|^2 + |\beta|^2} = |\alpha|^2. \end{aligned} \quad (\text{B.301})$$

**Exercise 7.21** We just show in Fig. B.8 an example for  $\mathcal{N} = 100$  trajectories. One can see that this number is sufficient to obtain rather smooth paths for the Bloch vector. The reader can compare with the exact analytical solution of the master equation (7.253), see exercise 7.16.

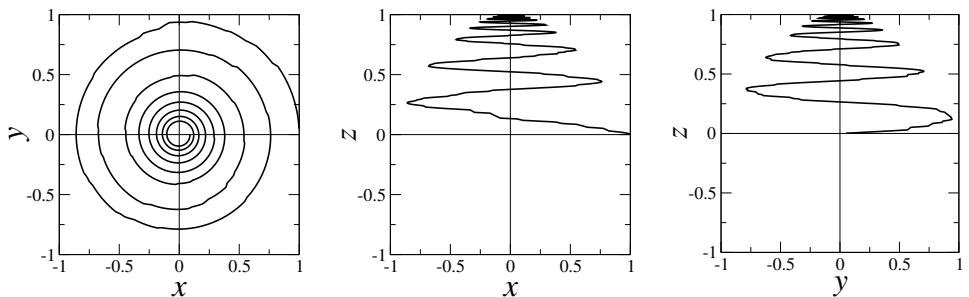


Fig. B.8 Projections of the Bloch-vector path, for a two-level atom whose dynamics is ruled by the master equation (7.253), with initial condition  $|\phi(t_0)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle)$ ,  $\gamma = 0.2$ ,  $\omega_0 = 2$  (we set  $\hbar = 1$ ). Time evolution is followed up to  $t = 8\pi$ , with a time step for quantum trajectories equal to  $dt = \pi/125$ .

## B.8 Chapter 8

### Exercise 8.1

$$\begin{aligned}
\log \left( \frac{n!}{\prod_{i=1}^k (np_i)!} \right) &= \log n! - \sum_{i=1}^k \log(np_i)! \\
&= n \log n - \frac{n}{\ln 2} - \sum_{i=1}^k \left[ np_i \log(np_i) - \frac{np_i}{\ln 2} \right] \\
&= -n \sum_{i=1}^k p_i \log p_i = nH(p_1, \dots, p_k). \tag{B.302}
\end{aligned}$$

Note that, in order to apply Stirling's formula, we have assumed that  $np_i \gg 1$ , for all  $i$ .

**Exercise 8.2** A generic two-letter input probability distribution is given by  $p(X=0) = q$ ,  $p(X=1) = 1-q$ , with  $0 \leq q \leq 1$ . The input entropy is  $H(X) = -q \log q - (1-q) \log(1-q) = H_{\text{bin}}(q)$ . The output probabilities are computed as  $p(Y=j) = \sum_{i=0}^1 p(X=i)p(Y=j|X=i)$ , with  $j = 1, 2$ , and therefore  $p(Y=0) = q+p-2pq$ ,  $p(Y=1) = 1-p(Y=0)$  and the output entropy  $H(Y) = H_{\text{bin}}(q+p-2qp)$ . We then compute the joint probabilities from  $p(X=i, Y=j) = p(X=i)p(Y=j|X=i)$ :  $p(0,0) = q(1-p)$ ,  $p(0,1) = qp$ ,  $p(1,0) = (1-q)p$ ,  $p(1,1) = (1-q)(1-p)$ . Hence the conditional entropy is given by  $H(Y|X) = -\sum_{x,y} p(x,y) \log p(y|x) = H_{\text{bin}}(p)$ . Finally we derive the mutual information

$$\mathcal{I}(X:Y) = H(Y) - H(Y|X) = H_{\text{bin}}(q+p-2qp) - H_{\text{bin}}(p). \tag{B.303}$$

The channel capacity  $C$  corresponds to the maximum of  $I$  over all possible input probability distributions  $\{p(x)\}$ . In this case, we must simply find the maximum of  $I$  over  $q$ , with  $q \in [0, 1]$ . The maximum is obtained for  $q = \frac{1}{2}$ , and correspondingly  $C = 1 - H_{\text{bin}}(p)$ .

**Exercise 8.3** The solutions follows the same steps as in the solution to exercise 8.2. We consider a generic two-letter input probability distribution given by  $p(X=0) = q$ ,  $p(X=1) = 1-q$ , with  $0 \leq q \leq 1$ . We then derive  $p(Y=0) = q(1-p)$ ,  $p(Y=1) = (1-q)(1-p)$ ,  $p(Y=e) = p$ , leading to  $H(Y) = H_{\text{bin}}(p) + (1-p)H_{\text{bin}}(q)$ . We then obtain  $X(Y|X) = H_{\text{bin}}(p)$ , and finally  $\mathcal{I}(X:Y) = (1-p)H_{\text{bin}}(q)$ . The mutual information is maximal for  $q = \frac{1}{2}$ , yielding the channel capacity  $C = 1-p$ . We note that the capacity of the erasure channel is greater than the capacity of the binary symmetric channel (for equal values of  $p$  for the two channels) and this result is intuitive since misidentifications are excluded in the erasure channel.

**Exercise 8.4** This quantum channel does not change the states sent by Alice, and therefore we recover the special case  $\theta = 0$  of the example discussed in Sec. 8.3.2.

**Exercise 8.5** The Bloch vectors  $\mathbf{r}_0 = (0, 0, 1)$  and  $\mathbf{r}_1 = (0, 0, -1)$ , corresponding to the states  $|0\rangle$  and  $|1\rangle$  sent by Alice, are modified by the quantum channel as follows:

$$\mathbf{r}_0 \rightarrow \tilde{\mathbf{r}}_0 = \mathbf{r}_0, \quad \mathbf{r}_1 \rightarrow \tilde{\mathbf{r}}_1 = (0, 0, \sin^2 \theta - \cos^2 \theta). \tag{B.304}$$

Let  $\tilde{\rho}_0$  and  $\tilde{\rho}_1$  denote the density matrices corresponding to the Bloch vectors  $\tilde{\mathbf{r}}_0$  and  $\tilde{\mathbf{r}}_1$ , respectively. Notice that for  $\theta = 0$  this quantum channel reduces to the identity.

In order to compute the mutual information of Alice and Bob, we first need the conditional probabilities

$$p(y|x) = \text{Tr}(\tilde{\rho}_x F_y) \quad (x, y = 0, 1, 2), \tag{B.305}$$

where the von Neumann projectors  $F_0$  and  $F_1$  are given by Eq. (8.38). If we choose the measurement axis  $\hat{n}$  in the  $(x, z)$  plane of the Bloch sphere (see Fig. 8.3); that is,  $\hat{n} = (\sin \bar{\theta}, 0, \cos \bar{\theta})$ , we obtain

$$\begin{aligned} p(0|0) &= \frac{1}{2}(1 + \cos \bar{\theta}), & p(1|0) &= \frac{1}{2}(1 - \cos \bar{\theta}), \\ p(0|1) &= \frac{1}{2}[1 - \cos 2\theta \cos \bar{\theta}], & p(1|1) &= \frac{1}{2}[1 + \cos 2\theta \cos \bar{\theta}]. \end{aligned} \quad (\text{B.306})$$

We now compute  $p(x, y) = p(x)p(y|x)$ . We know that the states  $|0\rangle$  and  $|1\rangle$  are sent by Alice with probabilities  $p(X = 0) = p$  and  $p(X = 1) = 1 - p$ , respectively. Therefore, we have

$$\begin{aligned} p(0, 0) &= \frac{1}{2}p(1 + \cos \bar{\theta}), & p(0, 1) &= \frac{1}{2}p(1 - \cos \bar{\theta}), \\ p(1, 0) &= \frac{1}{2}(1 - p)[1 - \cos 2\theta \cos \bar{\theta}], & p(1, 1) &= \frac{1}{2}(1 - p)[1 + \cos 2\theta \cos \bar{\theta}]. \end{aligned} \quad (\text{B.307})$$

Then we compute  $p(y) = \sum_x p(x, y)$  and obtain

$$\begin{aligned} p(Y = 0) &= \frac{1}{2}\{1 + \cos \bar{\theta}[p - (1 - p)\cos 2\theta]\}, \\ p(Y = 1) &= \frac{1}{2}\{1 - \cos \bar{\theta}[p - (1 - p)\cos 2\theta]\}. \end{aligned} \quad (\text{B.308})$$

Finally, we insert the expressions derived for  $p(x)$ ,  $p(y)$ , and  $p(x, y)$  into (6.33), obtaining the mutual information  $I(X:Y)$ .

A few examples of  $I(X:Y)$  as a function of the parameters  $p$ ,  $\theta$ , and  $\bar{\theta}$  are shown in Fig. B.9. As a general result, the mutual information is maximized for measurements along the  $z$ -axis; that is, when  $\bar{\theta} = 0, \pi$ . This result is demonstrated in Fig. B.9 (left) for fixed  $p$  and  $\theta$ . In Fig. B.9 (middle), we show that  $I = 0$  when  $\theta = \pi/2$ . This is quite natural since for  $\theta = \pi/2$  the quantum channel maps the entire Bloch sphere onto a single point, namely the north pole of the sphere. Therefore, whatever state Alice sends, Bob always receive the state  $|0\rangle$ . Hence, there is no transmission of information. Finally, in Fig. B.9 (right) we show the mutual information as a function of  $p$ , for given  $\theta$  and  $\bar{\theta}$ .

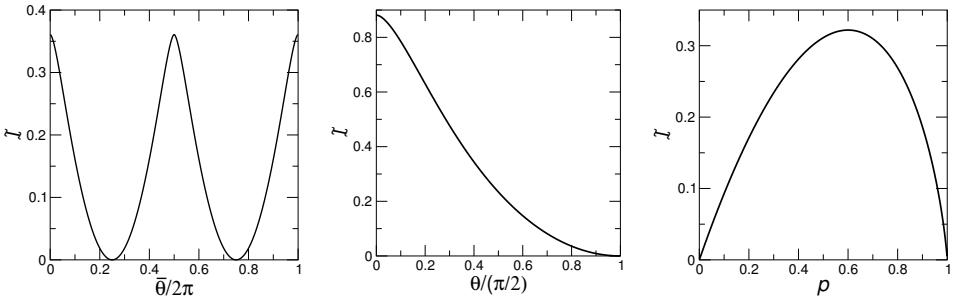


Fig. B.9 The mutual information  $I(X:Y)$  for a message transmitted as described in exercise 8.5, for  $p = 0.9$ ,  $\theta = \pi/8$  (left),  $p = 0.3$ ,  $\bar{\theta} = 0$  (middle), and  $\theta = \pi/4$ ,  $\bar{\theta} = 0$  (right).

**Exercise 8.6** Using an ensemble of pure states  $\{\rho_i\}$  and assuming a noiseless channel,  $\mathbb{S}(\rho_i) = \rho_i$ , we have

$$S\left[\mathbb{S}\left(\sum_i p_i \rho_i\right)\right] - \sum_i p_i S[\mathbb{S}(\rho_i)] = S\left(\sum_i p_i \rho_i\right) = S(\rho) \quad (\text{B.309})$$

and therefore  $C = \max_{\rho} S(\rho)$ . If qubits are sent down the channel, then it is sufficient to consider any ensemble of two orthogonal pure states with a priori probabilities  $p_1 = p_2 = \frac{1}{2}$  to obtain  $\rho = \frac{I}{2}$  and  $C = 1$ .

**Exercise 8.7** Assuming that the states  $\rho_1$  and  $\rho_2$  are chosen with a priori probabilities  $p_1$  and  $p_2 = 1 - p_1$ , we obtain

$$S[\mathbb{S}(\rho)] = S\left[\mathbb{S}\left(\sum_i p_i \rho_i\right)\right] = H_{\text{bin}}\left[\frac{p}{2} + (1-p)p_1\right], \quad (\text{B.310})$$

$$\sum_i p_i S[\mathbb{S}(\rho_i)] = p_1 H_{\text{bin}}\left(\frac{p}{2}\right) + (1-p_1) H_{\text{bin}}\left(\frac{p}{2}\right) = H_{\text{bin}}\left(\frac{p}{2}\right). \quad (\text{B.311})$$

Since  $\sum_i p_i S[\mathbb{S}(\rho_i)]$  is independent of  $p_1$ , it is sufficient to find the maximum of  $S[\mathbb{S}(\rho)]$  to obtain the Holevo information. The maximum is obtained for  $p_1 = \frac{1}{2}$  and correspondingly  $\chi(\mathbb{S}) = 1 - H_{\text{bin}}\left(\frac{p}{2}\right)$ . The capacity  $C(\mathbb{S}) = \chi(\mathbb{S})$ , shown in Fig. B.10, vanishes for  $p = 1$ , which corresponds to a useless channel, projecting any input state  $\rho_i$  onto the fully unpolarized state  $\frac{I}{2}$ . On the other hand,  $C = 1$  is obtained for the noiseless channel,  $p = 0$ .

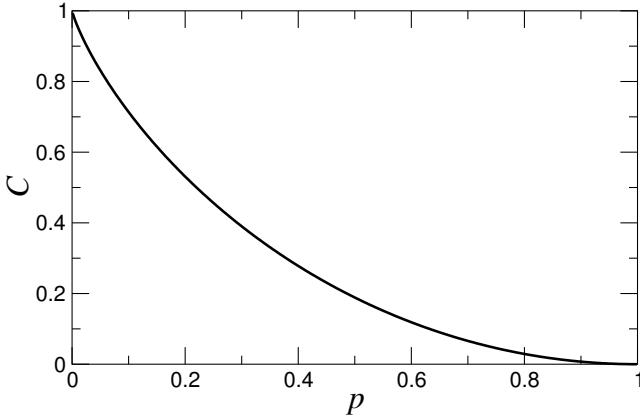


Fig. B.10 Classical capacity of the depolarizing channel.

**Exercise 8.8** We obtain

$$S[\mathbb{S}(\rho)] = S\left[\mathbb{S}\left(\sum_i p_i \rho_i\right)\right] = H_{\text{bin}}(p + p_1 - 2pp_1), \quad (\text{B.312})$$

$$\sum_i p_i S[\mathbb{S}(\rho_i)] = H_{\text{bin}}(p). \quad (\text{B.313})$$

The Holevo function  $S[\mathbb{S}(\rho)] - \sum_i p_i S[\mathbb{S}(\rho_i)]$  is maximized for  $p_1 = \frac{1}{2}$  and takes the value  $1 - H_{\text{bin}}(p) \equiv C_{\text{lb}}$ , which is a lower bound for the channel classical capacity  $C$ . This function has two maxima, for  $p = 0$  and  $p = 1$ , where  $C_{\text{lb}} = 1$ . The first case corresponds to the noiseless channel, the second to the deterministic bit-flip channel, namely all qubits are flipped by the channel and therefore there is no information degradation. The minimum  $C_{\text{lb}} = 0$  is obtained for  $p = \frac{1}{2}$ , where the qubits sent down the channel have probability

$\frac{1}{2}$  of being flipped. Hence the receiver's measurement is equivalent to a coin flipping experiments, and the initial information is lost.

Since the eigenstates  $|\pm\rangle$  of  $\sigma_x$  are not modified by the bit-flip channel, then it is sufficient to consider the ensemble  $\mathcal{E} = \{\rho_1 = |+\rangle\langle+|, \rho_2 = |-\rangle\langle-|; p_1 = \frac{1}{2}, p_2 = \frac{1}{2}\}$ , for which  $\chi(\mathcal{E}) = S(\epsilon) = 1$  to conclude that the channel capacity  $C = 1$ , namely the channel is noiseless with respect to the transmission of classical information.

**Exercise 8.9** The Kraus operators for the channel are given by

$$E_0 = \frac{1+g}{2} I, \quad E_z = \frac{1-g}{2} \sigma_z. \quad (\text{B.314})$$

Using Eq. (8.63), we then obtain

$$f_e(\rho_Q, \mathbb{S}) = \frac{1+g}{2} + \frac{1-g}{2} z^2, \quad (\text{B.315})$$

where  $z = \rho_{00} - \rho_{11}$  is the  $z$ -axis Bloch sphere coordinate for the input qubit. In particular, for a maximally entangled Bell pair  $\mathcal{RQ}$  we have  $\rho_Q = \frac{I}{2}$ , and therefore  $z = 0$ , leading to  $f_e = \frac{1+g}{2}$ .

**Exercise 8.10** It is convenient to compute matrix  $W$ , whose eigenvalues are  $\lambda_{1,2}^W = \frac{1}{2}(1 \pm \sqrt{g^2 + (1-g^2)z^2})$ , with  $z$  Bloch sphere coordinate for the input qubit. Hence

$$S_e(\rho_Q, \mathbb{S}) = S(W) = - \sum_{m=1}^2 \lambda_m^W \log \lambda_m^W. \quad (\text{B.316})$$

In particular,

$$S_e^{(z=0)} = H_{\text{bin}}\left(\frac{1+g}{2}\right) = H_{\text{bin}}(f_e^{(z=0)}). \quad (\text{B.317})$$

It is worth noticing that the entropy exchange takes its maximum for the completely dephasing channel ( $g = 0$ ), for which the entanglement fidelity  $f_e$  is minimum.

**Exercise 8.11** The coherent information is given by  $\mathcal{I}_c(\rho_Q, \mathbb{S}) = H_{\text{bin}}(\lambda_1^{\text{out}}) - H_{\text{bin}}(\lambda_1^W)$ , where the eigenvalues  $\lambda_{1,2}^W$  are provided in the solution to exercise 8.10 and  $\lambda_{1,2}^{\text{out}} = \frac{1}{2}(1 \pm \sqrt{z^2 + g^2\gamma^2})$  are the eigenvalues of  $\rho'_Q$ . Here  $\gamma^2 = x^2 + y^2$ , where the Bloch coordinates  $x = 2\text{Re}(\rho_{01})$  and  $y = -2\text{Im}(\rho_{01})$ . It is easy to show that the coherent information is maximized by the input state  $\rho_Q = \frac{1}{2}I$ , that is, for  $x = y = z = 0$ . In this case,

$$\mathcal{I}_c^{(z=\gamma=0)} = 1 - S_e^{(z=0)} = 1 - H_{\text{bin}}\left(\frac{1+g}{2}\right). \quad (\text{B.318})$$

## B.9 Chapter 9

**Exercise 9.1** The quantum circuit extracting the error syndrome without auxiliary qubits is shown in Fig. B.11. It can be readily checked that the output  $x_0 = 0, x_1 = 0$  corresponds to no error,  $x_0 = 1, x_1 = 1$  to error on the first qubit,  $x_0 = 1, x_1 = 0$  to error on the second qubit and  $x_0 = 0, x_1 = 1$  to error on the third qubit. In the case  $x_0 = x_1 = 1$  the logical state  $\alpha|0\rangle + \beta|1\rangle$  is recovered after a bit flip of the first qubit, in all the other cases the state of the first qubit is correct and no further action is required. At the end the three qubits are left in the state  $(\alpha|0\rangle + \beta|1\rangle)|x_0\rangle|x_1\rangle$ .

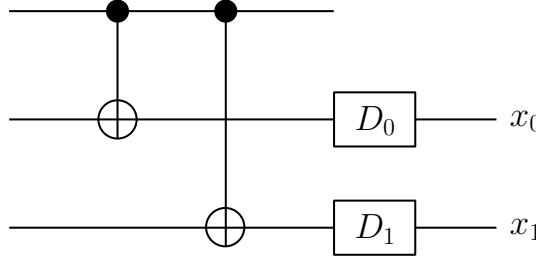


Fig. B.11 A quantum circuit extracting the error syndrome in the case of the three-qubit bit-flip errors, without using any ancillary qubits.

**Exercise 9.2** We start from a generic pure state,  $\alpha|0\rangle + \beta|1\rangle$ . After coding, the three-qubit state is  $\alpha|000\rangle + \beta|111\rangle$ . Let us assume that the error  $R_x(\delta)$  acts on the first qubit (the cases of the same error acting either on the second or on the third qubit can be treated similarly). The three-qubit state becomes

$$\alpha C|000\rangle - i\alpha S|100\rangle - i\beta S|011\rangle + \beta C|111\rangle, \quad (\text{B.319})$$

where  $C \equiv \cos \frac{\delta}{2}$  and  $S \equiv \sin \frac{\delta}{2}$ . After decoding (see Fig. B.11), we obtain

$$\alpha C|000\rangle - i\alpha S|111\rangle - i\beta S|011\rangle + \beta C|100\rangle, \quad (\text{B.320})$$

corresponding to the three-qubit density matrix

$$(\alpha C|000\rangle - i\alpha S|111\rangle - i\beta S|011\rangle + \beta C|100\rangle)(\alpha^* C|000\rangle + i\alpha^* S|111\rangle + i\beta^* S|011\rangle + \beta^* C|100\rangle). \quad (\text{B.321})$$

After tracing the third and third qubit we obtain the (mixed) state

$$\rho = \begin{bmatrix} |\alpha|^2 C^2 + |\beta|^2 S^2 & \alpha \beta^* C^2 + \alpha^* \beta S^2 \\ \alpha^* \beta C^2 + \alpha \beta^* S^2 & |\alpha|^2 S^2 + |\beta|^2 C^2 \end{bmatrix}. \quad (\text{B.322})$$

If we measure the second and the third qubit, we obtain 00 with probability  $C^2$  and 11 with probability  $S^2$ . The collapse postulate of quantum mechanics tells us that in the first case we obtain the state  $(\alpha|0\rangle + \beta|1\rangle)|00\rangle$ , in the latter  $(\alpha|1\rangle + \beta|0\rangle)|11\rangle$ . In the last case, it is sufficient to apply a NOT gate to the first qubit to correct the error.

**Exercise 9.3** If a single qubit, prepared in the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ , is sent through a bit-flip noisy channel, the final state of the qubit is described by

$$\rho = (1-\epsilon)(\alpha|0\rangle + \beta|1\rangle)(\alpha^* \langle 0| + \beta^* \langle 1|) + \epsilon(\alpha|1\rangle + \beta|0\rangle)(\alpha^* \langle 1| + \beta^* \langle 0|), \quad (\text{B.323})$$

where  $\epsilon$  is the bit-flip probability. Hence,

$$\rho = \begin{bmatrix} (1-\epsilon)|\alpha|^2 + \epsilon|\beta|^2 & (1-\epsilon)\alpha\beta^* + \epsilon\alpha^*\beta \\ (1-\epsilon)\alpha^*\beta + \epsilon\alpha\beta^* & (1-\epsilon)|\beta|^2 + \epsilon|\alpha|^2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1+z' & x'-iy' \\ x'+iy' & 1-z' \end{bmatrix}, \quad (\text{B.324})$$

where  $x' = x$ ,  $y' = (1-2\epsilon)y$ ,  $z' = (1-2\epsilon)z$  while  $(x, y, z)$  and  $(x', y', z')$  are the Bloch-sphere coordinates corresponding to the initial density matrix  $\rho_0 = |\psi\rangle\langle\psi|$  and to  $\rho$ , respectively.

Therefore, Eq. (7.49) for a bit-flip channel is recovered. The fidelity of the transmitted state is given by Eq. (5.19), namely,

$$\begin{aligned} f &= \langle \psi | \rho | \psi \rangle = \text{Tr}(\rho_0 \rho) = \text{Tr} \left( \frac{1}{2} \begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1+z' & x'-iy' \\ x'+iy' & 1-z' \end{bmatrix} \right) \\ &= \frac{1}{2} [1 + x^2 + (1 - 2\epsilon)(y^2 + z^2)] = 1 - \epsilon(y^2 + z^2). \end{aligned} \quad (\text{B.325})$$

Note that unit fidelity is recovered when  $\epsilon = 0$  or when the initial state is an eigenstate of  $\sigma_x$  ( $x = \pm 1$ ,  $y = z = 0$ ). For a generic initial state  $f = 1 - O(\epsilon)$ .

When the three qubit bit-flip code is applied the initial logical qubit is encoded into three physical qubits, which are sent through the bit-flip channel. After error correction and decoding the initial state is recovered, unless two or more qubits were flipped. Therefore, following Eq. (9.8) and the subsequent discussion, we can see that the final state of the logical qubit is given by

$$\begin{aligned} \rho &= [(1 - \epsilon)^3 + 3\epsilon(1 - \epsilon)^2] (\alpha|0\rangle + \beta|1\rangle)(\alpha^* \langle 0| + \beta^* \langle 1|) \\ &\quad + [3\epsilon^2(1 - \epsilon) + \epsilon^3] (\alpha|1\rangle + \beta|0\rangle)(\alpha^* \langle 1| + \beta^* \langle 0|). \end{aligned} \quad (\text{B.326})$$

Therefore, the fidelity of the transmitted state can be computed as in Eq. (B.325), after the substitution  $\epsilon \rightarrow \epsilon_c \equiv 3\epsilon^2 - 2\epsilon^3$ . We have  $f = 1 - \epsilon_c(y^2 + z^2) = 1 - (3\epsilon^2 - 2\epsilon^3)(y^2 + z^2)$ . For a generic state,  $f = 1 - O(\epsilon^2)$ . This has to be compared with  $f = 1 - O(\epsilon)$ , obtained by sending a single qubit without error correction. Therefore, the error correction procedure greatly improves the fidelity of the transmitted quantum information for  $\epsilon \ll 1$ .

**Exercise 9.4** Let us consider the state  $|0_L\rangle$  and an error affecting the first qubit (the state  $|1_L\rangle$ ) and errors affecting other qubits are treated in the same manner). We have

$$\begin{aligned} U|0_L\rangle|0\rangle_E &= \frac{1}{\sqrt{2}} (|0\rangle|e_0\rangle_E + |1\rangle|e_1\rangle_E)|00\rangle(\dots)(\dots) + \frac{1}{\sqrt{2}} (|0\rangle|e_2\rangle_E + |1\rangle|e_3\rangle_E)|11\rangle(\dots)(\dots) \\ &= \frac{1}{\sqrt{2}} (|000\rangle|e_0\rangle_E + |100\rangle|e_1\rangle_E + |011\rangle|e_2\rangle_E + |111\rangle|e_3\rangle_E)(\dots)(\dots), \end{aligned} \quad (\text{B.327})$$

where the state of the last six qubits has been simply denoted as  $(\dots)(\dots)$  since they remain de-entangled from the environment. The density matrix describing the first three qubits plus the environment reads

$$\begin{aligned} \rho &= \frac{1}{2} (|000\rangle|e_0\rangle_E + |100\rangle|e_1\rangle_E + |011\rangle|e_2\rangle_E + |111\rangle|e_3\rangle_E) \\ &\quad \times (\langle 000|_E \langle e_0| + \langle 100|_E \langle e_1| + \langle 011|_E \langle e_2| + \langle 111|_E \langle e_3|). \end{aligned} \quad (\text{B.328})$$

After tracing over the three qubits, we obtain the density matrix of the environment:

$$\rho_E = \frac{1}{2} (|e_0\rangle_{EE} \langle e_0| + |e_1\rangle_{EE} \langle e_1| + |e_2\rangle_{EE} \langle e_2| + |e_3\rangle_{EE} \langle e_3|). \quad (\text{B.329})$$

**Exercise 9.5** The correctable errors are

$$\begin{aligned} E_1 &= \sigma_1^x \otimes I_2 \otimes I_3 = E_1^\dagger, \\ E_2 &= I_1 \otimes \sigma_2^x \otimes I_3 = E_2^\dagger, \\ E_3 &= I_1 \otimes I_2 \otimes \sigma_3^x = E_3^\dagger. \end{aligned} \quad (\text{B.330})$$

Since  $\sigma_x^2 = I$ , we have  $E_1^\dagger E_1 = E_2^\dagger E_2 = E_3^\dagger E_3 = I$ . Thus, we obtain

$$\langle i_L | E_a^\dagger E_a | j_L \rangle = \delta_{ij}, \quad (\text{B.331})$$

where  $a = 1, 2, 3$ ,  $i = 0, 1$  ( $|0_L\rangle = |000\rangle$ ,  $|1_L\rangle = |111\rangle$ ). Finally, it is easy to check that, for  $a \neq b$ ,  $\langle i_L | E_a^\dagger E_b | j_L \rangle = 0$ . For instance, for  $a = 1$ ,  $b = 2$  we have

$$\langle 0_L | E_1^\dagger E_2 | 0_L \rangle = \langle 000 | \sigma_1^x \otimes \sigma_2^x \otimes I_3 | 000 \rangle = \langle 000 | 110 \rangle = 0. \quad (\text{B.332})$$

Similarly, we obtain  $\langle 1_L | E_1^\dagger E_2 | 1_L \rangle = 0$ . Thus, condition (9.29) is fulfilled with  $C_{ab} = \delta_{ab}$ .

**Exercise 9.6** The non-degeneracy of the three-qubit code was verified in the previous exercise (we have seen that  $C_{ab} = \delta_{ab}$ ). To show the degeneracy of the nine-qubit Shor code it is sufficient to find a case in which  $C_{ab} \neq \delta_{ab}$ . Let  $E_1$  and  $E_2$  denote the phase errors affecting the first and on the second qubit, respectively. We obtain

$$\begin{aligned} |0'_L\rangle &\equiv E_1|0_L\rangle = \frac{1}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \\ |1'_L\rangle &\equiv E_1|1_L\rangle = \frac{1}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \end{aligned} \quad (\text{B.333})$$

Similarly, we obtain

$$\begin{aligned} |0''_L\rangle &\equiv E_2|0_L\rangle = \frac{1}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle), \\ |1''_L\rangle &\equiv E_2|1_L\rangle = \frac{1}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \end{aligned} \quad (\text{B.334})$$

Since  $|0''_L\rangle = |0'_L\rangle$  and  $|1''_L\rangle = |1'_L\rangle$ , we have  $\langle i_L | E_1^\dagger E_2 | i_L \rangle = 1$ , with  $i = 0, 1$ . Thus,  $C_{12} = 1 \neq \delta_{12}$ .

**Exercise 9.7** Looking at the list of generators  $M_j \in \mathcal{S}$  in Table 9.1, we can observe that, for each of the nine qubits, there is at least one  $M_j$  which acts with  $\sigma_z$  and one which acts with  $\sigma_x$ . Then it is clear that any single-qubit operator  $\sigma_i^x$  (and also  $\sigma_i^y$ ) will anti-commute with the generator  $M_j$  containing  $\sigma_j^z$ . Conversely, any single-qubit operator  $\sigma_i^z$  will anti-commute with the generator  $M_j$  containing  $\sigma_j^x$ . This follows from the property (A.104) of the Pauli matrices.

**Exercise 9.8** It is not difficult to check that the operator  $O_{3w} = \sigma_1^x \sigma_2^x \sigma_3^x$  commutes with all the generators  $M_i$  in the stabilizer, that are listed in Table 9.1. Indeed, clearly  $O_{3w}$  commutes with operators that do not have support on the qubits 1, 2, and 3 (which are  $M_3$ ,  $M_4$ ,  $M_5$ ,  $M_6$ ), and with  $M_7$  and  $M_8$ , since  $[\sigma_j^x, \sigma_j^x] = 0$ . Moreover we have:

$$\begin{aligned} [O_{3w}, M_1] &= [\sigma_1^x \sigma_2^x, \sigma_1^z \sigma_2^z] \sigma_3^x = (\sigma_1^x [\sigma_2^x, \sigma_1^z] \sigma_2^z + \sigma_1^x \sigma_1^z [\sigma_2^x, \sigma_2^z] + [\sigma_1^x, \sigma_1^z] \sigma_2^z \sigma_2^x) \sigma_3^x \\ &= -(\sigma_1^x \sigma_1^z \sigma_2^y - \sigma_1^y \sigma_2^z \sigma_2^x) \sigma_3^x = (i\sigma_1^y \sigma_2^y - i\sigma_1^y \sigma_2^y) \sigma_3^x = 0, \end{aligned} \quad (\text{B.335})$$

where we exploited the fact that Pauli matrices on different qubits commute, we also used the property (A.100c) of the commutator, the commutation rules (A.105) involving Pauli matrices, and the relations (A.106). Proceeding in a similar way, we have:

$$\begin{aligned} [O_{3w}, M_2] &= [\sigma_1^x \sigma_3^x, \sigma_1^z \sigma_3^z] \sigma_2^x = (\sigma_1^x [\sigma_3^x, \sigma_1^z] \sigma_3^z + \sigma_1^x \sigma_1^z [\sigma_3^x, \sigma_3^z] + [\sigma_1^x, \sigma_1^z] \sigma_3^z \sigma_3^x) \sigma_2^x \\ &= -(\sigma_1^x \sigma_1^z \sigma_3^y - \sigma_1^y \sigma_3^z \sigma_3^x) \sigma_2^x = (i\sigma_1^y \sigma_3^y - i\sigma_1^y \sigma_3^y) \sigma_2^x = 0. \end{aligned} \quad (\text{B.336})$$

However, despite the fact that  $O_{3w}$  commutes with everything in  $\mathcal{S}$ , a simple inspection of the two codewords  $|0_L\rangle$  and  $|1_L\rangle$  in Eqs. (9.12) tells us that:

$$\langle 0_L | \sigma_1^x \sigma_2^x \sigma_3^x | 0_L \rangle = \langle 0_L | 0_L \rangle = +1 \quad (\text{B.337})$$

$$\langle 1_L | \sigma_1^x \sigma_2^x \sigma_3^x | 1_L \rangle = -\langle 1_L | 1_L \rangle = -1. \quad (\text{B.338})$$

Therefore, we can conclude that the weight-three operator  $O_{3w}$  does not satisfy Eq. (9.29).

**Exercise 9.9** If condition (1) holds, then, by definition of the stabilizer  $\mathcal{S}$ , one has that  $\langle \psi | E_a^\dagger E_b | \psi \rangle = \langle \psi | \psi \rangle = 1$ , and thus Eq. (9.41) is satisfied with  $C_{AB} = 1$ .

On the other hand, if condition (2) holds with  $M \in \mathcal{S}$  and  $ME_a^\dagger E_b = -E_a^\dagger E_b M$ , it is immediate to see that

$$\langle \psi | E_a^\dagger E_b | \psi \rangle = \langle \psi | E_a^\dagger E_b M | \psi \rangle = -\langle \psi | M E_a^\dagger E_b | \psi \rangle = -\langle \psi | E_a^\dagger E_b | \psi \rangle, \quad (\text{B.339})$$

and therefore  $\langle \psi | E_a^\dagger E_b | \psi \rangle = 0$ .

**Exercise 9.10** As an example, we consider the second line of Table 9.3. The effect of the error  $\sigma_3^x \sigma_3^z$  on the encoded states (9.44) leads to

$$\begin{aligned} \sigma_3^x \sigma_3^z |0_L\rangle &= \frac{1}{\sqrt{8}} (|00100\rangle + |01011\rangle - |10111\rangle - |11000\rangle \\ &\quad - |00010\rangle + |01101\rangle - |10001\rangle + |11110\rangle), \\ \sigma_3^x \sigma_3^z |1_L\rangle &= \frac{1}{\sqrt{8}} (-|11011\rangle - |10100\rangle - |01000\rangle - |00111\rangle \\ &\quad + |11101\rangle - |10010\rangle - |01110\rangle + |00001\rangle). \end{aligned} \quad (\text{B.340})$$

It can be checked by direct computation that the quantum circuit in Fig. 9.8 maps the error-affected state  $\sigma_3^x \sigma_3^z (\alpha |0_L\rangle + \beta |1_L\rangle)$  into  $-\alpha |11101\rangle + \beta |11001\rangle$ . Therefore, the detectors  $D_a$ ,  $D_b$ ,  $D_c$ , and  $D_d$  in Fig. 9.8 give outcome  $a = 1$ ,  $b = 1$ ,  $c = 0$ , and  $d = 1$  and the third qubit is in the state  $|\psi'\rangle = -\alpha |0\rangle + \beta |1\rangle$ , in agreement with Table 9.3.

**Exercise 9.11** It is sufficient to remember that an amplitude error in the computational basis  $\{|0\rangle, |1\rangle\}$  becomes a phase error in the basis  $\{|0\rangle_x, |1\rangle_x\}$ . Indeed, the mapping

$$|0\rangle_x \rightarrow |0\rangle_x, \quad |1\rangle_x \rightarrow e^{i\phi} |1\rangle_x \quad (\text{B.341})$$

may be expressed in the computational basis as

$$|0\rangle \rightarrow \frac{1+e^{i\phi}}{2} |0\rangle + \frac{1-e^{i\phi}}{2} |1\rangle, \quad |1\rangle \rightarrow \frac{1-e^{i\phi}}{2} |0\rangle + \frac{1+e^{i\phi}}{2} |1\rangle. \quad (\text{B.342})$$

For  $\phi = \pi$ , these latter relations reduce to the bit-flip error ( $|0\rangle \leftrightarrow |1\rangle$ ). We can code a single logical qubit by means of two physical qubits as follows:

$$\begin{aligned} |0_L\rangle &\equiv |0\rangle_x |1\rangle_x = \frac{1}{2} (|00\rangle - |01\rangle + |10\rangle - |11\rangle) \\ |1_L\rangle &\equiv |1\rangle_x |0\rangle_x = \frac{1}{2} (|00\rangle + |01\rangle - |10\rangle - |11\rangle). \end{aligned} \quad (\text{B.343})$$

If the same amplitude error (B.341) acts on both qubits, we have

$$|0_L\rangle \rightarrow e^{i\phi} |0_L\rangle, \quad |1_L\rangle \rightarrow e^{i\phi} |1_L\rangle, \quad (\text{B.344})$$

so that a generic state  $|\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$  just acquires an overall phase factor of no physical significance. The generalization to  $n$  qubits is analogous to the discussion in Sec. 9.7.

**Exercise 9.12** In the basis of the eigenstates of  $\sigma_z$  the Hamiltonian reads as follows:

$$H = \frac{1}{2}\hbar \begin{bmatrix} \omega & \Omega \\ \Omega & -\omega \end{bmatrix}. \quad (\text{B.345})$$

The corresponding Schrödinger equation can be solved as discussed in exercise 3.17. Alternatively, we can write the density matrix as  $\rho(t) = \frac{1}{2}(I + \mathbf{r}(t) \cdot \boldsymbol{\sigma})$  and obtain from the von Neumann equation (2.98) the following equations for components  $(x, y, z)$  of the Bloch vector  $\mathbf{r}$ :

$$\begin{cases} \dot{x} = -\omega y, \\ \dot{y} = \omega x - \Omega z, \\ \dot{z} = \Omega y. \end{cases} \quad (\text{B.346})$$

The solution to this linear system reads

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \\ \frac{\omega}{\Omega} \end{bmatrix} + B \begin{bmatrix} 1 \\ -i\frac{\sqrt{\omega^2 + \Omega^2}}{\omega} \\ -\frac{\Omega}{\omega} \end{bmatrix} e^{i\sqrt{\omega^2 + \Omega^2}t} + C \begin{bmatrix} 1 \\ i\frac{\sqrt{\omega^2 + \Omega^2}}{\omega} \\ -\frac{\Omega}{\omega} \end{bmatrix} e^{-i\sqrt{\omega^2 + \Omega^2}t}, \quad (\text{B.347})$$

with the constants  $A$ ,  $B$  and  $C$  determined by the initial condition. Starting from the state  $|0\rangle$  ( $x(0) = y(0) = 0$ ,  $z(0) = 1$ ) we obtain  $A = \frac{-\omega\Omega}{\omega^2 + \Omega^2}$  and  $B = C = -\frac{A}{2}$ , so that

$$\begin{cases} x(t) = \frac{\omega\Omega}{\omega^2 + \Omega^2} [1 - \cos(\sqrt{\omega^2 + \Omega^2}t)], \\ y(t) = -\frac{\Omega}{\sqrt{\omega^2 + \Omega^2}} \sin(\sqrt{\omega^2 + \Omega^2}t), \\ z(t) = \frac{\omega^2}{\omega^2 + \Omega^2} + \frac{\Omega^2}{\omega^2 + \Omega^2} \cos(\sqrt{\omega^2 + \Omega^2}t). \end{cases} \quad (\text{B.348})$$

The survival probability is given by

$$p(t) = \text{Tr}(\rho(t)|0\rangle\langle 0|) = \frac{1}{2}[1 + z(t)] = 1 - \frac{\Omega^2}{\omega^2 + \Omega^2} \sin^2\left(\sqrt{\omega^2 + \Omega^2} \frac{t}{2}\right). \quad (\text{B.349})$$

The short-time expansion of this equation leads to  $p(t) = 1 - \frac{t^2}{t_Z^2}$ , with the Zeno time

$$t_Z = \frac{2}{\Omega} = \frac{\hbar}{\sqrt{\langle 0 | H_{\text{int}}^2 | 0 \rangle}}. \quad (\text{B.350})$$

A graphical visualization of the Zeno effect is shown in Fig. B.12: the survival probability is enhanced by frequent measurements of  $\sigma_z$ , performed at time intervals  $\tau \ll t_Z$ . In such a case, the survival probability is bounded below (interpolated at  $t = N\tau$ ) by the curve

$$p(t) = \exp\left(-\frac{\tau}{t_Z^2} t\right). \quad (\text{B.351})$$

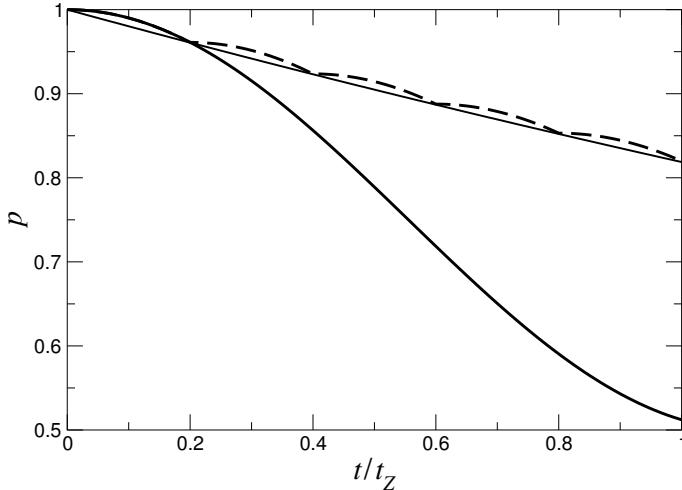


Fig. B.12 Evolution of the survival probability with (above) and without (below) measurements, with  $\omega = \Omega$  and  $\tau = \frac{1}{5}t_Z$ . The thin solid line shows the exponential decay (B.351).

**Exercise 9.13** The eigenvalues  $E_j$  and the corresponding eigenvectors  $|\varphi_j\rangle$  of Hamiltonian (9.90) read

$$E_0 = 0, \quad E_1 = \sqrt{\Omega_1^2 + \Omega_2^2}, \quad E_2 = -\sqrt{\Omega_1^2 + \Omega_2^2}, \quad (\text{B.352})$$

$$|\varphi_0\rangle = \begin{bmatrix} \frac{\Omega_2}{\sqrt{\Omega_1^2 + \Omega_2^2}} \\ 0 \\ -\frac{\Omega_1}{\sqrt{\Omega_1^2 + \Omega_2^2}} \end{bmatrix}, \quad |\varphi_1\rangle = \begin{bmatrix} \frac{\Omega_1}{\sqrt{2(\Omega_1^2 + \Omega_2^2)}} \\ \frac{1}{\sqrt{2}} \\ \frac{\Omega_2}{\sqrt{2(\Omega_1^2 + \Omega_2^2)}} \end{bmatrix}, \quad |\varphi_2\rangle = \begin{bmatrix} \frac{\Omega_1}{\sqrt{2(\Omega_1^2 + \Omega_2^2)}} \\ -\frac{1}{\sqrt{2}} \\ \frac{\Omega_2}{\sqrt{2(\Omega_1^2 + \Omega_2^2)}} \end{bmatrix}. \quad (\text{B.353})$$

The wave function  $|\psi(0)\rangle = |0\rangle$  may be expanded over the (orthonormal) basis  $\{|\varphi_0\rangle, |\varphi_1\rangle, |\varphi_2\rangle\}$ :

$$|\psi(0)\rangle = \sum_{j=0}^2 c_j |\varphi_j\rangle, \quad c_j = \langle \varphi_j | \psi(0) \rangle = \langle \varphi_j | 0 \rangle. \quad (\text{B.354})$$

We then obtain

$$|\psi(t)\rangle = \sum_{j=0}^2 c_j \exp\left(-\frac{i}{\hbar} E_j t\right) |\varphi_j\rangle. \quad (\text{B.355})$$

Therefore, the survival probability at time  $t$  reads

$$p(t) = |\langle 0 | \psi(t) \rangle|^2 = \left| \sum_{j=0}^2 \exp\left(-\frac{i}{\hbar} E_j t\right) |\langle 0 | \varphi_j \rangle|^2 \right|^2, \quad (\text{B.356})$$

directly leading to Eq. (9.91).

## B.10 Chapter 10

**Exercise 10.1** It is convenient to project the solution  $\psi(\mathbf{r}, t) = \langle \mathbf{r} | \psi(t) \rangle$  of the Schrödinger equation (10.2) onto the basis of eigenfunctions of  $H_0$ . We obtain

$$\psi(\mathbf{r}, t) = \sum_i c_i(t) \phi_i(\mathbf{r}) \exp\left(-i \frac{E_i}{\hbar} t\right), \quad (\text{B.357})$$

where  $\phi_i(\mathbf{r})$  is the eigenfunction of  $H_0$  corresponding to the eigenvalue  $E_i$  (i.e.,  $H_0 \phi_i(\mathbf{r}) = E_i \phi_i(\mathbf{r})$ ). We are interested in the case in which only two states of the atom ( $\phi_g(\mathbf{r})$  and  $\phi_e(\mathbf{r})$ ) are relevant for its dynamical evolution, the corresponding energies being  $E_g = \hbar\omega_g$  and  $E_e = \hbar\omega_e$ . After substitution of (B.357) into (10.2) we obtain

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) &= i\hbar \sum_{i=g,e} [\dot{c}_i(t) - i\omega_i c_i(t)] \phi_i(\mathbf{r}) e^{-i\omega_i t} \\ &= \sum_{k=g,e} (H_0 + H_I) c_k(t) \phi_k(\mathbf{r}) e^{-i\omega_k t} = \sum_{k=g,e} (\hbar\omega_k + H_I) c_k(t) \phi_k(\mathbf{r}) e^{-i\omega_k t}. \end{aligned} \quad (\text{B.358})$$

We now write

$$E(t) = E_0 \cos(\omega t + \phi) = \frac{1}{2} E_0 \left( e^{i(\omega t + \phi)} + e^{-i(\omega t + \phi)} \right) = \alpha e^{i\omega t} + \alpha^* e^{-i\omega t}, \quad (\text{B.359})$$

where  $\alpha = \frac{1}{2} E_0 e^{i\phi}$ . Thus, Eq. (B.358) reads

$$i\hbar \sum_{i=g,e} \phi_i(\mathbf{r}) e^{-i\omega_i t} \dot{c}_i(t) = \sum_{k=g,e} c_k(t) e^{-i\omega_k t} \left[ -e z \left( \alpha e^{i\omega t} + \alpha^* e^{-i\omega t} \right) \right] \phi_k(\mathbf{r}), \quad (\text{B.360})$$

corresponding to two first-order ordinary differential equations in the variables  $c_0$  and  $c_1$ . After multiplication on the left of both members of (B.360) by  $\phi_j^*(\mathbf{r})$  and integration over  $\mathbf{r}$  we obtain, for  $j = g, e$ ,

$$i\hbar e^{-i\omega_j t} \dot{c}_j(t) = \sum_{k=g,e} c_k(t) e^{-i\omega_k t} \int d\mathbf{r} \phi_j^*(\mathbf{r}) \left[ -e z \left( \alpha e^{i\omega t} + \alpha^* e^{-i\omega t} \right) \right] \phi_k(\mathbf{r}), \quad (\text{B.361})$$

where we have taken advantage of the standard orthonormality relation  $\int d\mathbf{r} \phi_j^*(\mathbf{r}) \phi_i(\mathbf{r}) = \delta_{ij}$ . We must evaluate the four integrals

$$D_{jk} = -e \int d\mathbf{r} \phi_j^*(\mathbf{r}) z \phi_k(\mathbf{r}), \quad (j, k = g, e). \quad (\text{B.362})$$

Since  $V(r)$  is spherically symmetric, the eigenfunctions  $\phi_i(\mathbf{r})$  are symmetric or antisymmetric under the inversion  $\mathbf{r} \rightarrow -\mathbf{r}$ . Thus,  $D_{gg} = D_{ee} = 0$  and Eq. (B.361) reads

$$\begin{cases} i\hbar \dot{c}_g(t) = c_e(t) \left[ D \alpha e^{i(\omega - \omega_0)t} + D \alpha^* e^{-i(\omega + \omega_0)t} \right], \\ i\hbar \dot{c}_e(t) = c_g(t) \left[ D^* \alpha e^{i(\omega + \omega_0)t} + D^* \alpha^* e^{-i(\omega - \omega_0)t} \right], \end{cases} \quad (\text{B.363})$$

where we have defined  $D = D_{ge} = D_{eg}^*$  and  $\omega_0 = \omega_e - \omega_g$ . The terms depending on  $\omega + \omega_0$  oscillate very rapidly and can therefore be neglected (*rotating wave approximation*).

Setting

$$\Delta = \omega - \omega_0, \quad \Omega = \frac{D\alpha}{\hbar}, \quad (\text{B.364})$$

we obtain

$$\begin{cases} i\dot{c}_g(t) = \Omega e^{i\Delta t} c_e, \\ i\dot{c}_e(t) = \Omega^* e^{-i\Delta t} c_g. \end{cases} \quad (\text{B.365})$$

To solve this system, it is convenient to differentiate the first equation and substitute the result into the second. We have

$$i\ddot{c}_g = i\Delta\Omega e^{i\Delta t} c_e + \Omega e^{i\Delta t} \dot{c}_e = i\Delta i\dot{c}_g + \Omega e^{i\Delta t} \frac{1}{i} \Omega^* e^{-i\Delta t} c_g, \quad (\text{B.366})$$

that is,

$$\ddot{c}_g - i\Delta\dot{c}_g + |\Omega|^2 c_g = 0. \quad (\text{B.367})$$

This last equation is easily solved by setting  $c_0 = Ae^{i\xi t}$ , which leads to the algebraic equation

$$\xi^2 - \Delta\xi - |\Omega|^2 = 0, \quad (\text{B.368})$$

whose solutions are

$$\xi_{\pm} = \frac{\Delta}{2} \pm \sqrt{\frac{\Delta^2}{4} + |\Omega|^2}. \quad (\text{B.369})$$

Therefore, the general solution of (B.363) is

$$\begin{cases} c_g(t) = A_+ e^{i\xi_+ t} + A_- e^{i\xi_- t}, \\ c_e(t) = -\frac{1}{\Omega} (\xi_+ A_+ e^{i(\xi_+ - \Delta)t} + \xi_- A_- e^{i(\xi_- - \Delta)t}). \end{cases} \quad (\text{B.370})$$

The constants  $A_+$  and  $A_-$  are determined from the initial conditions  $c_g^{(0)} = c_g(t=0)$  and  $c_e^{(0)} = c_e(t=0)$ . We obtain

$$A_+ = \frac{\xi_- c_g^{(0)} + \Omega c_e^{(0)}}{\xi_- - \xi_+}, \quad A_- = \frac{\xi_+ c_g^{(0)} + \Omega c_e^{(0)}}{\xi_+ - \xi_-}. \quad (\text{B.371})$$

We finally substitute these relations into (B.370) and obtain

$$\begin{bmatrix} c_g(t) \\ c_e(t) \end{bmatrix} = U \begin{bmatrix} c_g^{(0)} \\ c_e^{(0)} \end{bmatrix} = \begin{bmatrix} U_{gg} & U_{ge} \\ U_{eg} & U_{ee} \end{bmatrix} \begin{bmatrix} c_g^{(0)} \\ c_e^{(0)} \end{bmatrix}. \quad (\text{B.372})$$

Here  $U$  is a unitary matrix with matrix elements

$$\begin{aligned} U_{gg} &= e^{i\frac{\Delta}{2}t} \left[ \cos(at) - \frac{i\Delta \sin(at)}{2a} \right] = U_{ee}^*, \\ U_{ge} &= e^{i\frac{\Delta}{2}t} \left[ \frac{-i\Omega \sin(at)}{a} \right] = -U_{eg}^*, \end{aligned} \quad (\text{B.373})$$

where

$$a = \sqrt{\frac{\Delta^2}{4} + |\Omega|^2}, \quad \Delta = \omega - \omega_0. \quad (\text{B.374})$$

We are particularly interested in the resonant case in which the detuning parameter  $\Delta = 0$ . It is easy to see that in this case the unitary matrix  $U$  reduces to (10.1), with  $|\Omega|t = \frac{\theta}{2}$ .

**Exercise 10.2** We can expand the solution to the Schrödinger equation for the Jaynes–Cummings model as

$$|\psi(t)\rangle = \sum_{i=g,e} \sum_{n=0}^{\infty} c_{i,n}(t) |i, n\rangle, \quad (\text{B.375})$$

where the indices  $i$  and  $n$  label the atomic state and the number of photons. Therefore, the Schrödinger equation reads

$$i\hbar \sum_{i=g,e} \sum_{n=0}^{\infty} \dot{c}_{i,n}(t) |i, n\rangle = \sum_{i'=g,e} \sum_{n'=0}^{\infty} H c_{i',n'}(t) |i', n'\rangle, \quad (\text{B.376})$$

which implies

$$i\hbar \dot{c}_{i,n}(t) |i, n\rangle = \sum_{i'=g,e} \sum_{n'=0}^{\infty} \langle i, n | H | i', n' \rangle c_{i',n'}(t). \quad (\text{B.377})$$

The matrix elements of the Jaynes–Cummings Hamiltonian can be easily computed and we obtain

$$\begin{aligned} \langle g, n | H | g, n' \rangle &= \hbar \left[ -\frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] \delta_{n,n'}, \\ \langle g, n | H | e, n' \rangle &= \hbar \lambda^* \sqrt{n} \delta_{n,n'+1}, \\ \langle e, n | H | g, n' \rangle &= \hbar \lambda \sqrt{n+1} \delta_{n,n'-1}, \\ \langle e, n | H | e, n' \rangle &= \hbar \left[ \frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] \delta_{n,n'}. \end{aligned} \quad (\text{B.378})$$

After substitution of these matrix elements into (B.377) we obtain

$$\begin{cases} i\hbar \dot{c}_{g,n} = \hbar \left[ -\frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] c_{g,n}(t) + \hbar \lambda^* \sqrt{n} c_{e,n-1}(t), \\ i\hbar \dot{c}_{e,n} = \hbar \left[ \frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] c_{e,n}(t) + \hbar \lambda \sqrt{n+1} c_{g,n+1}(t). \end{cases} \quad (\text{B.379})$$

It is convenient to write the second equation for  $n \rightarrow n-1$ . Then system (B.379) reads

$$\begin{cases} i\dot{c}_{g,n} = \left[ -\frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] c_{g,n}(t) + \lambda^* \sqrt{n} c_{e,n-1}(t), \\ i\dot{c}_{e,n-1} = \left[ \frac{\omega_0}{2} + \omega \left( n - \frac{1}{2} \right) \right] c_{e,n-1}(t) + \lambda \sqrt{n+1} c_{g,n}(t). \end{cases} \quad (\text{B.380})$$

It is now clear that the level  $|g, n\rangle$  is only coupled to  $|e, n-1\rangle$ . In order to solve these equations, we first separate the time dependence due to  $H_0$ , setting

$$\begin{cases} c_{g,n}(t) = \exp \left\{ -i \left[ -\frac{\omega_0}{2} + \omega \left( n + \frac{1}{2} \right) \right] t \right\} \tilde{c}_{g,n}(t), \\ c_{e,n-1}(t) = \exp \left\{ -i \left[ \frac{\omega_0}{2} + \omega \left( n - \frac{1}{2} \right) \right] t \right\} \tilde{c}_{e,n-1}(t). \end{cases} \quad (\text{B.381})$$

We insert these relations into (B.380) and, after defining

$$\Delta = \omega - \omega_0, \quad \Omega_n = \lambda^* \sqrt{n}, \quad (\text{B.382})$$

we obtain

$$\begin{cases} i\dot{\tilde{c}}_{g,n} = \Omega_n \exp(i\Delta t) \tilde{c}_{e,n-1}(t), \\ i\dot{\tilde{c}}_{e,n-1} = \Omega_n^* \exp(-i\Delta t) \tilde{c}_{g,n}(t). \end{cases} \quad (\text{B.383})$$

These equations are analogous to (B.365), obtained for a two-level atom interacting with a classical field and can be solved in the same manner. Thus, in the resonant case  $\Delta = 0$  there are Rabi oscillations between the states  $|g, n\rangle$  and  $|e, n-1\rangle$ . The frequency  $|\Omega_n|$  ( $\Omega_n = |\Omega_n|e^{i\phi_n}$ ) of these oscillations is proportional to the atom–field interaction strength  $|\lambda|$  and to the square root of the number  $n$  of photons in the resonant cavity.

**Exercise 10.3** It is clear from exercise 10.2 that the overall state of the atom–field system at time  $t$  is given by

$$|\psi(t)\rangle = c_{g,n}(t)|g, n\rangle + c_{e,n-1}(t)|e, n-1\rangle. \quad (\text{B.384})$$

The coefficients  $c_{g,n}(t)$  and  $c_{e,n-1}(t)$  are determined by the initial conditions  $c_{g,n}^{(0)} = c_{g,n}(t=0)$ ,  $c_{e,n-1}^{(0)} = c_{e,n-1}(t=0)$ . At resonance ( $\omega = \omega_0$ ) we have

$$\begin{bmatrix} c_{g,n}(t) \\ c_{e,n-1}(t) \end{bmatrix} = \begin{bmatrix} \cos(|\Omega_n|t) & -ie^{i\phi_n} \sin(|\Omega_n|t) \\ -ie^{-i\phi_n} \sin(|\Omega_n|t) & \cos(|\Omega_n|t) \end{bmatrix} \begin{bmatrix} c_{g,n}^{(0)} \\ c_{e,n-1}^{(0)} \end{bmatrix}. \quad (\text{B.385})$$

Since  $|\psi_0\rangle = |g, n\rangle$ , then

$$|\psi(t)\rangle = \cos(|\Omega_n|t)|g, n\rangle - ie^{-i\phi_n} \sin(|\Omega_n|t)|e, n-1\rangle. \quad (\text{B.386})$$

The overall atom–field system at time  $t$  is in a pure state,  $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$ . Thus, the entanglement between the atom and the field is quantified by the reduced von Neumann entropy

$$S_a(t) = -\text{Tr}[\rho_a(t) \log \rho_a(t)], \quad (\text{B.387})$$

where

$$\rho_a(t) = \text{Tr}_f[\rho(t)] = \cos^2(|\Omega_n|t)|g\rangle\langle g| + \sin^2(|\Omega_n|t)|e\rangle\langle e| \quad (\text{B.388})$$

is the reduced density matrix describing the state of the two-level atom (the trace is taken over the field degree of freedom). Therefore,  $S_a(t) = -\sum_{i=1}^2 \lambda_i(t) \log \lambda_i(t)$ , where  $\lambda_1(t) = \cos^2(|\Omega_n|t)$  and  $\lambda_2 = \sin^2(|\Omega_n|t)$ . are the eigenvalues of  $\rho_a$ . The temporal evolution of  $\rho_a(t)$  is shown in Fig. B.13: it oscillates between  $S_a = 0$  (for separable states) and  $S_a = 1$  (for maximally entangled Bell states).

**Exercise 10.4** The initial atom–field state is

$$|\psi_0\rangle = |g\rangle \otimes \sum_{n=0}^{\infty} c_n |n\rangle, \quad (\text{B.389})$$

where

$$c_n = e^{-|\alpha|^2/2} \frac{\alpha^n}{\sqrt{n!}}, \quad \alpha \in \mathbb{C}. \quad (\text{B.390})$$

The temporal evolution of the atom–field state in the Jaynes–Cummings model was discussed in exercise 10.2. Given the initial condition (B.389), we obtain

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} [c_n \cos(\lambda\sqrt{n}t)|g\rangle - ic_{n+1} \sin(\lambda\sqrt{n+1}t)|e\rangle] |n\rangle, \quad (\text{B.391})$$

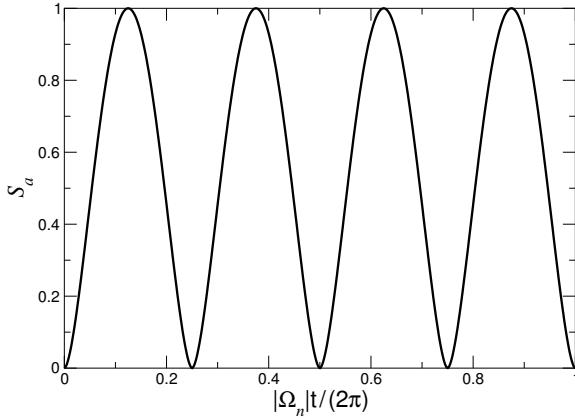


Fig. B.13 The evolution of the entropy  $S_a$  for a two-level atom coupled to a single mode of the electromagnetic field at resonance ( $\omega = \omega_0$ ).

where we have assumed  $\lambda$  real, the detuning parameter  $\Delta = 0$  and set  $\hbar = 1$ . The density matrix describing the atomic state  $\rho^{(a)}$  at time  $t$  is obtained after tracing over the field degree of freedom the overall density matrix  $|\psi(t)\rangle\langle\psi(t)|$ . The matrix elements of  $\rho^{(a)}(t)$  in the  $\{|g\rangle, |e\rangle\}$  basis read as follows:

$$\begin{aligned} \rho_{gg}^{(a)} &= \sum_{n=0}^{\infty} |c_n|^2 \cos^2(\lambda\sqrt{n}t), \\ \rho_{ee}^{(a)} &= 1 - \rho_{gg}^{(a)}, \\ \rho_{ge}^{(a)} &= (\rho_{eg}^{(a)})^* = i \sum_{n=0}^{\infty} c_n^* c_{n+1} \cos(\lambda\sqrt{n}t) \sin(\lambda\sqrt{n+1}t). \end{aligned} \quad (\text{B.392})$$

After writing explicitly the coefficients  $c_n$ , we obtain

$$\begin{aligned} \rho_{gg}^{(a)} &= \sum_{n=0}^{\infty} e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \cos^2(\lambda\sqrt{n}t) \\ \rho_{ee}^{(a)} &= \sum_{n=0}^{\infty} e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \sin^2(\lambda\sqrt{n}t) \\ \rho_{ge}^{(a)} &= (\rho_{eg}^{(a)})^* = i \sum_{n=0}^{\infty} e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \frac{\alpha^*}{\sqrt{n+1}} \cos(\lambda\sqrt{n}t) \sin(\lambda\sqrt{n+1}t). \end{aligned} \quad (\text{B.393})$$

Given the density matrix  $\rho^{(a)}(t)$ , it is easy to determine the Bloch-sphere coordinates  $(x(t), y(t), z(t))$  of the two-level atom and the entropy  $S(t) = -\sum_{i=1}^2 \lambda_i(t) \log \lambda_i(t)$ , where  $\lambda_1(t), \lambda_2(t)$  are the eigenvalues of  $\rho^{(a)}(t)$ . The von Neumann entropy  $S(t)$  is a measure of the atom-field entanglement.

Examples of the temporal evolution of a two-level atom interacting with a single-mode field, initially prepared in a coherent state, are shown in Figs. B.14–B.15. The first figure refers to the short-time motion. The representative point describing the atomic state exhibits a motion similar to a spiral and *collapses* to the centre of the Bloch sphere. Thus, the state of the atom is no longer pure and its entropy is non-zero. The usual Rabi

oscillations between the atomic states  $|g\rangle$  and  $|e\rangle$  are damped. Note that the number of oscillations before damping dominates increases when the average number of photons  $\bar{n}$  in the field is larger. This is quite natural as the transition to the classical electromagnetic field takes place when the number of photons  $\bar{n} \rightarrow \infty$ . The behaviour of the atomic state at longer times is shown in Fig. B.15. The main feature of this figure is the existence of *revivals*; that is, at times much longer than the damping time, the amplitude of the Rabi oscillations increases.

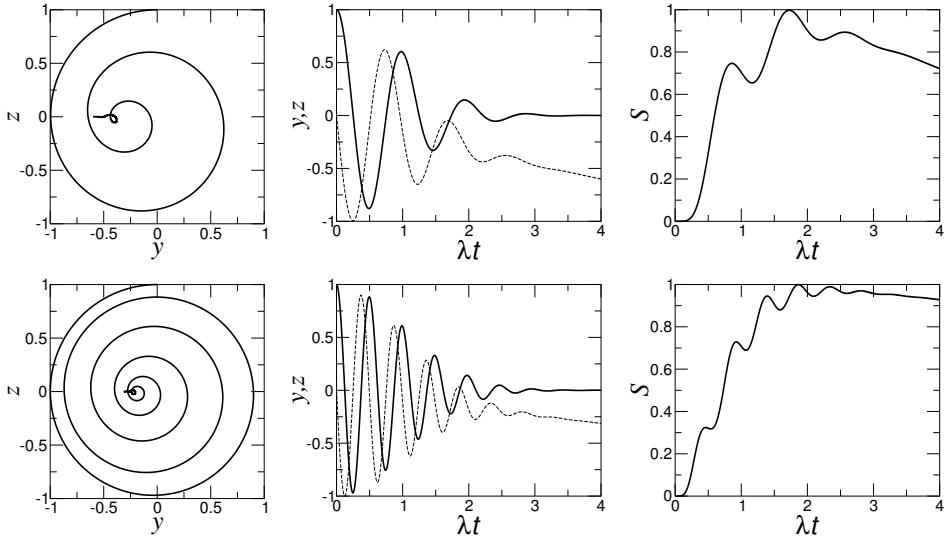


Fig. B.14 The temporal evolution of a two-level atom interacting with an initially coherent single-mode field with  $\alpha = \sqrt{\bar{n}} = \sqrt{10}$  (top) and  $\sqrt{40}$  (bottom): Bloch sphere trajectory (left), Bloch-sphere coordinates  $y$  (dashed curve) and  $z$  (full curve) versus time (middle) and entropy  $S$  versus time (right). Note that for the chosen initial conditions  $x = 0$  at all times.

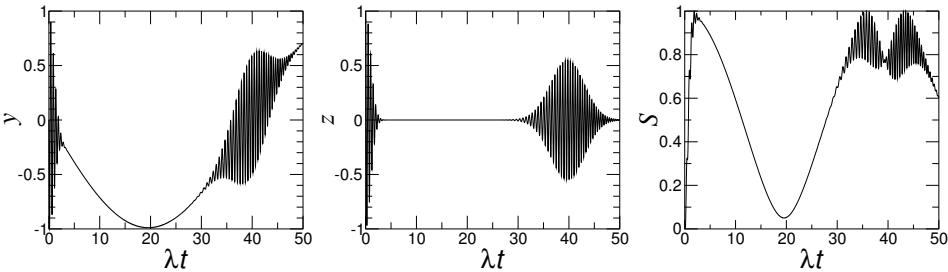


Fig. B.15 The same as in Fig. B.14 but at longer times, for  $\bar{n} = 40$ .

The phenomenon of collapse and revival is repeated with increasing time and can be qualitatively understood as follows. Rabi oscillations of  $z(t)$  are determined by the sum of oscillating terms  $\propto \cos^2(\lambda\sqrt{n}t)$  appearing in  $\rho_{gg}(t) = \frac{1}{2}(1 + z(t))$ . Each term in the summation represents Rabi oscillations with frequency  $\propto \sqrt{n}$ . At time  $t = 0$  all these terms are correlated. As time goes on, the Rabi oscillations associated with different values of  $n$

have different frequencies and therefore become uncorrelated leading to the phenomenon of collapse. Correlation between these contributions is restored, at least partially, at longer times and revivals occur. Note that revivals are purely quantum phenomena and are due to the discrete structure of the photon distribution (only integer  $n$  values are allowed).

The relevant time scales can be estimated as follows: the period  $t_R$  of Rabi oscillations is given by the inverse of the Rabi frequency at  $n = \bar{n}$ ; that is,  $t_R \sim 1/\Omega_{\bar{n}} \sim 1/\lambda\sqrt{\bar{n}}$ . These oscillations continue until the collapse time  $t_c$  when the terms associated with different  $n$  values become dephased. Since in the initial coherent state approximately  $\Delta n = \sqrt{\bar{n}}$  Fock states are relevant,  $t_c$  can be estimated from the condition  $(\Omega_{\bar{n}+\Delta n} - \Omega_{\bar{n}-\Delta n})t_c \sim 1$ , leading to  $t_c \sim 1/\lambda$ . Finally, the revival times  $t_r^{(m)}$  can be estimated by requiring that the phases of the oscillations corresponding to neighbouring photon numbers differ by an integral multiple of  $2\pi$ ; that is, when  $(\Omega_{\bar{n}} - \Omega_{\bar{n}-1})t_r^{(m)} = 2\pi m$ . This implies  $t_r^{(m)} \sim m\sqrt{\bar{n}}/\lambda$ . In particular, the time of the first revival ( $m = 1$ ) is  $\sim \sqrt{\bar{n}}/\lambda$ . Therefore, revivals, which are a purely quantum phenomenon, require longer and longer times when  $\bar{n} \rightarrow \infty$ .

It is instructive to evaluate the entropy at the first Rabi oscillation; that is, for  $\lambda\sqrt{\bar{n}}\tilde{t} = \frac{\pi}{2}$ , in the limit of large mean photon number  $\bar{n}$  (note that  $\tilde{t}$  corresponds to half of the period of Rabi oscillation; that is,  $\tilde{t} = \frac{1}{2}t_R$ ). We first evaluate the matrix elements of  $\rho^{(a)}(t)$  in (B.393), assuming that only the terms with  $n - \bar{n} \ll \bar{n}$  contribute significantly. This is the case as for a coherent state the root mean square deviation  $\Delta n$  in the photon number is equal to  $\sqrt{\bar{n}} \ll \bar{n}$ . We then obtain  $\rho_{gg}^{(a)}(\tilde{t}) \approx \frac{\pi^2}{16\bar{n}}$ ,  $|\rho_{ge}^{(a)}(\tilde{t})| \approx \frac{\pi}{8\bar{n}}$ , from which we can compute  $\lambda_1(\tilde{t}) = \frac{\pi^2}{16\bar{n}^2}$  and  $\lambda_2(\tilde{t}) = 1 - \frac{\pi^2}{16\bar{n}^2}$ . Note that  $S(\tilde{t}) \propto (1/\bar{n}^2)\log(1/\bar{n}^2) \rightarrow 0$  when  $\bar{n} \rightarrow \infty$ . This is expected since there is no decoherence induced by a classical electromagnetic field.

**Exercise 10.5** We apply the unitary transformation (10.1) to both atoms, with  $\theta = \frac{\pi}{2}$  and phases  $\phi_1$  and  $\phi_2$ . The global unitary transformation is given, in the  $\{|g_1, g_2\rangle, |g_1, e_2\rangle, |e_1, g_2\rangle, |e_1, e_2\rangle\}$  basis, by

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -ie^{i\phi_1} \\ -ie^{-i\phi_1} & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -ie^{i\phi_2} \\ -ie^{-i\phi_2} & 1 \end{bmatrix}. \quad (\text{B.394})$$

After application of  $U$  to the Bell state  $\frac{1}{\sqrt{2}}(|e_1, g_2\rangle - |g_1, e_2\rangle)$  we obtain the state

$$\begin{aligned} \frac{1}{2\sqrt{2}} & \left[ -i(e^{i\phi_2} - e^{i\phi_1})|g_1, g_2\rangle - (1 + e^{i(\phi_1 - \phi_2)})|g_1, e_2\rangle \right. \\ & \left. + (1 + e^{i(\phi_2 - \phi_1)})|e_1, g_2\rangle + i(e^{-i\phi_1} - e^{-i\phi_2})|e_1, e_2\rangle \right], \quad (\text{B.395}) \end{aligned}$$

from which the probabilities (10.15) directly follow.

**Exercise 10.6** (i) The electric field is given by  $\mathbf{E} = -\nabla\Phi$  and the equations of motion for the ion are  $M\ddot{\mathbf{r}} = q\mathbf{E}$ . Thus, we obtain

$$\ddot{x} = -\frac{q}{M} \left[ \frac{U_0}{R^2} \cos(\omega_{RFT}) - \epsilon \frac{V_0}{R^2} \right] x, \quad (\text{B.396a})$$

$$\ddot{y} = \frac{q}{M} \left[ \frac{U_0}{R^2} \cos(\omega_{RFT}) + (1 - \epsilon) \frac{V_0}{R^2} \right] y, \quad (\text{B.396b})$$

$$\ddot{z} = -\frac{q}{M} \frac{V_0}{R^2} z. \quad (\text{B.396c})$$

Therefore, the motion along  $z$  is harmonic, with frequency

$$\omega_z = \sqrt{\frac{qV_0}{MR^2}}, \quad (\text{B.397})$$

while the motion along  $y$  and  $z$  is governed by the Mathieu equation (10.17), where

$$a_x = -\frac{4q\epsilon V_0}{M\omega_{RF}^2 R^2}, \quad q_x = \frac{2qU_0}{M\omega_{RF}^2 R^2} \quad (\text{B.398})$$

for  $\xi = x$  and

$$a_y = -\frac{4q(1-\epsilon)V_0}{M\omega_{RF}^2 R^2}, \quad q_y = -\frac{2qU_0}{M\omega_{RF}^2 R^2} \quad (\text{B.399})$$

for  $\xi = y$ . The analytic treatment of the Mathieu equation can be found, for instance, in Leibfried *et al.* (2003a). Here in Fig. B.16 we simply draw, for small  $q, a$ , the stability region in the  $q$ - $a$  plane, as obtained from the numerical integration of the Mathieu equation.

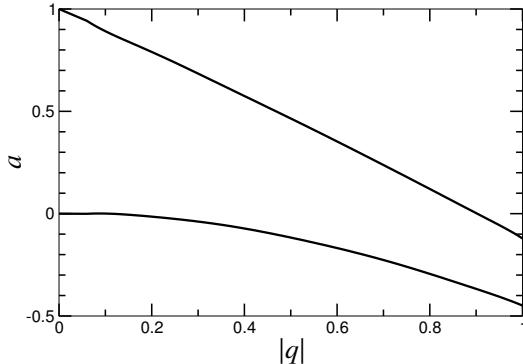


Fig. B.16 The stability diagram for the Mathieu equation. The two solid curves bound the stability region.

(ii) As an example, we compare in Fig. B.17 the approximate solution (10.18) to the Mathieu equation (10.17) for  $q = 0.3$  and various  $a$  values. It is clear that (10.18) is a good approximation for small  $a$  (and  $q$ ), while, as expected, it differs more and more from the exact solution when the stability border is approached.

It is interesting that the approximate solution (10.18) gives harmonic oscillations of size  $\xi_0$ , at a frequency  $\omega_\xi = \beta_\xi \omega_{RF}/2 \ll \omega_{RF}$  (the *secular motion*), superposed with smaller driven excursions of size  $\xi_0 q_\xi/2 \ll \xi_0$ , at a frequency  $\omega_{RF}$  (the *micromotion*). If the fast and small oscillations of the micromotion are neglected, the motion can be approximated by that of a harmonic oscillator at a frequency  $\omega_\xi$ . Such a harmonic approximation is used in the rest of this section, starting from the ion trap Hamiltonian (10.19).

**Exercise 10.7** The balance between the harmonic and the Coulomb forces gives the equilibrium positions. This leads, for  $N = 2$ , to the following equations:

$$-M\omega_z^2 z_1 - \frac{q^2}{4\pi\epsilon_0(z_2 - z_1)^2} = 0, \quad -M\omega_z^2 z_2 + \frac{q^2}{4\pi\epsilon_0(z_2 - z_1)^2} = 0. \quad (\text{B.400})$$

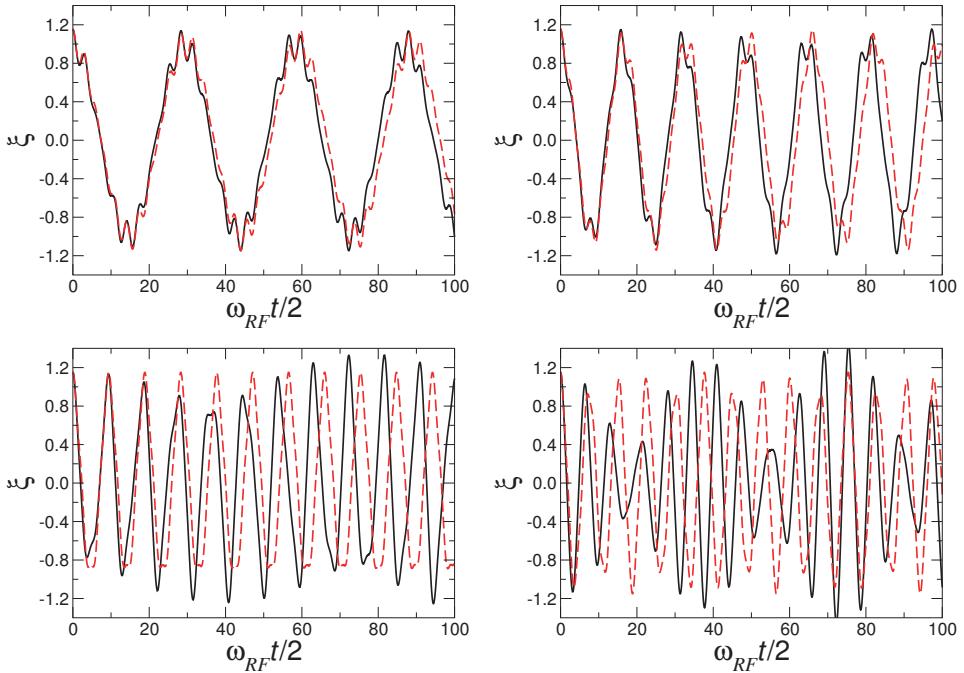


Fig. B.17 A comparison between the numerical integration of Eq. (10.17) (solid line) and the approximate solution (10.18) (dashed line), for  $q = 0.3$ ,  $a = 0$  (top left),  $0.1$  (top right),  $0.4$  (bottom left) and  $0.65$  (bottom right).

It is convenient to employ units in which  $q^2/(4\pi\epsilon_0 M\omega_z^2) = 1$ , so that the above equations become

$$-z_1 - \frac{1}{(z_2 - z_1)^2} = 0, \quad -z_2 + \frac{1}{(z_2 - z_1)^2} = 0, \quad (\text{B.401})$$

with solutions (equilibrium positions)

$$z_1 = z_1^{(0)} = -\left(\frac{1}{4}\right)^{1/3}, \quad z_2 = z_2^{(0)} = -z_1^{(0)} = \left(\frac{1}{4}\right)^{1/3}. \quad (\text{B.402})$$

For  $N = 3$  ions, we obtain

$$\begin{aligned} -z_1 - \frac{1}{(z_2 - z_1)^2} - \frac{1}{(z_3 - z_1)^2} &= 0, \\ -z_2 + \frac{1}{(z_2 - z_1)^2} - \frac{1}{(z_3 - z_2)^2} &= 0, \\ -z_3 + \frac{1}{(z_3 - z_1)^2} + \frac{1}{(z_3 - z_2)^2} &= 0. \end{aligned} \quad (\text{B.403})$$

The sum of these three equations gives the condition  $z_1 + z_2 + z_3 = 0$ . Symmetry considerations tell us that  $z_2 = 0$ ,  $z_1 = -z_3$ , thus obtaining the equilibrium positions

$$z_1 = z_1^{(0)} = -\left(\frac{5}{4}\right)^{1/3}, \quad z_2 = z_2^{(0)} = 0, \quad z_3 = z_3^{(0)} = -z_1^{(0)} = \left(\frac{5}{4}\right)^{1/3}. \quad (\text{B.404})$$

Let us now compute the normal modes of vibration of the string about the equilibrium positions. Starting from Hamiltonian (10.19) we obtain the equations of motion

$$\ddot{z}_i = -z_i + \sum_{k=1}^{i-1} \frac{1}{(z_i - z_k)^2} - \sum_{k=i+1}^N \frac{1}{(z_k - z_i)^2}, \quad (i = 1, \dots, N), \quad (\text{B.405})$$

where we have set the unit of time in such a manner that  $\omega_z = 1$ . By linearizing (B.405) around equilibrium positions and setting

$$z_i = z_i^{(0)} + \xi_i, \quad (i = 1, \dots, N), \quad (\text{B.406})$$

we obtain

$$\ddot{\xi}_i = -\xi_i - 2 \sum_{k=1}^{i-1} \frac{\xi_i - \xi_k}{(z_i^{(0)} - z_k^{(0)})^3} + 2 \sum_{k=i+1}^N \frac{\xi_k - \xi_i}{(z_k^{(0)} - z_i^{(0)})^3}, \quad (i = 1, \dots, N). \quad (\text{B.407})$$

For  $N = 2$  Eqs. (B.407) read as follows:

$$\ddot{\xi}_1 = -\xi_1 + 2 \frac{\xi_2 - \xi_1}{(z_2^{(0)} - z_1^{(0)})^3}, \quad \ddot{\xi}_2 = -\xi_2 - 2 \frac{\xi_2 - \xi_1}{(z_2^{(0)} - z_1^{(0)})^3}. \quad (\text{B.408})$$

Looking for normal modes,  $\xi_i(t) = a_i e^{i\omega t}$ , we obtain

$$-\omega^2 \xi_1 + \xi_1 - 2 \frac{\xi_2 - \xi_1}{(z_2^{(0)} - z_1^{(0)})^3} = 0, \quad -\omega^2 \xi_2 + \xi_2 + 2 \frac{\xi_2 - \xi_1}{(z_2^{(0)} - z_1^{(0)})^3} = 0. \quad (\text{B.409})$$

After substitution of the equilibrium positions (B.402) into (B.409) we arrive to the eigenvalue equation  $(2 - \omega^2)^2 - 1 = 0$ , whose solutions are

$$\omega_1 = 1, \quad \omega_2 = \sqrt{3}. \quad (\text{B.410})$$

The corresponding eigenstates are  $(\xi_1 = 1, \xi_2 = 1)$  (centre-of-mass mode) and  $(-1, 1)$  (stretch mode).

For  $N = 3$ , we similarly derive the eigenvalue equation

$$\det \begin{bmatrix} -\omega^2 + \frac{14}{5} & -\frac{8}{5} & -\frac{1}{5} \\ -\frac{8}{5} & -\omega^2 + \frac{21}{5} & -\frac{8}{5} \\ -\frac{1}{5} & -\frac{8}{5} & -\omega^2 + \frac{14}{5} \end{bmatrix} = 0, \quad (\text{B.411})$$

whose solutions are

$$\omega_1 = 1, \quad \omega_2 = \sqrt{3}, \quad \omega_3 = \sqrt{\frac{29}{5}}. \quad (\text{B.412})$$

The corresponding eigenstates are  $(\xi_1 = 1, \xi_2 = 1, \xi_3 = 1)$  (centre-of-mass mode),  $(-1, 0, 1)$  (stretch mode) and  $(1, -2, 1)$ .

**Exercise 10.8** We look for a unitary operator  $U$  such that  $U|g, 0\rangle = \sum_{n=0}^N c_n |g, n\rangle$ . We first construct  $U^{-1}$  step-by-step; that is, starting from a generic superposition  $\sum_{n=0}^N c_n |g, n\rangle$ , the operator  $U^{-1}$  maps this state into the ground state  $|g, 0\rangle$ . As a first step, a red detuned laser transfers the entire population of the state  $|g, N\rangle$  into  $|e, N-1\rangle$ . Then a tuned laser moves the entire population of  $|e, N-1\rangle$  into  $|g, N-1\rangle$ . We then transfer  $|g, N-1\rangle$  into  $|e, N-2\rangle$  and so on. Note that at each transformation we must keep track of what is happening to all populated levels of the trapped ion. Inverting the procedure illustrated above, we obtain the operator  $U$ . Note that  $2N$  single-qubit (-ion) gates are required to construct a generic  $N$ -level state. Further details can be found in Gardiner *et al.* (1997).

**Exercise 10.9** (i) The term  $e^{ikz}$  in (10.25) couples the radiation to the motion of the trapped ion. Therefore, at resonance it is possible to induce transitions  $|g, n\rangle \leftrightarrow |e, n'\rangle$ , with renormalized Rabi frequency

$$\Omega_{n,n'} = \Omega |\langle n' | e^{ikz} | n \rangle|. \quad (\text{B.413})$$

In order to evaluate the matrix element  $\langle n' | e^{ikz} | n \rangle$ , we employ the formula

$$e^{A+B} = e^A e^B e^{-\frac{[A,B]}{2}}, \quad (\text{B.414})$$

which is valid when the operators  $A$  and  $B$  are such that

$$[A, [A, B]] = 0 = [B, [A, B]]. \quad (\text{B.415})$$

Therefore, taking  $A = i\eta a^\dagger$  and  $B = i\eta a$ , we obtain

$$e^{ikz} = e^{ikz_0(a^\dagger + a)} = e^{-\frac{1}{2}\eta^2} e^{i\eta a^\dagger} e^{i\eta a}. \quad (\text{B.416})$$

Since

$$e^{i\eta a} |n\rangle = \sum_{m=0}^n \frac{1}{m!} (i\eta)^m \sqrt{\frac{n!}{(n-m)!}} |n-m\rangle, \quad (\text{B.417})$$

we obtain

$$\langle n' | e^{ikz} | n \rangle = e^{-\frac{1}{2}\eta^2} \sum_{m'=0}^{n'} \frac{(-i\eta)^{m'}}{m'!} \sqrt{\frac{n'!}{(n'-m')!}} \langle n' - m' | \sum_{m=0}^n \frac{(i\eta)^m}{m!} \sqrt{\frac{n!}{(n-m)!}} |n-m\rangle. \quad (\text{B.418})$$

Due to the orthogonality of the Fock states,  $\langle n' - m' | n - m \rangle = \delta_{n'-m', n-m}$ . Thus, Eq. (B.418) becomes

$$\langle n' | e^{ikz} | n \rangle = e^{-\frac{1}{2}\eta^2} \sum_{m=\max(0, n-n')}^n \frac{(-i\eta)^{(m+n'-n)} (i\eta)^m}{m!(m+n'-n)!} \frac{\sqrt{n'! n!}}{(n-m)!}. \quad (\text{B.419})$$

(ii) When  $\eta \ll 1$ ; that is, the wavelength of the laser is much larger than the extension of the ion wave function, we can expand the exponential in (10.25) to the first order in  $\eta$ :

$$e^{ikz} = e^{i\eta(a^\dagger + a)} \approx 1 + i\eta a^\dagger + i\eta a. \quad (\text{B.420})$$

The three terms on the right-hand side of this equation are associated with the carrier resonance and the first blue and red sidebands.

**Exercise 10.10** First of all, we write down the equations of motion for the three-level system:

$$\begin{aligned}\dot{c}_g &= i c_e \Omega_2 \exp[-i(\omega_{gc} - \omega_2)t + i\phi_2], \\ \dot{c}_e &= i c_g \Omega_1 \exp[-i(\omega_{gc} - \omega_0 - \omega_1)t + i\phi_1], \\ \dot{c}_c &= i c_g \Omega_2 \exp[i(\omega_{gc} - \omega_2)t - i\phi_2] + i c_e \Omega_1 \exp[i(\omega_{gc} - \omega_0 - \omega_1)t - i\phi_1].\end{aligned}\quad (\text{B.421})$$

Note that, when  $\Omega_1 = 0$  (or  $\Omega_2 = 0$ ), we recover the well-known equations of motion for a two-level system in a classical electromagnetic field.

We now make the following assumptions:

$$\Delta \equiv \omega_{gc} - \omega_2 \gg \Omega_1, \Omega_2, \quad \omega_{gc} - \omega_2 \approx \omega_{gc} - \omega_0 - \omega_1. \quad (\text{B.422})$$

Therefore,  $c_g(t)$  and  $c_e(t)$  change in time much more slowly than  $c_c(t)$  and we can integrate the last equation in (B.421) neglecting the variation in time of  $c_g$  and  $c_e$ . We also assume that  $c_c(t = 0) = 0$ . This leads to

$$c_c = \frac{c_g \Omega_2}{\Delta} e^{i(\Delta t - \phi_2)} + \frac{c_e \Omega_1}{\Delta} e^{i(\Delta t - \phi_1)}. \quad (\text{B.423})$$

After substitution of this expression for  $c_c$  into the first two equations of (B.421), we obtain

$$\dot{c}_g = i \frac{\Omega_2^2}{\Delta} c_g + i \frac{\Omega_1 \Omega_2}{\Delta} e^{i(\phi_2 - \phi_1)} c_e, \quad \dot{c}_e = i \frac{\Omega_1 \Omega_2}{\Delta} e^{-i(\phi_2 - \phi_1)} c_g + i \frac{\Omega_1^2}{\Delta} c_e. \quad (\text{B.424})$$

It is clear that the Rabi frequency  $\Omega_R$  for this two-level system is given by Eq. (10.29).

Note that, since the exact equation for  $\dot{c}_c$  (B.421) contains terms oscillating with frequency  $\Delta$ , the Raman approximation is valid when

$$\Omega_R = \frac{2\Omega_1 \Omega_2}{\Delta} \ll \Delta. \quad (\text{B.425})$$

Such a condition is fulfilled when  $\Omega_1, \Omega_2 \ll \Delta$ . Moreover, we can estimate from (B.423) that

$$|c_c| \approx \frac{\Omega_2}{\Delta} |c_g|, \quad |c_c| \approx \frac{\Omega_1}{\Delta} |c_e|. \quad (\text{B.426})$$

Since  $|c_g|, |c_e| \leq 1$ , we have that the population of level  $|c\rangle$  does not exceed

$$|c_c| \approx \frac{\Omega}{\Delta} \ll 1, \quad (\text{B.427})$$

where we have considered the special case  $\Omega \equiv \Omega_1 = \Omega_2$ . Therefore, level  $c$  is very weakly populated.

A comparison between the Raman approximation and the direct numerical integration of the equations of motion (B.421) is shown in Fig. B.18. It is clear that the Raman approximation is quite good in the case  $\Omega/\Delta = 0.1 \ll 1$ .

**Exercise 10.11** The motion of the particle is bounded within the interval  $[0, a]$ ; that is, its wave function  $\phi(x)$  must be zero outside this interval. Continuity of the wave function at  $x = 0$  and  $x = a$  implies

$$\lim_{x \rightarrow 0^+} \phi(x) = \lim_{x \rightarrow a^-} \phi(x) = 0. \quad (\text{B.428})$$

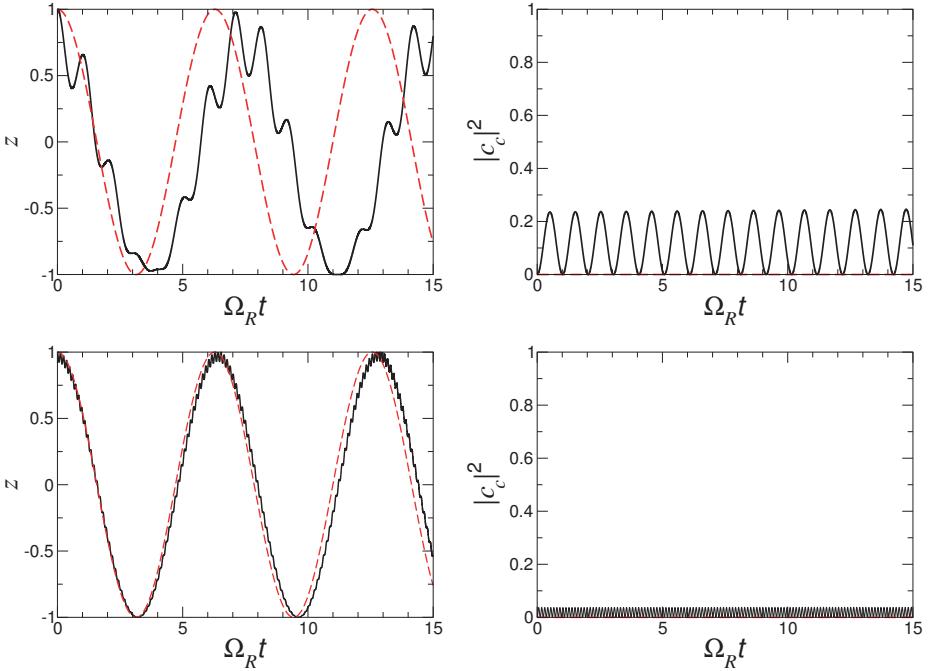


Fig. B.18 A comparison between the numerical integration of Eqs. (B.421) (solid line) and the Raman approximation (dashed line), for  $\Omega_1 = \Omega_2 = 0.1$ ,  $\phi_1 = \phi_2 = 0$ , initial conditions  $c_g = 1$ ,  $c_e = c_c = 0$ ,  $\Delta = 0.3$  (above) and  $\Delta = 1$  (below): temporal evolution of the  $z$  coordinate of the Bloch sphere for the qubit spanned by the levels  $|g\rangle$  and  $|e\rangle$  (left figures) and population of level  $|c\rangle$  (right figures).

The solutions to the stationary Schrödinger equation  $H\phi(x) = E\phi(x)$  inside the interval  $[0, a]$  can be written as

$$\phi(x) = Ae^{ikx} + A'e^{-ikx}, \quad (\text{B.429})$$

where  $A$  and  $A'$  are complex constants. Since  $\phi(0) = 0$ , it follows that  $A' = -A$ , and therefore

$$\phi(x) = 2iA \sin(kx). \quad (\text{B.430})$$

Moreover,  $\phi(a) = 0$ , which leads to

$$k = \frac{n\pi}{a}, \quad (\text{B.431})$$

where  $n$  is an arbitrary positive integer. If we normalize (B.430); that is, we require  $\int_{-\infty}^{+\infty} dx |\phi(x)|^2 = \int_0^a dx |\phi(x)|^2 = 1$ , and we take into account (B.431), we then obtain the stationary wave functions

$$\phi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi n}{a} x\right), \quad (\text{B.432})$$

with energies

$$E_n = \frac{\pi^2 \hbar^2}{2ma^2} n^2. \quad (\text{B.433})$$

The general solution of the Schrödinger equation  $i\hbar \frac{\partial}{\partial t} \psi(x, t) = H\psi(x, t)$  is

$$\psi(x, t) = \sum_{n=1}^{\infty} c_n e^{-\frac{i}{\hbar} E_n t} \phi_n(x), \quad (\text{B.434})$$

where the coefficients  $c_n$  are determined by the initial condition  $\psi(x, 0) = \psi_0(x)$ :

$$c_n = \int_0^a dx \psi_0(x) \phi_n(x). \quad (\text{B.435})$$

**Exercise 10.12** Since we are looking for bound states, we limit ourselves to studying the case  $-V_0 < E < 0$ . Taking into account the boundary condition  $\lim_{x \rightarrow \pm\infty} \phi(x) = 0$ , we obtain

$$\phi(x) = \begin{cases} A \exp(kx), & x < -a, \\ B \cos(k'x) + C \sin(k'x), & -a \leq x \leq a, \\ D \exp(-kx), & x > a, \end{cases} \quad (\text{B.436})$$

with

$$k = \frac{1}{\hbar} \sqrt{-2mE}, \quad k' = \frac{1}{\hbar} \sqrt{2m(V_0 + E)}. \quad (\text{B.437})$$

By requiring the continuity of  $\phi$  and  $d\phi/dx$  at  $x = \pm a$  we derive four linear homogeneous equations in the variables  $A, B, C, D$ :

$$\begin{aligned} A \exp(-ka) &= B \cos(k'a) - C \sin(k'a), \\ D \exp(-ka) &= B \cos(k'a) + C \sin(k'a), \\ kA \exp(-ka) &= k'B \sin(k'a) + k'C \cos(k'a), \\ -kD \exp(-ka) &= -k'B \sin(k'a) + k'C \cos(k'a). \end{aligned} \quad (\text{B.438})$$

After appropriately adding and subtracting these relations we have

$$\begin{aligned} (A + D) \exp(-ka) - 2B \cos(k'a) &= 0, \\ k(A + D) \exp(-ka) - 2k'B \sin(k'a) &= 0, \\ (A - D) \exp(-ka) + 2C \sin(k'a) &= 0, \\ k(A - D) \exp(-ka) - 2k'C \cos(k'a) &= 0. \end{aligned} \quad (\text{B.439})$$

Non-trivial solutions are obtained when

$$\det \begin{bmatrix} \exp(-ka) & -2 \cos(k'a) & 0 & 0 \\ k \exp(-ka) & -2k' \sin(k'a) & 0 & 0 \\ 0 & 0 & \exp(-ka) & 2 \sin(k'a) \\ 0 & 0 & k \exp(-ka) & -2k' \cos(k'a) \end{bmatrix}, \quad (\text{B.440})$$

that is, when one of the following two equations is satisfied:

$$k' \sin(k'a) - k \cos(k'a) = 0, \quad (\text{B.441a})$$

$$k' \cos(k'a) + k \sin(k'a) = 0. \quad (\text{B.441b})$$

If (B.441a) is satisfied, then

$$D = A, \quad C = 0, \quad B = \frac{\exp(-ka)}{\cos(k'a)} A, \quad (\text{B.442})$$

with  $A$  determined from the normalization condition  $\int_{-\infty}^{+\infty} dx |\phi(x)|^2 = 1$ . Note that this solution is even, namely  $\phi(-x) = \phi(x)$ . On the other hand, if (B.441b) is satisfied we obtain

$$D = -A, \quad B = 0, \quad C = -\frac{\exp(-ka)}{\cos(k'a)} A \quad (\text{B.443})$$

and again  $A$  is determined from the normalization of the wave function. This solution is odd,  $\phi(-x) = -\phi(x)$ .

We now find the energy levels. First of all, we observe that (B.437) leads to

$$k^2 + k'^2 = \frac{2mV_0}{\hbar^2} \equiv k_0^2. \quad (\text{B.444})$$

In the case of even eigenfunctions, the energy levels are determined by the intersections of (B.444) with (B.441a), in the case of odd eigenfunctions by the intersections of (B.444) with (B.441b). The solutions can be found graphically, as shown in Fig. B.19. Note that the number of bound states depends on the parameter  $\sqrt{k_0 a}$ , that is, on the depth  $V_0$  of the square well. If  $k_0 a \leq \frac{\pi}{2}$ , namely  $V_0 \leq \bar{V} \equiv \frac{\pi^2 \hbar^2}{8ma^2}$ , there exists only one bound state of the particle, corresponding to an even wave function. If  $\frac{\pi}{2} < k_0 a \leq \pi$ , that is,  $\bar{V} < V_0 \leq 4\bar{V}$ , we have two bound states since a level corresponding to an odd wave function appears, and so on. Note that energy levels corresponding to even and odd wave functions appear alternatively as  $V_0$  increases.

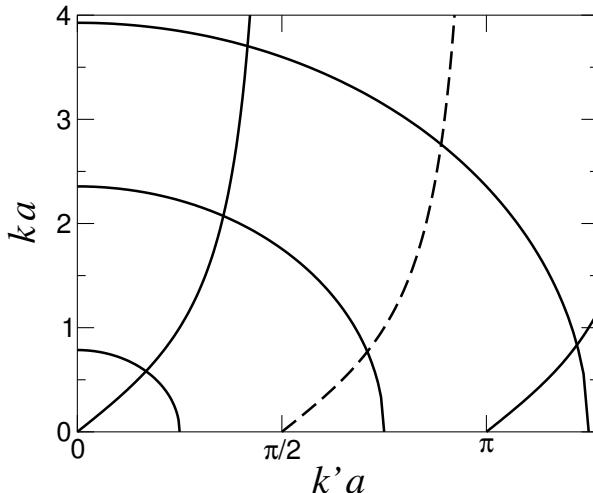


Fig. B.19 The energy levels of a particle in a square well potential determined graphically. The circular arcs have radius  $k_0 a = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}$ . The intersections of these arcs with the solid and dashed curves determine the energy levels corresponding to even and odd eigenfunctions, respectively.

**Exercise 10.13** It is useful to define  $\alpha = \frac{1+i}{\sqrt{2}}$  and  $\beta = \frac{1-i}{\sqrt{2}}$ . Then we have

$$\begin{aligned} \sqrt{\text{SWAP}} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \beta^* & \beta & 0 \\ 0 & \beta & \beta^* & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\sqrt{\text{SWAP}})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \beta & \beta^* & 0 \\ 0 & \beta^* & \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ \text{SWAP} &= (\sqrt{\text{SWAP}})^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned} \quad (\text{B.445})$$

$$\begin{aligned} I \otimes R_z(\pi/2) &= \text{diag}(\alpha^*, \alpha, \alpha^*, \alpha), \quad I \otimes R_z(-\pi/2) = (I \otimes R_z(\pi/2))^*, \\ I \otimes R_z(\pi) &= \text{diag}(-i, i, -i, i). \end{aligned} \quad (\text{B.446})$$

If we multiply the above matrices, as in (10.32), up to an irrelevant global phase factor, we obtain the CMINUS quantum gate.

**Exercise 10.14** We wish to compute the transmission and reflection probabilities for an incident particle, with momentum  $\hbar k$  (and energy  $E = \frac{\hbar^2 k^2}{2m}$ ), propagating from left to right (that is, coming from  $x \rightarrow -\infty$ ). We consider the case  $0 < E < V_0$ . The solution to the stationary Schrödinger equation has the form

$$\phi(x) = \begin{cases} \exp(ikx) + R \exp(-ikx), & x < 0, \\ A \cosh(k'x) + B \sinh(k'x), & 0 \leq x \leq a, \\ T \exp(ikx), & x > a, \end{cases} \quad (\text{B.447})$$

with

$$k = \frac{1}{\hbar} \sqrt{2mE}, \quad k' = \frac{1}{\hbar} \sqrt{2m(V_0 - E)}. \quad (\text{B.448})$$

The first term in the first line of (B.447) is the plane wave describing the incident particle, the second term in the same line corresponds to a reflected particle, with momentum  $-\hbar k$ . Finally, the only term in the last equation of (B.447) is associated with a transmitted particle. The terms  $T$  and  $R$  define the transmission and reflection coefficients, respectively.

Continuity of the wave function and of its derivative at  $x = 0$  and  $x = a$  leads to the following equations:

$$\begin{aligned} 1 + R &= A, \\ A \cosh(k'a) + B \sinh(k'a) &= T \exp(ika), \\ ik(1 - R) &= k'B, \\ k'A \sinh(k'a) + k'B \cosh(k'a) &= ikT \exp(ika). \end{aligned} \quad (\text{B.449})$$

We can solve these equations, thus obtaining

$$\begin{aligned} T &= \frac{2ikk' \exp(-ika)}{2ikk' \cosh(k'a) + [k^2 - k'^2] \sinh(k'a)}, \\ R &= \frac{[k^2 + k'^2] \sinh(k'a)}{2ikk' \cosh(k'a) + [k^2 - k'^2] \sinh(k'a)}, \\ A &= 1 + R, \\ B &= i \frac{k}{k'} (1 - R). \end{aligned} \quad (\text{B.450})$$

It is easy to verify that  $|R|^2 + |T|^2 = 1$ ; namely, the sum of the reflection and transmission probabilities is equal to unity. We emphasize that, in contrast to the classical predictions, the particle has a non-zero probability of crossing the potential barrier even though its energy  $E$  is smaller than the height  $V_0$  of the barrier (tunnel effect).

**Exercise 10.15** The time-evolution operator over a time interval  $T$  reads

$$U(T) = \exp\left[-\frac{iT}{\hbar} \left(\frac{\omega_0 \sigma_z}{2} + \Delta \sigma_x\right)\right] = \cos\left(\frac{\theta(T)}{2}\right) I - i \sin\left(\frac{\theta(T)}{2}\right) \mathbf{n} \cdot \boldsymbol{\sigma}, \quad (\text{B.451})$$

where we have defined

$$\theta(T) = \frac{2T}{\hbar} \sqrt{\frac{1}{4} \omega_0^2 + \Delta^2}, \quad (\text{B.452})$$

$$\mathbf{n} = \left( \frac{\Delta}{\sqrt{\frac{1}{4} \omega_0^2 + \Delta^2}}, 0, \frac{\omega_0}{2\sqrt{\frac{1}{4} \omega_0^2 + \Delta^2}} \right). \quad (\text{B.453})$$

Therefore, the pulse induces a rotation through an angle  $\theta(T)$  about the  $\mathbf{n}$ -axis of the Bloch sphere.

**Exercise 10.16** In the  $\{|g, 0\rangle, |e, 0\rangle, |g, 1\rangle, |e, 1\rangle, \dots\}$  basis, the matrix representation of Hamiltonian  $H_I(t)$  reads as follows:

$$H_I(t) = \hbar \lambda \begin{bmatrix} 0 & 0 & 0 & e^{-2i\omega t} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & \sqrt{2}e^{-2i\omega t} & 0 & 0 & \dots \\ e^{2i\omega t} & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{3}e^{-2i\omega t} & \dots \\ 0 & 0 & \sqrt{2}e^{2i\omega t} & 0 & 0 & 0 & \sqrt{3} & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \sqrt{3} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \sqrt{3}e^{2i\omega t} & 0 & 0 & 0 & \dots \\ \vdots & \vdots \end{bmatrix}. \quad (\text{B.454})$$

Note that the terms beyond the RWA are proportional to  $e^{\pm 2i\omega t}$  and connect  $2 \times 2$  matrix blocks which are disconnected within the RWA.

To compute the first-order terms, we first observe that  $H_I(t)|\psi(0)\rangle = \hbar\lambda e^{2i\omega t}|e, 1\rangle$ . After integrating over time  $H_I(t)|\psi(0)\rangle$  according to Eq. (10.43), we obtain

$$|\psi^{(1)}(t)\rangle = \frac{\lambda}{2\omega} \left(1 - e^{2i\omega t}\right) |e, 1\rangle. \quad (\text{B.455})$$

The second-order terms are computed from Eq. (10.44). We obtain

$$|\psi^{(2)}(t)\rangle = i \frac{\lambda^2}{2\omega} \left[ t + \frac{i}{2\omega} (1 - e^{-2i\omega t}) \right] |g, 0\rangle + i \frac{\sqrt{2}\lambda^2}{2\omega} \left[ -t + \frac{i}{2\omega} (1 - e^{2i\omega t}) \right] |g, 2\rangle. \quad (\text{B.456})$$

It is also interesting to remark in the latter term the  $\sqrt{2}$  factor, which is due to the stimulated emission of the second photon by the first one. The procedure can then be iterated to higher orders.

We can then obtain the mean number of photons  $\langle n \rangle$ . To the first order,

$$\langle n \rangle(t) = \frac{\lambda^2}{\omega^2} \sin^2(\omega t). \quad (\text{B.457})$$

A comparison between the results of perturbation theory (up to fourth order) and numerical results is shown in Fig. B.20.

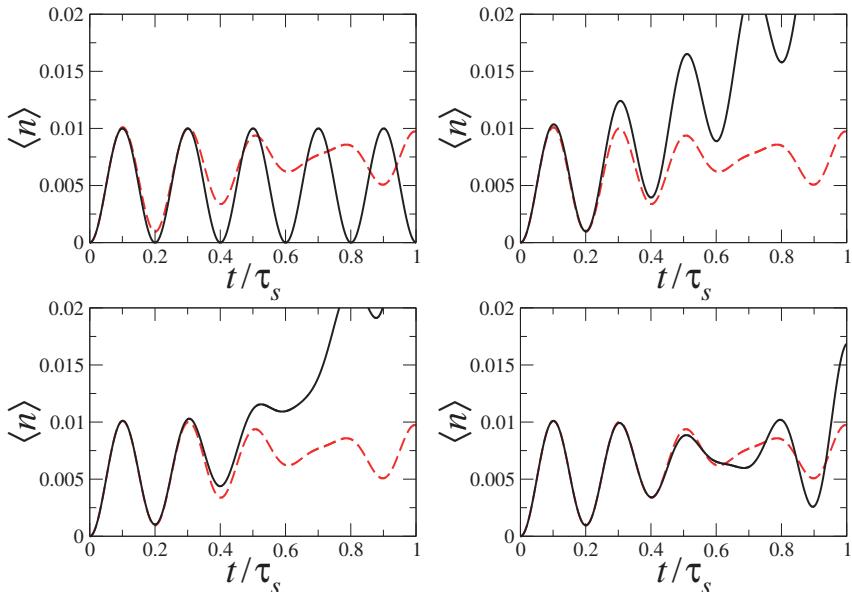


Fig. B.20 Mean number of generated photons as a function of time, measured in units of the swapping time  $\tau_s = \pi/2\lambda$ , which is the time needed to transfer within RWA an excitation from the qubit to the field or vice versa ( $|e, 0\rangle \leftrightarrow |g, 1\rangle$ ). The dotted curve represents the numerical results, the solid curve the perturbative results to first (top left), second (top right), third (bottom left), and fourth (bottom right) order. Here  $\lambda = 0.1\omega$ .

**Exercise 10.17** Let us first consider the circuit in Fig. 10.12 (left). The sequence of gates  $U_P(\phi = -\frac{\pi}{2}) U_B(\theta = \frac{\pi}{4}, \phi = -\frac{\pi}{2}) U_P(\phi = -\frac{\pi}{2})$  transforms the input states  $|0\rangle$  and

$|1\rangle$  as follows:

$$\begin{aligned} |0\rangle &= |1\rangle_0|0\rangle_1 \rightarrow |1\rangle_0|0\rangle_1 \rightarrow \frac{1}{\sqrt{2}}(|1\rangle_0|0\rangle_1 + i|0\rangle_0|1\rangle_1) \\ &\rightarrow \frac{1}{\sqrt{2}}(|1\rangle_0|0\rangle_1 + |0\rangle_0|1\rangle_1) = \frac{1}{\sqrt{2}}(|0'\rangle + |1'\rangle), \\ |1\rangle &= |0\rangle_0|1\rangle_1 \rightarrow -i|0\rangle_0|1\rangle_1 \rightarrow \frac{-i}{\sqrt{2}}(i|1\rangle_0|0\rangle_1 + |0\rangle_0|1\rangle_1) \\ &\rightarrow \frac{1}{\sqrt{2}}(|1\rangle_0|0\rangle_1 - |0\rangle_0|1\rangle_1) = \frac{1}{\sqrt{2}}(|0'\rangle - |1'\rangle). \end{aligned} \quad (\text{B.458})$$

This is exactly a Hadamard transformation.

We now show that the circuit in Fig. 10.12 (right) implements the CNOT gate (up to a sign factor). Indeed, we have

$$\begin{aligned} |0\rangle|h\rangle &= |1\rangle_0|0\rangle_1|h\rangle \rightarrow |1\rangle_{0'}|0\rangle_{1'}|h\rangle = |0'\rangle|h\rangle, \\ |0\rangle|v\rangle &= |1\rangle_0|0\rangle_1|v\rangle \rightarrow |1\rangle_{0'}|0\rangle_{1'}|v\rangle = |0'\rangle|v\rangle, \\ |1\rangle|h\rangle &= |0\rangle_0|1\rangle_1|h\rangle \rightarrow |0\rangle_{0'}|1\rangle_{1'}|v\rangle = |1'\rangle|v\rangle, \\ |1\rangle|v\rangle &= |0\rangle_0|1\rangle_1|v\rangle \rightarrow -|0\rangle_{0'}|1\rangle_{1'}|h\rangle = -|1'\rangle|h\rangle. \end{aligned} \quad (\text{B.459})$$

**Exercise 10.18** We obtain

$$\begin{aligned} HU_P(\phi)H &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= e^{i\frac{\phi}{2}} \begin{bmatrix} \cos \frac{\phi}{2} & -i \sin \frac{\phi}{2} \\ -i \sin \frac{\phi}{2} & \cos \frac{\phi}{2} \end{bmatrix} = e^{i\frac{\phi}{2}} U_B \left( \frac{\phi}{2}, \frac{\pi}{2} \right), \end{aligned} \quad (\text{B.460})$$

with  $U_P$  and  $U_B$  defined in (10.46) and (10.47). Therefore, the entire Mach–Zehnder interferometer corresponds to a beam splitter of transmittance  $T = \cos^2(\frac{\theta}{2})$ . Note that, if the phase shift  $\phi = 0$ , then the photon leaves the interferometer in the same direction as it entered, as expected from the fact that  $H^2 = I$ .

**Exercise 10.19** It is convenient to write the matrix  $U$  in (10.54) as follows:

$$U = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}. \quad (\text{B.461})$$

The transformation of the creation operators is given by

$$a_l^\dagger \rightarrow \sum_m U_{ml} a_m^\dagger, \quad (\text{B.462})$$

namely

$$\begin{aligned} a_1^\dagger &\rightarrow Aa_1^\dagger + Da_2^\dagger + Ga_3^\dagger, \\ a_2^\dagger &\rightarrow Ba_1^\dagger + Ea_2^\dagger + Ha_3^\dagger, \\ a_3^\dagger &\rightarrow Ca_1^\dagger + Fa_2^\dagger + Ka_3^\dagger. \end{aligned} \quad (\text{B.463})$$

Therefore, the initial state  $|\psi\rangle|1\rangle|0\rangle$  is mapped by  $U$  into

$$\left[ \alpha + \beta(Aa_1^\dagger + Da_2^\dagger + Ga_3^\dagger) + \frac{1}{\sqrt{2}}\gamma(Aa_1^\dagger + Da_2^\dagger + Ga_3^\dagger)^2 \right] (Ba_1^\dagger + Ea_2^\dagger + Ha_3^\dagger) |000\rangle. \quad (\text{B.464})$$

Since we accept only measurement outcomes with one photon in mode 2 and no photons in mode 3, in (B.464) we must keep only the terms with mode 2 in the state  $|1\rangle$  and mode 3 in  $|0\rangle$ , thus obtaining

$$\begin{aligned} & \left[ \alpha E + \beta(AE + BD)a_1^\dagger + \frac{1}{\sqrt{2}}\gamma(A^2E + 2ADB)(a_1^\dagger)^2 \right] |010\rangle \\ &= \frac{1}{2} \left[ \alpha + \beta a_1^\dagger - \frac{1}{\sqrt{2}}\gamma(a_1^\dagger)^2 \right] |010\rangle = \frac{1}{2} |\psi'\rangle |10\rangle. \end{aligned} \quad (\text{B.465})$$

Therefore, the final state  $|\psi'\rangle$  is obtained with probability  $\frac{1}{4}$ .

**Exercise 10.20** The initial state can be written as

$$\alpha\gamma|0101\rangle + \alpha\delta|0110\rangle + \beta\gamma|1001\rangle + \beta\delta|1010\rangle. \quad (\text{B.466})$$

Note that, for the sake of simplicity, we omit the indices  $1, \dots, 4$  specifying the modes. Using (10.51) we can see that the first beam splitter transforms (B.466) into

$$\begin{aligned} & \alpha\gamma|0101\rangle - \frac{1}{\sqrt{2}}\alpha\delta|1100\rangle + \frac{1}{\sqrt{2}}\alpha\delta|0110\rangle + \frac{1}{\sqrt{2}}\beta\gamma|1001\rangle \\ &+ \frac{1}{\sqrt{2}}\beta\gamma|0011\rangle - \frac{1}{\sqrt{2}}\beta\delta|2000\rangle + \frac{1}{\sqrt{2}}\beta\delta|0020\rangle. \end{aligned} \quad (\text{B.467})$$

After the two non-linear sign shift gates we have (with overall success probability  $(\frac{1}{4})^2 = \frac{1}{16}$ )

$$\begin{aligned} & \alpha\gamma|0101\rangle - \frac{1}{\sqrt{2}}\alpha\delta|1100\rangle + \frac{1}{\sqrt{2}}\alpha\delta|0110\rangle + \frac{1}{\sqrt{2}}\beta\gamma|1001\rangle \\ &+ \frac{1}{\sqrt{2}}\beta\gamma|0011\rangle + \frac{1}{\sqrt{2}}\beta\delta|2000\rangle - \frac{1}{\sqrt{2}}\beta\delta|0020\rangle. \end{aligned} \quad (\text{B.468})$$

After the final beam splitter we obtain

$$\alpha\gamma|0101\rangle + \alpha\delta|0110\rangle + \beta\gamma|1001\rangle - \beta\delta|0101\rangle. \quad (\text{B.469})$$

## B.11 Chapter 11

**Exercise 11.1** We first demonstrate the commutation relations  $[\sigma_i^+, \sigma_j^-] = \delta_{ij}\sigma_j^z$  and  $[\sigma_i^z, \sigma_j^\pm] = \pm 2\delta_{ij}\sigma_j^\pm$ , between the raising and lowering spin operators  $\sigma_j^\pm = \frac{1}{2}(\sigma_j^x \pm i\sigma_j^y)$ . These can be readily obtained by first using the definitions of  $\sigma_j^\pm$ :

$$[\sigma_i^+, \sigma_j^-] = \frac{1}{4} ([\sigma_i^x, \sigma_j^x] + i[\sigma_i^y, \sigma_j^x] - i[\sigma_i^x, \sigma_j^y] + [\sigma_i^y, \sigma_j^y]), \quad (\text{B.470})$$

$$[\sigma_i^z, \sigma_j^\pm] = \frac{1}{2} ([\sigma_i^z, \sigma_j^x] \pm i[\sigma_i^z, \sigma_j^y]). \quad (\text{B.471})$$

Now from Eq. (A.104) it is clear that, on different sites ( $i \neq j$ ), the above expressions are zero. Conversely Eq. (A.105) tells us that, on the same site ( $i = j$ ), we have:  $[\sigma_j^+, \sigma_j^-] = \frac{1}{4}(i[\sigma_j^y, \sigma_j^x] - i[\sigma_j^x, \sigma_j^y]) = \frac{i}{2}[\sigma_j^y, \sigma_j^x] = \frac{i}{2}(-2i\sigma_j^z) = \sigma_j^z$ ;  $[\sigma_j^z, \sigma_j^\pm] = \frac{1}{2}(2i\sigma_j^y \pm i(-2i\sigma_j^x)) = \pm\sigma_j^x + i\sigma_j^y = \pm(\sigma_j^x \pm i\sigma_j^y) = \pm 2\sigma_j^\pm$ .

For our purpose, it is also useful to remind that  $\sigma_j^+ = a_j^\dagger$ ,  $\sigma_j^- = a_j$ , and  $c_j^\dagger c_j = a_j^\dagger a_j$ . The JWT in Eq. (11.22) can be thus easily inverted to give

$$c_j = e^{-i\pi \sum_{i < j} a_i^\dagger a_i} a_j = \left[ \prod_{i=1}^{j-1} (1 - 2n_i) \right] a_j = \left[ \prod_{i=1}^{j-1} (-\sigma_i^z) \right] a_j. \quad (\text{B.472})$$

Using the above equalities and defining  $K_{\alpha,\beta} = \prod_{i=\alpha}^{\beta-1} (-\sigma_i^z)$ , we have that ( $i \neq j$ ):

$$\begin{aligned}\{c_i, c_j^\dagger\} &= K_{1,i} \sigma_i^- K_{1,j} \sigma_j^+ + K_{1,j} \sigma_j^+ K_{1,i} \sigma_i^- = \sigma_i^- K_{i,j} \sigma_j^+ + \sigma_j^+ K_{i,j} \sigma_i^- \\ &= -(\sigma_i^z \sigma_i^- + 2\sigma_i^-) K_{i+1,j} \sigma_j^+ + K_{i,j} \sigma_j^+ \sigma_i^- = 2K_{i,j} \sigma_i^- \sigma_j^+ - 2\sigma_i^- K_{i+1,j} \sigma_j^+.\end{aligned}\quad (\text{B.473})$$

It is now easy to see that, for  $i \neq j$ , the last equality applied on any quantum state must give zero. Indeed, on the  $i$ -th site we have:  $-2(\sigma_i^z \sigma_i^- + \sigma_i^-) = 0$ . Conversely, if  $i = j$  the string  $K_{1,j}$  appears twice in the anti-commutator, and thus it disappears:  $\{c_j, c_j^\dagger\} = \sigma_j^- \sigma_j^+ + \sigma_j^+ \sigma_j^- = I$ , where  $I$  is the identity matrix.

Proceeding in an analogous way, we can calculate

$$\begin{aligned}\{c_i, c_j\} &= K_{1,i} \sigma_i^- K_{1,j} \sigma_j^- + K_{1,j} \sigma_j^- K_{1,i} \sigma_i^- = \sigma_i^- K_{i,j} \sigma_j^- + \sigma_j^- K_{i,j} \sigma_i^- \\ &= -(\sigma_i^z \sigma_i^- + 2\sigma_i^-) K_{i+1,j} \sigma_j^- + K_{i,j} \sigma_j^- \sigma_i^- = 2K_{i,j} \sigma_i^- \sigma_j^- - 2\sigma_i^- K_{i+1,j} \sigma_j^-.\end{aligned}\quad (\text{B.474})$$

For the same reason as before, if  $i \neq j$  acting on the  $i$ -th site we find zero. Conversely, if  $i = j$  obviously we have:  $\{c_j, c_j\} = 2\sigma_j^- \sigma_j^- = 0$ .

**Exercise 11.2** As explained in the text, when mapping the Ising ring into a free-fermion model, we require periodic boundary conditions (PBC:  $c_{L+1} = c_1$ ) for an odd number  $N_F$  of fermions, and anti-periodic boundary conditions (ABC:  $c_{L+1} = -c_1$ ) for an even number  $N_F$  of fermions. Equivalently, from the definition of the Fourier transform in Eq. (11.28), we need  $e^{i \frac{2\pi}{L} k L} = \pm 1$  where the plus sign holds for  $N_F$  odd and the minus sign holds for  $N_F$  even.

We distinguish two cases, according to the parity of the number  $L$  of sites in the chain.

- CASE I: if  $L$  is even, we can label the sites with

$$j = -\frac{L}{2} + 1, \frac{L}{2} + 2, \dots, \frac{L}{2}, \quad (\text{e.g. } L = 10 \implies j = -4, -3, \dots, 5). \quad (\text{B.475})$$

There are now two sub-cases: for  $N_F$  odd, we can take

$$k = \frac{L}{2} + 1, -\frac{L}{2} + 2, \dots, \frac{L}{2} \implies \frac{2\pi}{L} k = -\pi + \frac{2\pi}{L}, -\pi + \frac{4\pi}{L}, \dots, \pi \quad (\text{B.476})$$

thus enforcing PBC:  $e^{i \frac{2\pi}{L} k L} = 1$  (e.g.  $L = 10 \implies \frac{2\pi}{L} k = -\frac{8}{10}\pi, -\frac{6}{10}\pi, \dots, \pi$ ). Conversely for  $N_F$  even, we can take

$$k = -\frac{L-1}{2}, -\frac{L-3}{2}, \dots, \frac{L-1}{2} \implies \frac{2\pi}{L} k = -\frac{\pi(L-1)}{L}, -\frac{\pi(L-3)}{L}, \dots, \frac{\pi(L-1)}{L} \quad (\text{B.477})$$

thus enforcing ABC:  $e^{i \frac{2\pi}{L} k L} = -1$  (e.g.  $L = 10 \implies \frac{2\pi}{L} k = -\frac{9}{10}\pi, -\frac{7}{10}\pi, \dots, \frac{9}{10}\pi$ ).

- CASE II: if  $L$  is odd, we can label the sites with

$$j = -\frac{L-1}{2}, \frac{L-3}{2}, \dots, \frac{L-1}{2}, \quad (\text{e.g. } L = 11 \implies j = -5, -4, \dots, 5). \quad (\text{B.478})$$

There are again two sub-cases: for  $N_F$  odd, we can take the same  $k$  values as in Eq. (B.477), which now enforce PBC:  $e^{i \frac{2\pi}{L} k L} = 1$  (e.g.  $L = 11 \implies \frac{2\pi}{L} k = -\frac{10}{11}\pi, -\frac{8}{11}\pi, \dots, \frac{10}{11}\pi$ ). Conversely for  $N_F$  even, we can take the same  $k$  values as in Eq. (B.476), which now enforce ABC:  $e^{i \frac{2\pi}{L} k L} = -1$  (e.g.  $L = 11 \implies \frac{2\pi}{L} k = -\frac{9}{11}\pi, -\frac{7}{11}\pi, \dots, \pi$ ).

Summarizing, we obtained that the choice of momenta as in Eq. (B.476) holds for  $L$  even and  $N_F$  odd, or for  $L$  odd and  $N_F$  even (which means  $L + N_F$  odd). On the other

hand, the choice of momenta as in Eq. (B.477) holds for both  $L$  and  $N_F$  even, or  $L$  and  $N_F$  odd (which means  $L + N_F$  even).

**Exercise 11.3** We wish to find a canonical transformation that diagonalizes the Hamiltonian in Eq. (11.43), mapping it into

$$H = \sum_q \varepsilon_q (\eta_q^\dagger \eta_q - \frac{1}{2}) . \quad (\text{B.479})$$

This can be done by employing a generalized Bogoliubov rotation of the form, following the method first proposed by Lieb *et al.* (1961):

$$\eta_q = \sum_j (g_{qj} c_j + h_{qj} c_j^\dagger) , \quad (\text{B.480})$$

where  $g_{qj}$  and  $h_{qj}$  can be seen as the entries of two  $L \times L$  matrices  $G$  and  $H$ , respectively. We stress that such matrices can be chosen to be real, such that Eq. (B.480) implies:  $\eta_q^\dagger = \sum_j (g_{qj} c_j^\dagger + h_{qj} c_j)$ . In a compact form, Eq. (B.480) can be rewritten as:  $\boldsymbol{\eta} = G \vec{c} + H \vec{c}^\dagger$ , where  $\boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_L]^T$  and  $\vec{c} = [c_1, c_2, \dots, c_L]^T$ .

Since the  $\eta_q^{(\dagger)}$  operators must have a fermionic character, they are requested to satisfy the anti-commutation relations  $\{\eta_q, \eta_p^\dagger\} = \delta_{qp}$  and  $\{\eta_q, \eta_p\} = \{\eta_q^\dagger, \eta_p^\dagger\} = 0$ . These translate into the following two constraints:

$$\sum_j (g_{qj} g_{pj} + h_{qj} h_{pj}) = \delta_{qp} \implies G G^T + H H^T = I_L , \quad (\text{B.481})$$

$$\sum_j (g_{qj} h_{pj} - g_{pj} h_{qj}) = 0 \implies G H^T - H G^T = 0 , \quad (\text{B.482})$$

where  $I_L$  is the  $L \times L$  identity matrix. Now if Eq. (B.479) holds, then the following condition must be satisfied:

$$\{\eta_q, H\} = \varepsilon_q \eta_q . \quad (\text{B.483})$$

Plugging in the definition of  $\eta_q$  in Eq. (B.480), we obtain

$$\sum_j (g_{qj} A_{ji} - h_{qj} B_{ji}) = \varepsilon_q g_{qi} \quad \text{and} \quad \sum_j (g_{qj} B_{ji} - h_{qj} A_{ji}) = \varepsilon_q h_{qi} . \quad (\text{B.484})$$

These equations can be written in a compact form, after defining the  $2L$  vectors  $\{\Phi_q\}_{q=1,\dots,L}$  and  $\{\Psi_q\}_{q=1,\dots,L}$ , with components:

$$[\Phi_q]_j = g_{qj} + h_{qj} , \quad [\Psi_q]_j = g_{qj} - h_{qj} , \quad (\text{B.485})$$

such that  $g_{qj} = \frac{1}{2}([\Phi_q]_j + [\Psi_q]_j)$  and  $h_{qj} = \frac{1}{2}([\Phi_q]_j - [\Psi_q]_j)$ . Thus, we finally obtain

$$\Phi_q (A - B) = \varepsilon_q \Psi_q , \quad (\text{B.486})$$

$$\Psi_q (A + B) = \varepsilon_q \Phi_q . \quad (\text{B.487})$$

which can be also cast into:

$$(A + B) (A - B) \Phi_q = \varepsilon_q^2 \Phi_q , \quad (\text{B.488})$$

$$\Psi_q (A + B) (A - B) = \varepsilon_q^2 \Psi_q . \quad (\text{B.489})$$

For  $\varepsilon_q \neq 0$ , it is then possible to solve either Eq. (B.488) or Eq. (B.489) for  $\Phi_q$  or  $\Psi_q$ , respectively. The other vector is thus obtained from the previous Eq. (B.486) or Eq. (B.487). For  $\varepsilon_q = 0$ , it is more convenient to find  $\Phi_q$  and  $\Psi_q$  directly from Eq. (B.486) and Eq. (B.487).

**Exercise 11.4** Let us start with the case of PBC. In order to contract a sequence of  $L$  transfer matrices  $\mathbb{E}_{O_s}^{[s]}$ , with  $s = 1, \dots, L$ , we can proceed recursively. Namely, we start from the leftmost transfer matrix  $\mathbb{E}_{O_1}^{[1]}$  and add to the right of it the second one  $\mathbb{E}_{O_2}^{[2]}$ . The brute force approach would consist in simply performing a row-by-column matrix multiplication the two  $\chi^2 \times \chi^2$  matrices, thus resulting in a number of  $\chi^6$  elementary operations. We can however simplify this operation by decomposing the building blocks of the final contracted tensor  $\mathbb{E}_{O_1}^{[1]} \mathbb{E}_{O_2}^{[2]}$  in elementary pieces, following the five-step scheme of Fig. B.21. Considering that typically  $\chi \gg d$ , the most expensive steps of this approach are the third and the fifth one, each requiring  $O(d\chi^5)$  operations. The contraction scheme then has to be repeated  $L - 1$  times, thus requiring in total  $O((L - 1)d\chi^5)$  operations.

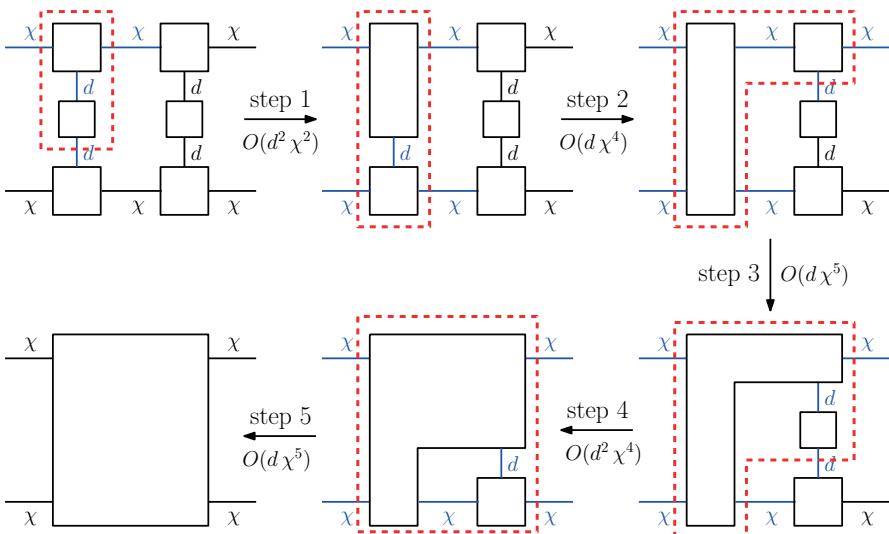


Fig. B.21 Sequence of the various contractions to be performed in order to multiply two nearby transfer matrices. The number of physical links associated to each leg is indicated in the figure. Each step consists in contracting the tensors grouped in the red dashed boxes; the corresponding number of operations is equal to the product of all the legs inside the boxes and going out of them (blue links).

For OBC, we can proceed exactly in the same way, starting either from the left-boundary (or even from the right-boundary) transfer matrix. The only difference is that the two legs on the left of the first transfer matrix are no longer present, and thus the full calculation requires a factor  $O(\chi^2)$  less operations than before.

**Exercise 11.5** We start from the proof of the equalities in Eq. (11.196). Using the definition (11.195) of the identity superoperator, and the superoperator representation of

observables we get:

$$(A \otimes \mathbb{I})|\mathbb{I}\rangle = \left( \sum_{i,j} A_{ij}|i\rangle\langle j| \otimes \sum_k |k\rangle\langle k| \right) \sum_\ell |\ell\rangle \otimes |\ell\rangle = \sum_{i,j,k,\ell} A_{ij}\delta_{j\ell}\delta_{k\ell}|i\rangle \otimes |k\rangle, \quad (\text{B.490})$$

$$(\mathbb{I} \otimes A^T)|\mathbb{I}\rangle = \left( \sum_i |i\rangle\langle i| \otimes \sum_{j,k} A_{kj}|j\rangle\langle k| \right) \sum_\ell |\ell\rangle \otimes |\ell\rangle = \sum_{i,j,k,\ell} A_{kj}\delta_{i\ell}\delta_{k\ell}|i\rangle \otimes |j\rangle, \quad (\text{B.491})$$

and, from the last step, it is clear that they both coincide with  $|A\rangle = \sum_{i,j} A_{ij}|i\rangle \otimes |j\rangle$ .

Equation (11.197) is a direct consequence of Eq. (11.196), indeed:

$$(A \otimes \mathbb{I})(B \otimes \mathbb{I})|\mathbb{I}\rangle = (A \otimes \mathbb{I})|B\rangle = (A \otimes \mathbb{I})(\mathbb{I} \otimes B^T)|\mathbb{I}\rangle = (\mathbb{I} \otimes B^T)|A\rangle, \quad (\text{B.492})$$

where, in the last passage, we used the fact that the identity matrix  $\mathbb{I}$  commutes with any other matrix, and thus  $(A \otimes \mathbb{I})(\mathbb{I} \otimes B^T) = (\mathbb{I} \otimes B^T)(A \otimes \mathbb{I})$ .

Finally, Eq. (11.198) follows from the previous equality:

$$|ABC\rangle = (\mathbb{I} \otimes C^T)|AB\rangle = (\mathbb{I} \otimes C^T)(A \otimes \mathbb{I})|B\rangle = (A \otimes C^T)|B\rangle, \quad (\text{B.493})$$

where, in the first passage, we used Eq. (11.197) substituting  $A \rightarrow AB$  and  $B \rightarrow C$ .

## B.12 Appendix A

**Exercise A.1** The linearity of the operators  $A$  and  $B$  implies that

$$[(A + B)^\dagger]_{ij} = (A + B)_{ji}^* = A_{ji}^* + B_{ji}^* = A_{ij}^\dagger + B_{ij}^\dagger. \quad (\text{B.494})$$

Since

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}, \quad (\text{B.495})$$

we have

$$(AB)_{ij}^\dagger = (AB)_{ji}^* = \sum_k A_{jk}^* B_{ki}^*. \quad (\text{B.496})$$

Furthermore, we obtain

$$(B^\dagger A^\dagger)_{ij} = \sum_k (B^\dagger)_{ik} (A^\dagger)_{kj} = \sum_k B_{ki}^* A_{jk}^* = \sum_k A_{jk}^* B_{ki}^*, \quad (\text{B.497})$$

and therefore comparison of Eqs. (B.496) and (B.497) proves that  $(AB)^\dagger = B^\dagger A^\dagger$ . Finally, we obtain

$$[(A^\dagger)^\dagger]_{ij} = (A^\dagger)_{ji}^* = [A_{ij}^*]^* = A_{ij}. \quad (\text{B.498})$$

**Exercise A.2** From Eq. (A.38) we find that the projector  $P$  has matrix elements

$$P_{ij} = \langle i|P|j\rangle = \sum_l \langle i|\alpha_l\rangle \langle \alpha_l|j\rangle, \quad (\text{B.499})$$

where  $\{|i\rangle\}$  is a basis for the Hilbert space. Therefore,

$$P_{ji}^* = \sum_l \langle j|\alpha_l\rangle^* \langle \alpha_l|i\rangle^* = \sum_l \langle \alpha_l|j\rangle \langle i|\alpha_l\rangle = P_{ij}. \quad (\text{B.500})$$

Hence, a projector is a Hermitian operator. If we exclude the trivial case in which  $P = I$ , a projector cannot be inverted, since it has zero eigenvalues, corresponding to the eigenvectors residing in the subspace orthogonal to the subspace onto which  $P$  projects. Therefore,  $\det P = 0$  and  $P$  cannot be inverted.

**Exercise A.3** We can immediately check that the Pauli matrices are Hermitian: since  $\sigma_i = (\sigma_i^T)^*$  ( $i = x, y, z$ ). By computing  $\sigma_i^{-1}$ , we can also check that  $\sigma_i^{-1} = (\sigma_i^T)^*$  and therefore the Pauli matrices are also unitary.

**Exercise A.4**

$$\sigma_x \otimes \sigma_y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}. \quad (\text{B.501})$$

$$I \otimes \sigma_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (\text{B.502})$$

**Exercise A.5** Let us call  $|\phi\rangle$  the vector obtained by application of the operator  $A$  to a generic vector  $|\psi\rangle$ ; that is,  $|\phi\rangle = A|\psi\rangle$ . All vectors are transformed by means of the matrix  $S^{-1}$  (see Eq. (A.71)), in particular  $|\psi'\rangle = S^{-1}|\psi\rangle$  and  $|\phi'\rangle = S^{-1}|\phi\rangle$ . Therefore,

$$S|\phi\rangle = SA|\psi\rangle = SAS^{-1}S|\psi\rangle, \quad (\text{B.503})$$

where we have used the relation  $S^{-1}S = I$ . Finally, we obtain

$$|\phi'\rangle = SAS^{-1}|\psi'\rangle = A'|\psi'\rangle, \quad (\text{B.504})$$

which implies the final result

$$A' = S^{-1}AS. \quad (\text{B.505})$$

**Exercise A.6** Assume that  $\{|\alpha_{i1}\rangle, |\alpha_{i2}\rangle, \dots, |\alpha_{ik}\rangle\}$  are eigenvectors of the linear operator  $A$  corresponding to the eigenvalue  $\alpha_i$  and that  $|\alpha_j\rangle$  is an eigenvector corresponding to the eigenvalue  $\alpha_j$ . Furthermore, assume that  $\alpha_i \neq \alpha_j$  and that the eigenvectors  $\{|\alpha_{il}\rangle\}$  and  $|\alpha_j\rangle$  are linearly dependent; that is,

$$|\alpha_j\rangle = \sum_{l=1}^k c_l |\alpha_{jl}\rangle. \quad (\text{B.506})$$

We then have

$$A|\alpha_j\rangle = \sum_{l=1}^k c_l A|\alpha_{il}\rangle = \sum_{l=1}^k c_l \alpha_i |\alpha_{il}\rangle = \alpha_i |\alpha_j\rangle. \quad (\text{B.507})$$

Since  $A|\alpha_j\rangle = \alpha_j|\alpha_j\rangle$ , we find  $\alpha_i = \alpha_j$ , which contradicts the hypothesis that the eigenvalues  $\alpha_i$  and  $\alpha_j$  are distinct.

**Exercise A.7** We have  $\sigma'_i = S^{-1}\sigma_i S$  ( $i = x, y, z$ ), where the matrix  $S$  relating the old basis vectors  $\{|0\rangle, |1\rangle\}$  to the new  $\{|+\rangle, |-\rangle\}$  is given by Eq. (A.97). Note that the matrix  $S$  is self-inverse, that is,  $S^{-1} = S$ . Therefore, we obtain

$$\sigma'_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \sigma_z. \quad (\text{B.508})$$

Similarly, we obtain  $\sigma'_y = -\sigma_y$  and  $\sigma'_z = \sigma_x$ .

**Exercise A.8** Since  $A^\dagger = A$  and  $B^\dagger = B$ , we obtain:

$$\{i[A, B]\}^\dagger = \{iAB - iBA\}^\dagger = -iB^\dagger A^\dagger + iA^\dagger B^\dagger = i[A, B]. \quad (\text{B.509})$$

**Exercise A.9** Let us define  $S \equiv \sqrt{M^T M}$ . The matrix  $S$  is by definition symmetric. Therefore, it is diagonalizable and we can write its spectral decomposition  $S = \sum_i s_i |i\rangle\langle i|$ . As  $S$  is a non-negative operator, we have  $s_i \geq 0$ . We now define  $|\psi_i\rangle \equiv M|i\rangle$ . We see that  $\langle\psi_i|\psi_i\rangle = \langle i|M^T M|i\rangle = s_i^2$ . For  $s_i \neq 0$ , we define  $|\alpha_i\rangle \equiv |\psi_i\rangle/s_i$ . The vectors  $|\alpha_i\rangle$  are normalized and orthogonal. We employ the Gram–Schmidt decomposition to complete the orthonormal basis  $\{|\alpha_i\rangle\}$ . Finally, let us define the operator  $O \equiv \sum_i |\alpha_i\rangle\langle i|$ . This operator is orthogonal since  $O^T O = \sum_{i,j} |j\rangle\langle j|\alpha_j\rangle\langle i| = \sum_i |i\rangle\langle i| = I$ . When  $s_i \neq 0$  we have  $OS|i\rangle = s_i O|i\rangle = s_i |\alpha_i\rangle = |\psi_i\rangle = M|i\rangle$ ; when  $s_i = 0$ ,  $OS|i\rangle = 0 = |\psi_i\rangle = M|i\rangle$ . Since the actions of the linear operators  $M$  and  $OS$  on the basis  $\{|i\rangle\}$  coincide, then  $M = OS$ .

## Bibliography

- Abbott, D., Davies, P. C. W., and Pati, A. K. (Eds.) (2008), Quantum aspects of life, Imperial College Press, London.
- Abrams, D. S. and Lloyd, S. (1997), Simulation of many-body Fermi systems on a universal quantum computer, *Phys. Rev. Lett.* **79**, 2586.
- Abrams, D. S. and Lloyd, S. (1999), Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors, *Phys. Rev. Lett.* **83**, 5162.
- Adesso, G., Bromley, T. R., and Cianciaruso, M. (2016), Measures and applications of quantum correlations, *J. Phys. A: Math. Theor.* **49**, 473001.
- Adesso, G. and Datta, A. (2010), Quantum versus classical correlations in Gaussian states, *Phys. Rev. Lett.* **105**, 030501.
- Affleck, I., Kennedy, T., Lieb, E. H., and Tasaki, H. (1987), Rigorous results on valence-bond ground states in antiferromagnets, *Phys. Rev. Lett.* **59**, 799.
- Agaian, S. S. and Klappenecker, A. (2002), Quantum computing and a unified approach to fast unitary transforms, arXiv:quant-ph/0201120, in Proc. SPIE 4667, Image Processing: Algorithms and Systems, p. 1, 2002.
- Agrawal, M., Kayal, N., and Saxena, N. (2004), PRIMES is in P, *Ann. Math.* **160**, 781.
- Aharanov, Y. and Bohm, D. (1959), Significance of electromagnetic potentials in the quantum theory, *Phys. Rev.* **115**, 485.
- Albash, T. and Lidar, D. A. (2018), Adiabatic quantum computation, *Rev. Mod. Phys.* **90**, 015002.
- Alber, G., Beth, T., Horodecki, M., Horodecki, P., Horodecki, R., Rötteler, M., Weinfurter, H., Werner, R., and Zeilinger, A. (2001), Quantum information – An introduction to theoretical concepts and experiments, Springer–Verlag.
- Alekseev, V. M. and Jacobson, M. V. (1981), Symbolic dynamics and hyperbolic dynamic systems, *Phys. Rep.* **75**, 287.
- Amico, L., Fazio, R., Osterloh, A., and Vedral, V. (2008), Entanglement in many-body systems, *Rev. Mod. Phys.* **80**, 517.
- Andresen, B. (2011), Current trends in finite-time thermodynamics, *Angew. Chem. Int. Ed.* **50**, 2690.

- Angelakis, D. G., Santos, M. F., and Bose, S. (2007), Photon-blockade-induced Mott transitions and  $XY$  spin models in coupled cavity arrays, *Phys. Rev. A* **76**, 031805(R).
- Aspect, A., Grangier, P., and Roger, G. (1981), Experimental tests of realistic local theories via Bell's theorem, *Phys. Rev. Lett.* **47**, 460.
- Aspuru-Guzik, A., Dutoi, A. D., Love, P. J., and Head-Gordon, M. (2005), Simulated quantum computation of molecular energies, *Science* **309**, 1704.
- Balazs, N. L. and Voros, A. (1989), The quantized baker's transformation, *Ann. Phys. (N.Y.)* **190**, 1.
- Bandyopadhyay, S. (2000), Qubit- and entanglement-assisted optimal entanglement concentration, *Phys. Rev. A* **62**, 032308.
- Barahona, F. (1982), On the computational complexity of Ising spin glass models, *J. Phys. A: Math. Gen.* **15**, 3241.
- Barenco, A. (1995), A universal two-bit gate for quantum computation, *Proc. R. Soc. Lond. A* **449**, 679.
- Barenco, A., Bennett, C. H., Cleve, R., DiVincenzo, D. P., Margolus, N., Shor, P., Sleator, T., Smolin, J. A., and Weinfurter, H. (1995), Elementary gates for quantum computation, *Phys. Rev. A* **52**, 3457.
- Barends, R. *et al.* (2016), Digitized adiabatic quantum computing with a superconducting qubit, *Nature* **534**, 222.
- Barnett, S. M. and Radmore, P. M. (1997), Methods in theoretical quantum optics, Clarendon Press, Oxford.
- Barnum, H., Fuchs, C. A., Jozsa, R., and Schumacher, B. (1996), General fidelity limit for quantum channels, *Phys. Rev. A* **54**, 4707.
- Barnum, H., Nielsen, M. A., and Schumacher, B. (1998), Information transmission through a noisy quantum channel, *Phys. Rev. A* **57**, 4153.
- Barouch, E., McCoy, B. M., and Dresden, M. (1970), Statistical mechanics of the  $XY$  model. I, *Phys. Rev. A* **2**, 1075.
- Barouch, E. and McCoy, B. M. (1971a), Statistical mechanics of the  $XY$  model. II. Spin correlation functions, *Phys. Rev. A* **3**, 786.
- Barouch, E. and McCoy, B. M. (1971b), Statistical mechanics of the  $XY$  model. III, *Phys. Rev. A* **3**, 2137.
- Barrett, M. D. *et al.* (2004), Deterministic quantum teleportation of atomic qubits, *Nature* **429**, 737.
- Bassi, A., Großardt, A., and Ulbricht, H. (2017), Gravitational decoherence, *Class. Quantum Grav.* **34**, 193002.
- Beauregard, S. (2003), Circuit for Shor's algorithm using  $2n+3$  qubits, *Quantum Computation and Information* **3**, 175.
- Beckman, B., Chari, A. N., Devabhaktuni, S., and Preskill, J. (1996), Efficient networks for quantum factoring, *Phys. Rev. A* **54**, 1034.
- Bekenstein, J. D. (1973), Black holes and entropy, *Phys. Rev. D* **7**, 2333.
- Bell, J. S. (1964), On the Einstein–Podolsky–Rosen paradox, *Physics* **1**, 195.

- Benenti, G., Casati, G., Montangero, S., and Shepelyansky, D. L. (2001), Efficient quantum computing of complex dynamics, *Phys. Rev. Lett.* **87**, 227901.
- Benenti, G., Casati, G., Montangero, S., and Shepelyansky, D. L. (2003), Dynamical localization simulated on a few-qubit quantum computer, *Phys. Rev. A* **67**, 052312.
- Benenti, G. and Strini, G. (2008), Quantum simulation of the single-particle Schrödinger equation, *Am. J. Phys.* **76**, 657.
- Benenti, G. and Strini, G. (2009a), Gaussian wave packets in phase space: The Fermi  $g_F$  function, *Am. J. Phys.* **77**, 546.
- Benenti, G. and Strini, G. (2009b), Simple representation of quantum process tomography, *Phys. Rev. A* **80**, 022318.
- Benenti, G., Siccaldi, S., and Strini, G. (2014a), Exotic states in the dynamical Casimir effect, *Eur. Phys. J. D* **68**, 139.
- Benenti, G., D'Arrigo, A., Siccaldi, S., and Strini, G. (2014b), Dynamical Casimir effect in quantum-information processing, *Phys. Rev. A* **90**, 052313.
- Benenti, G., Casati, G., Saito, K., and Whitney, R. S. (2017), Fundamental aspects of steady-state conversion of heat to work at the nanoscale, *Phys. Rep.* **694**, 1.
- Bengtsson, I. and Życzkowski, K. (2017), Geometry of quantum states. An introduction to quantum entanglement, Cambridge University Press, Cambridge, UK.
- Bennett, C. H. (1973), Logical reversibility of computation, *IBM J. Res. Dev.* **17**, 525.
- Bennett, C. H. (1982), The thermodynamics of computation – A review, *Int. J. Theor. Phys.* **21**, 905.
- Bennett, C. H. (1987), Demons, engines and the second law, *Sci. Am.* **257:5**, 108, November 1987.
- Bennett, C. H. (1992), Quantum cryptography using any two nonorthogonal states, *Phys. Rev. Lett.* **68**, 3121.
- Bennett, C. H. and Brassard, G. (1984), Quantum cryptography: Public key distribution and coin tossing, in Proc. of IEEE Int. Conf. on Computers, Systems and Signal Processing, p. 175, IEEE, New York, 1984.
- Bennett, C. H., Brassard, G., Crépau, C., Jozsa, R., Peres, A., and Wootters, W. K. (1993), Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels, *Phys. Rev. Lett.* **70**, 1895.
- Bennett, C. H., Brassard, G., and Ekert, A. K. (1991), Quantum cryptography, *Sci. Am.*, **267:4**, 50, October 1992.
- Bennett, C. H., Brassard, G., and Mermin, N. D. (1992), Quantum cryptography without Bell's theorem, *Phys. Rev. Lett.* **68**, 557.
- Bennett, C. H. and Wiesner, S. J. (1992), Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states, *Phys. Rev. Lett.* **69**, 2881.
- Bennett, C. H., Bernstein, H. J., Popescu, S., and Schumacher, B. (1996a), Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046.

- Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., and Wootters, W. K. (1996b), Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824.
- Bera, A., Das, T., Sadhukhan, D., Singha Roy, S., Sen (De), A., and Sen, U. (2018), Quantum discord and its allies: A review of recent progress, *Rep. Prog. Phys.* **81**, 024001.
- Bernstein, E. and Vazirani, U. (1997), Quantum complexity theory, *SIAM J. Comput.* **26**, 1411.
- Berry, M. V. (1984), Quantal phase-factor accompanying adiabatic changes, *Proc. Roy. Soc. A* **392**, 45.
- Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, A., Dillenschneider, R., and Lutz, E. (2012), Experimental verification of Landauer's principle linking information and thermodynamics, *Nature* **483**, 187.
- Biamonte, J. and Bergholm, V. (2017), Tensor networks in a nutshell, arXiv:1708.00006.
- Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., and Lloyd, S. (2017), Quantum machine learning, *Nature* **549**, 195.
- Biham, E., Brassard, G., Kenigsberg, D., and Mor, T. (2004), Quantum computing without entanglement, *Theor. Comput. Sci.* **320**, 15.
- Binder, F., Correa, L. A., Gogolin, C., Anders, J., and Adesso, G. (Eds.) (2018), Thermodynamics in the quantum regime. Fundamental aspects and new directions, Springer International Publishing.
- Blais, A., Huang, R.-S., Wallraff, A., Girvin, S. M., and Schoelkopf, R. J. (2004), Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, *Phys. Rev. A* **69**, 062320.
- Blatt, R. and Wineland, D. (2008), Entangled states of trapped atomic ions, *Nature* **453**, 1008.
- Bloch, I., Dalibard, J., and Zwerger, W. (2008), Many-body physics with ultracold gases, *Rev. Mod. Phys.* **80**, 885.
- Bohr, N. (1935), Can quantum-mechanical description of physical reality be considered complete?, *Phys. Rev.* **48**, 696.
- Boixo, S., Albash, T., Spedalieri, F. M., Chancellor, N., and Lidar, D. A. (2013), Experimental signature of programmable quantum annealing, *Nat. Commun.* **4**, 3067.
- Boixo, S., Rønnow, T. F., Isakov, S. V., Wang, Z., Wecker, D., Lidar, D. A., Martinis J. M., and Troyer, M. (2014), Evidence for quantum annealing with more than one hundred qubits, *Nat. Phys.* **10**, 218.
- Boixo, S., Isakov, S. V., Smelyanskiy, V. N., Babbush, R., Ding, N., Jiang, Z., Bremner, M. J., Martinis, J. M., and Neven, H. (2018), Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595.
- Born, M. and Fock, V. (1928), Beweis des Adiabatensatzes, *Z. Phys.* **51**, 165.
- Borrelli, M., Sabín, C., Adesso, G., Plastina, F., and Maniscalco, S. (2012), Dynamics of atom–atom correlations in the Fermi problem, *New J. Phys.* **14**, 103010.

- Boschi, D., Branca, S., De Martini, F., Hardy, L., and Popescu, S. (1998), Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein–Podolsky–Rosen channels, *Phys. Rev. Lett.* **80**, 1121.
- Bouwmeester, D., Pan, J.-W., Mattle, K., Eibl, M., Weinfurter, H., and Zeilinger, A. (1997), Experimental quantum teleportation, *Nature* **390**, 575.
- Bowden, C. M., Chen, G., Diao, Z., and Klappenecker, A. (2002), The universality of the quantum Fourier transform in forming the basis of quantum computing algorithms, *J. Math. Anal. Appl.* **274**, 69.
- Boyer, B., Brassard, G., Høyer, P., and Tapp, A. (1998), Tight bounds on quantum searching, *Fortschr. Phys.* **46**, 493.
- Braak, D. (2011), Integrability of the Rabi model, *Phys. Rev. Lett.* **107**, 100401.
- Braginsky, V. B. and Khalili, F. Ya. (1992), Quantum measurement, Cambridge University Press, Cambridge.
- Brassard, G., Braunstein, S. L., and Cleve, R. (1998), Teleportation as a quantum computation, *Physica D* **120**, 43.
- Brassard, G., Høyer, P., Mosca, M., and Tapp, A. (2002), Quantum amplitude amplification and estimation, arXiv:quant-ph/0005055, in *Contemp. Math.* **305**, AMS Special Session: Quantum Computation and Information, Lomonaco, S. J. and Brandt, H. E. (Eds.), American Mathematical Society, Providence, RI.
- Brassard, G., Horodecki, P., Mor, T. (2004), TelePOVM – A generalized quantum teleportation scheme, *IBM J. Res. Dev.* **48**, 87.
- Braunstein, S. L. and van Loock, P. (2005), Quantum information with continuous variables, *Rev. Mod. Phys.* **77**, 513.
- Breuer, H.-P., Kappler, B., and Petruccione, F. (1999), Stochastic wave-function method for non-Markovian quantum master equations, *Phys. Rev. A* **59**, 1633.
- Breuer, H.-P., Laine, E.-M., and Piilo, J. (2009), Measure for the degree of non-Markovian behavior of quantum processes in open systems, *Phys. Rev. Lett.* **103**, 210401.
- Breuer, H.-P., Laine, E.-M., Piilo, J., and Vacchini, B. (2016), Non-Markovian dynamics in open quantum systems, *Rev. Mod. Phys.* **88**, 021002.
- Breuer, H.-P. and Petruccione, F. (2002), The theory of open quantum systems, Oxford University Press.
- Briegel, H.-J., Dür, W., Cirac, J. I., and Zoller, P. (1998), Quantum repeaters: The role of imperfect local operations in quantum communication, *Phys. Rev. Lett.* **81**, 5932.
- Brun, T. A. (2002), A simple model of quantum trajectories, *Am. J. Phys.* **70**, 719.
- Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V., and Wehner, S. (2014), Bell nonlocality, *Rev. Mod. Phys.* **86**, 419.
- Bruß, D., DiVincenzo, D. P., Ekert, A., Fuchs, C. A., Macchiavello, C., and Smolin, J. A. (1998), Optimal universal and state-dependent quantum cloning, *Phys. Rev. A* **57**, 2368.

- Bruß, D. and Lütkenhaus, N. (2000), Quantum key distribution: from principles to practicalities, arXiv:quant-ph/9901061, in *Appl. Algebra Eng. Commun. Comput. (AAECC)* **10**, 383.
- Bruß, D. (2002), Characterizing entanglement, *J. Math. Phys.* **43**, 4237.
- Buluta, I. and Nori, F. (2009), Quantum Simulators, *Science* **326**, 108.
- Burkard, G. and Loss, D. (2002), Spin qubits in solid-state structures, *Europhys. News* **33/5**, 166.
- Bužek, V. and Hillery, M. (1996), Quantum copying: Beyond the no-cloning theorem, *Phys. Rev. A* **54**, 1844; Universal optimal cloning of qubits and quantum registers, arXiv:quant-ph/9801009.
- Bužek, V., Hillery, M., and Werner, R. F. (1999), Optimal manipulations with qubits: Universal-NOT gate, *Phys. Rev. A* **60**, R2626; Universal-NOT gate, *J. Mod. Opt.* **47**, 211 (2000).
- Caldeira, A. O. and Leggett, A. J. (1983), Quantum tunnelling in a dissipative system, *Ann. Phys. (N.Y.)* **153**, 445.
- Calderbank, A. R. and Shor, P. W. (1996), Good quantum error-correcting codes exist, *Phys. Rev. A* **54**, 1098.
- Campisi, M., Hänggi, P., and Talkner, P. (2011), Colloquium: Quantum fluctuation relations: Foundations and applications, *Rev. Mod. Phys.* **83**, 1653.
- Carleo, G. and Troyer, M. (2017), Solving the quantum many-body problem with artificial neural networks, *Science* **335**, 602.
- Carlo, G. G., Benenti, G., Casati, G., and Mejía-Monasterio, C. (2004), Simulating noisy quantum protocols with quantum trajectories, *Phys. Rev. A* **69**, 062317.
- Carmichael, H. J. (1993), An open systems approach to quantum optics, Lecture Notes in Physics, Springer, Berlin.
- Carrasquilla, J. and Melko, R. G. (2017), Machine learning phases of matter, *Nat. Phys.* **13**, 431.
- Caruso, F., Giovannetti, V., Lupo, C., and Mancini, S. (2014), Quantum channels and memory effects, *Rev. Mod. Phys.* **86**, 1203.
- Cerf, N. J., Adami, C., and Kwiat, P. G. (1998), Optical simulation of quantum logic, *Phys. Rev. A* **57**, R1477.
- Chang, S.-J. and Shi, K.-J. (1986), Evolution and exact eigenstates of a resonant quantum system, *Phys. Rev. A* **34**, 7.
- Chiaverini, J. et al. (2004), Realization of quantum error correction, *Nature* **432**, 602.
- Chin, C., Grimm, R., Julienne, P., and Tiesinga, E. (2010), Feshbach resonances in ultracold gases, *Rev. Mod. Phys.* **82**, 1225.
- Church, A. (1936), An unsolvable problem of elementary number theory, *Am. J. Math.* **58**, 345.
- Cirac, J. I. and Verstraete, F. (2009), Renormalization and tensor product states in spin chains and lattices, *J. Phys. A* **42**, 504004.

- Cirac, J. I. and Zoller, P. (1995), Quantum computations with cold trapped ions, *Phys. Rev. Lett.* **74**, 4091.
- Cirac, J. I. and Zoller, P. (2004), New frontiers in quantum information with atoms and ions, *Phys. Today*, March 2004, p. 38.
- Clarke, J. and Wilhelm, F. K. (2008), Superconducting quantum bits, *Nature* **453**, 1031.
- Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A., (1969), Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880.
- Cleve, R., Ekert, A., Macchiavello, C., and Mosca, M. (1998), Quantum algorithms revisited, *Proc. R. Soc. Lond. A* **454**, 339.
- Coffman, V., Kundu, J., and Wootters, W. (2000), Distributed entanglement, *Phys. Rev. A* **61**, 052306.
- Cohen-Tannoudji, C., Diu, B., and Laloë, F. (1977), Quantum mechanics, Vols. I and II, Hermann, Paris.
- Coppersmith, D. (1994), An approximate Fourier transform useful in quantum factoring, IBM Research Report No. RC 19642, arXiv:quant-ph/0201067.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2001), Introduction to algorithms, MIT Press, Cambridge, Massachusetts.
- Cover, T. M. and Thomas, J. A. (1991), Elements of information theory, John Wiley & Sons, New York.
- Cubitt, T. S., Ruskai, M. B., and Smith, G. (2008), The structure of degradable quantum channels, *J. Math. Phys.* **49**, 102104.
- Dakić, B., Vedral, V., and Brukner, C. (2010), Necessary and sufficient condition for nonzero quantum discord, *Phys. Rev. Lett.* **105**, 190502.
- D'Arrigo, A., Benenti, G., and Falci, G. (2007), Quantum capacity of dephasing channels with memory, *New J. Phys.* **9**, 310.
- D'Arrigo, A., Benenti, G., and Falci, G. (2008), Memory effects in quantum information transmission across a Hamiltonian dephasing channel, *Eur. Phys. J. Special Topics* **160**, 83.
- D'Arrigo, A., Lo Franco, R., Benenti, G., Paladino, E., and Falci, G. (2014), Recovering entanglement by local operations, *Ann. Phys.* **350**, 211.
- Datta, A. (2008), Studies on the role of entanglement in mixed-state quantum computation, arXiv:0807.4490.
- Davidov, A. S. (1982), Biology and quantum mechanics, Pergamon Press, Oxford.
- De Chiara, G., Rizzi, M., Rossini, D., and Montangero, S. (2008), Density matrix renormalization group for dummies, *J. Comput. Theor. Nanosci.* **5**, 1.
- De Martini, F., Bužek, V., Sciarrino, F., and Sias, C. (2002), Experimental realization of the quantum universal-NOT gate, *Nature* **419**, 815.
- Deng, D.-L., Li, X., and Das Sarma, S. (2017), Quantum entanglement in neural network states, *Phys. Rev. X* **7**, 021021.
- Deutsch, D. (1985), Quantum theory, the Church–Turing principle and the universal quantum computer, *Proc. R. Soc. Lond. A* **400**, 97.

- Deutsch, D. (1989), Quantum computational networks, *Proc. R. Soc. Lond. A* **425**, 73.
- Deutsch, D., Barenco, A., and Ekert, A. (1995), Universality in quantum computation, *Proc. R. Soc. Lond. A* **449**, 669.
- Deutsch, D. and Jozsa, R. (1992), Rapid solution of problems by quantum computation, *Proc. R. Soc. Lond. A* **439**, 553.
- Devetak, I. (2005), The private classical capacity and quantum capacity of a quantum channel, *IEEE Trans. Inf. Theory* **51**, 44.
- Devetak, I. and Shor, P. W. (2005), The capacity of a quantum channel for simultaneous transmission of classical and quantum information, *Comm. Math. Phys.* **256**, 287.
- Devoret, M. H. and Schoelkopf, R. J. (2913), Superconducting circuits for quantum information: An outlook, *Science* **339**, 1169.
- De Vos, A. (2010), Reversible computing: Fundamentals, quantum computing, and applications, Wiley-VCH.
- DiCarlo, L. *et al.* (2009), Demonstration of two-qubit algorithms with a superconducting quantum processor, *Nature* **460**, 240.
- Dieks, D. (1982), Communication by EPR devices, *Phys. Lett. A* **92**, 271.
- Diósi, L., Kulacsy, K., Lukács, B., and Rácz, A. (1996), Thermodynamic length, time, speed, and optimum path to minimize entropy production, *J. Chem. Phys.* **105**, 11220.
- Dittrich, T., Hänggi, P., Ingold, G.-L., Kramer, B., Schön, S., and Zwerger, W. (1998), Quantum transport and dissipation, Wiley-VCH, Weinheim.
- DiVincenzo, D. P. (1995), Two-bit gates are universal for quantum computation, *Phys. Rev. A* **51**, 1015.
- DiVincenzo, D. P. (2000a), The physical implementation of quantum computation, *Fortschr. Phys.* **48**, 771.
- DiVincenzo, D. P., Bacon, D., Kempe, J., Burkard, G., and Whaley, K. B. (2000b), Universal quantum computation with the exchange interaction, *Nature* **408**, 339.
- Dodonov, V. V., Kurmyshev, E. V., and Man'ko, V. I. (1988), Correlated coherent states, in "Classical and quantum effects in electrodynamics", Komar, A. A. (Ed.), Nova Science Publishers.
- Donoghue, J. F., Ivanov, M. M., and Shkerin, A. (2017), EPFL lectures on general relativity as a quantum field theory, arXiv:1702.00319.
- Draper, T. G. (2000), Addition on a quantum computer, arXiv:quant-ph/0008033.
- Duan, L.-M., Giedke, G., Cirac, J. I., and Zoller, P. (2000), Inseparability criterion for continuous variable systems, *Phys. Rev. Lett.* **84**, 2722.
- Duan, L.-M., Cirac, J. I., and Zoller, P. (2001), Geometric manipulation of trapped ions for quantum computation, *Science* **292**, 1695.
- Dubi, Y. and Di Ventra, M. (2011), Colloquium: Heat flow and thermoelectricity in atomic and molecular junctions, *Rev. Mod Phys.* **83**, 131.

- Dunjko, V. and Briegel, H. J. (2018), Machine learning & artificial intelligence in the quantum domain: A review of recent progress, *Rep. Prog. Phys.* **81**, 074001.
- Dür, W., Vidal, G., and Cirac, J. I. (2000), Three qubits can be entangled in two inequivalent ways, *Phys. Rev. A* **62**, 062314.
- Dziarmaga, J. (2010), Dynamics of a quantum phase transition and relaxation to a steady state, *Adv. Phys.* **59**, 1063.
- Eibenberger, S., Gerlich, S., Arndt, M., Mayor, M., and Tüxen, J. (2013), Matter-wave interference of particles selected from a molecular library with masses exceeding 10 000 amu, *Phys. Chem. Chem. Phys.* **15**, 14696.
- Einstein, A. (1905), Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt, *Ann. Physik* **17**, 132; english translation (*On a heuristic point of view about the creation and conversion of light*) in Ter Haar, D. (1967), *The old quantum theory* (Pergamon Press, Oxford, 1967).
- Einstein, A. (1917), Zur quantentheorie des strahlung, *Phys. Z.* **18**, 121; english translation (*On the quantum theory of radiation*), in Ter Haar, D. (1967), *The old quantum theory* (Pergamon Press, Oxford, 1967).
- Einstein, A., Podolsky, B., and Rosen, N. (1935), Can quantum-mechanical description of physical reality be considered complete?, *Phys. Rev.* **47**, 777.
- Eisert, J., Cramer, M., and Plenio, M. B. (2010), Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277.
- Ekert, A. K. (1991), Quantum cryptography based on Bell's theorem, *Phys. Rev. Lett.* **67**, 661.
- Ekert, A., Hayden, P. M., and Inamori, H. (2001), Basic concepts in quantum computation, arXiv:quant-ph/0011013, in “Coherent atomic matter waves”, Les Houches Summer Schools, Session LXXII, Kaiser, R., Westbrook, C., and David, F. (Eds.), Springer–Verlag.
- Ekert, A. and Jozsa, R. (1996), Quantum computation and Shor's factoring algorithm, *Rev. Mod. Phys.* **68**, 733.
- Ekert, A. and Jozsa, R. (1998), Quantum algorithms: Entanglement enhanced information processing, *Phil. Trans. R. Soc. Lond. A* **356**, 1769.
- Elzerman, J. M., Kouwenhoven, L. P., and Vandersypen, L. M. K. (2006), Electron spin qubits in quantum dots, in Proceedings of the “E. Fermi” Varenna School on “Quantum computers, algorithms and chaos”, Casati, G., Shepelyansky, D. L., Zoller, P., and Benenti G. (Eds.), IOS Press and SIF, Bologna.
- Emerson, J., Lloyd, S., Poulin, D., and Cory, D. G. (2004), Estimation of the local density of states on a quantum computer, *Phys. Rev. A* **69**, 050305(R).
- Emerson, J., Weinstein, Y. S., Lloyd, S., and Cory, D. G. (2002), Fidelity decay as an efficient indicator of quantum chaos, *Phys. Rev. Lett.* **89**, 284102.
- Essler, F. H. L., Frahm, H., Göhmann, F., Klümper, A., and Korepin, V. E. (2005), The one-dimensional Hubbard model, Cambridge University Press, Cambridge.
- Euler, L. (1736), Solutio problematis ad geometriam situs pertinentis, *Comment. Acad. Sci. U. Petrop.* **8**, 128.

- Facchi, P. and Pascazio, S. (2003), Three different manifestations of the quantum Zeno effect, arXiv:quant-ph/0303161, in “Irreversible quantum dynamics”, Benatti, F. and Floreanini, R. (Eds.), Lecture Notes in Physics, Vol. 622, p. 141, Springer–Verlag.
- Facchi, P., Lidar, D. A., and Pascazio, S. (2004), Unification of dynamical decoupling and the quantum Zeno effect, *Phys. Rev. A* **69**, 032314.
- Fahri, E., Goldstone, J., Gutmann, S., and Sipser, M. (2000), Quantum computation by adiabatic evolution, arXiv:quant-ph/0001106.
- Fahri, E., Goldstone, J., Gutmann, S., Lapan, J., Lundgren, A., and Preda, D. (2001), A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, *Science* **292**, 472.
- Fannes, M., Nachtergaelie, B., and Werner R. F. (1992), Finitely correlated states on quantum spin chains, *Commun. Math. Phys.* **144**, 443.
- Fano, U. (1957), Description of states in quantum mechanics by density matrix and operator techniques, *Rev. Mod. Phys.* **29**, 74.
- Fano, U. (1983), Pairs of two-level systems, *Rev. Mod. Phys.* **55**, 855.
- Fazio, R. and van der Zant, H. (2001), Quantum phase transitions and vortex dynamics in superconducting networks, *Phys. Rep.* **355**, 235.
- Fermi, E. (1930), L’interpretazione del principio di causalità nella meccanica quantistica, *Nuovo Cimento* **7**, 361. An English translation of this paper is available upon request to: [giuliano.strini@mi.infn.it](mailto:giuliano.strini@mi.infn.it)
- Fermi, E. (1932), Quantum theory of radiation, *Rev. Mod. Phys.* **4**, 87.
- Ferraro, A., Aolita, L., Cavalcanti, D., Cucchietti, F. M., and Acín, A. (2010), Almost all quantum states have nonclassical correlations, *Phys. Rev. A* **81**, 052318.
- Feynman, R. P. (1982), Simulating Physics with Computers, *Int. J. Theor. Phys.* **21**, 467.
- Fijany, A. and Williams, C. P. (1998), Quantum wavelet transforms: fast algorithms and complete circuits, arXiv:quant-ph/9809004, in Lecture Notes in Computer Science, No. 1509, p. 10, Springer–Verlag.
- Fisher, M. P. A., Weichman, P. B., Grinstein, G., and Fisher, D. S. (1989), Boson localization and the superfluid-insulator transition, *Phys. Rev. B* **40**, 546.
- Fitzpatrick, M., Sundaresan, N. M., Li, A. C. Y., Koch, J., and Houck, A. A. (2017), Observation of a dissipative phase transition in a one-dimensional circuit QED lattice, *Phys. Rev. X* **7**, 011016.
- Fleming, G. R., Huelga, S. F., and Plenio, M. B. (2011), Focus on quantum effects and noise in biomolecules, *New J. Phys.* **13**, 115002.
- Foong, S. K. and Kanno, S. (1994), Proof of Page’s conjecture on the average entropy of a subsystem, *Phys. Rev. Lett.* **72**, 1148.
- Ford, J. (1983), How random is a coin toss?, *Phys. Today*, April 1983, p. 40.
- Forn-Díaz, P., García-Ripoll, J. J., Peropadre, B., Orgiazzi, J.-L-, Yurtalan, M. A., Belyansky, R., Wilson, C. M., and Lupascu, A. (2017), Ultrastrong coupling of a single artificial atom to an electromagnetic continuum in the nonperturbative regime, *Nature Phys.* **13**, 39.

- Franco, M. I., Turin, L., Mershin, A., and Skoulakis, E. M. C. (2011), Molecular vibration-sensing component in *Drosophila melanogaster* olfaction, *Proc. Natl. Acad. Sci.* **108**, 3797.
- Fredkin, E. and Toffoli, T. (1982), Conservative logic, *Int. J. Theor. Phys.* **21**, 219.
- Fuchs, C. A. and Caves, C. M. (1994), Ensemble-dependent bounds for accessible information in quantum mechanics, *Phys. Rev. Lett.* **73**, 3047.
- Fürer, M. (2007), Faster Integer Multiplication, in Proceedings of the 39th ACM Symposium on Theory of Computing, p. 57.
- Gaioli, F. H., Garcia Alvarez, E. T., and Guevara, J. (1997), Quantum Brownian motion, *Int. J. Theor. Phys.* **36**, 2167.
- Galindo, A. and Martin-Delgado, M. A. (2002), Information and computation: Classical and quantum aspects, *Rev. Mod. Phys.* **74**, 347.
- Gambetta, J. M., Chow, J. M., and Steffen, M. (2017), Building logical qubits in a superconducting quantum computing system, *npj Quantum Information* **3**, 2.
- Gardiner, S. A., Cirac, J. I., and Zoller, P. (1997), Quantum chaos in an ion trap: The delta-kicked harmonic oscillator, *Phys. Rev. Lett.* **79**, 4790.
- Gardiner, C. W. and Zoller, P. (2000), Quantum noise (2nd Ed.), Springer–Verlag.
- Garey, M. R. and Johnson, D. S. (1979), Computers and intractability: A guide to the theory of NP-completeness, W. H. Freeman, New York. MIT Press, Cambridge, Massachusetts.
- Gelbwaser-Klimovsky, D., Niedenzu, W., and Kurizki, G. (2015), Thermodynamics of quantum systems under dynamical control, *At. Mol. Opt. Phys.* **64**, 329.
- Georgeot, B. and Shepelyansky, D. L. (2001a), Exponential gain in quantum computing of quantum chaos and localization, *Phys. Rev. Lett.* **86**, 2890.
- Georgeot, B. and Shepelyansky, D. L. (2001b), Stable quantum computation of unstable classical chaos, *Phys. Rev. Lett.* **85**, 5393; see also *Phys. Rev. Lett.* **88**, 219802 (2002).
- Georgescu, I. M., Ashhab, S., and Nori, F. (2014), Quantum simulation, *Rev. Mod. Phys.* **86**, 153.
- Gerlach, W. and Stern, O. (1922), Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld, *Z. Phys.* **9**, 349.
- Ghirardi, G. C. (2013), Entanglement, nonlocality, superluminal signaling and cloning, *Z. Phys.* **9**, 349. in “Advances in Quantum Mechanics”, P. Bracken (Ed.), InTech.
- Giazotto, F., Heikkilä, T. T., Luukanen, A., Savin, A. M., and Pekola, J. P. (2006), Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications, *Rev. Mod. Phys.* **78**, 217.
- Giorda, P. and Paris, M. G. A. (2010), Gaussian quantum discord, *Phys. Rev. Lett.* **105**, 020503.
- Gisin, N. and Massar, S. (1997), Optimal quantum cloning machines, *Phys. Rev. Lett.* **79**, 2153.
- Gisin, N., Ribordy, G., Tittel, W., and Zbinden, H. (2002), Quantum cryptography, *Rev. Mod. Phys.* **74**, 145.

- Goldman, I. I., Krivchenkov, V. D., Kogan, V. I., and Galitskii, V. M. (1960), Problems in quantum mechanics, Infosearch Limited, London.
- Golub, G. H. and Van Loan, C. F. (2013), Matrix computations (4th Ed.), The Johns Hopkins University Press, Baltimore.
- Goold, J., Huber, M., Riera, A., del Rio, L., and Skrzypczyk, P. (2016), The role of quantum information in thermodynamics – A topical review, *J. Phys. A* **49**, 143001.
- Gorbachev, V. N. and Trubillo, A. I. (2000), Quantum teleportation of EPR pair by three-particle entanglement, *J. Exp. Theor. Phys.* **91**, 894.
- Gorini, V., Kossakowski, A., and Sudarshan, E. C. G. (1976), Completely positive dynamical semigroups of N-level systems, *J. Math. Phys.* **17**, 821.
- Gossett, P. (1998), Quantum carry–save arithmetic, arXiv:quant-ph/9808061.
- Gottesman, D. (1997), Stabilizer Codes and Quantum Error Correction, arXiv:quant-ph/9705052.
- Gottesman, D. (2000), An introduction to quantum error correction, arXiv:quant-ph/0004072.
- Gottesman, D. and Chuang, I. L. (1999), Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390.
- Grassi, A. M. and Strini, G. (1999), Some extensions of the Deutsch problem, in *Mysteries, puzzles, and paradoxes in quantum mechanics*, Bonifacio, R. (Ed.), AIP Conf. Proc. 461, p. 291.
- Gray, R. M. (1990), Entropy and information theory, Springer–Verlag.
- Greentree, A. D., Tahan, C., Cole, J. H., and Hollenberg, L. C. L. (2006), Quantum phase transitions of light, *Nat. Phys.* **2**, 856.
- Greif, D., Uehlinger, T., Jotzu, F., Tarruell, L., and Esslinger, T. (2013), Short-Range Quantum Magnetism of Ultracold Fermions in an Optical Lattice, *Science* **340**, 1307.
- Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. W., and Bloch, I. (2002), Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms, *Nature* **415**, 39.
- Griffiths, D. J. (2005), Introduction to quantum mechanics (2nd Ed.), Prentice Hall, NY, USA.
- Grover, L. K. (1996), A fast quantum mechanical algorithm for database search, arXiv:quant-ph/9605043, in Proc. of the 28th Annual ACM Symposium on the Theory of Computing, p. 212, ACM Press, New York.
- Grover, L. K., (1997), Quantum Mechanics helps in searching for a needle in a haystack, *Phys. Rev. Lett.* **79**, 325.
- Häffner, H. *et al.* (2005), Scalable multiparticle entanglement of trapped ions, *Nature* **438**, 643.
- Häffner, H., Roos, C. F., and Blatt, R. (2008), Quantum computing with trapped ions, *Phys. Rep.* **469**, 155.

- Hagley, E., Maître, X., Nogues, G., Wunderlich, C., Brune, M., Raimond, J. M., and Haroche, S. (1997), Generation of Einstein–Podolsky–Rosen pairs of atoms, *Phys. Rev. Lett.* **79**, 1.
- Hameroff, S. R. (1987), Ultimate computing: Biomolecular consciousness and nanotechnology, Elsevier Science Publishers.
- Hanson, R., Kouwenhoven, L. P., Petta, J. R., Tarucha, S., and Vandersypen, L. M. K. (2007), Spins in few-electron quantum dots, *Rev. Mod. Phys.* **79**, 1217.
- Hardy, L. (2017), Proposal to use Humans to switch settings in a Bell experiment, arXiv:1705.04620.
- Haroche, S. and Raimond, J.-M. (2006), Exploring the quantum: Atoms, cavities, and photons, Oxford University Press.
- Hartmann, M. J., Brandā o, F. G. S. L., and Plenio, M. B. (2006), Strongly interacting polaritons in coupled arrays of cavities, *Nat. Phys.* **2**, 849.
- Hastings, M. B. (2007), An area law for one dimensional quantum systems, *J. Stat. Mech.*, P08024.
- Hastings, M. B. (2009), Superadditivity of communication capacity using entangled inputs, *Nature Phys.* **5**, 255.
- Hawking, S. W. (1975), Particle creation by black holes, *Commun. Math. Phys.* **43**, 199.
- Hayden, P., Leung, D. W., and Winter, A. (2006), Aspects of generic entanglement, *Commun. Math. Phys.* **265**, 95.
- Hegerfeldt, G. C. (1994), Causality problems for Fermi’s two-atom system, *Phys. Rev. Lett.* **72**, 596.
- Henderson, L. and Vedral, V. (2001), Classical, quantum and total correlations *J. Phys. A: Math. Gen.* **34**, 6899.
- Hioe, F. T. and Eberly, J. H. (1981),  $N$ -level coherence vector and higher conservation laws in quantum optics and quantum mechanics, *Phys. Rev. Lett.* **47**, 838.
- Holevo, A. S. (1973), Bounds for the quantity of information transmitted by a quantum communication channel, *Prob. Inf. Transm.* **9**, 177.
- Holevo, A. S. (1998), The capacity of the quantum channel with general signal states, *IEEE Trans. Inf. Theory* **44**, 269.
- Holevo, A. S. (2011), Probabilistic and statistical aspects of quantum theory, Edizioni della Normale, Pisa, Italy.
- Horodecki, M., Horodecki, P., and Horodecki, R. (1996), Separability of mixed states: Necessary and sufficient conditions, *Phys. Lett. A* **223**, 1.
- Horodecki, P. (1997), Separability criterion and inseparable mixed states with positive partial transposition, *Phys. Lett. A* **232**, 333.
- Horodecki, M., Horodecki, P., and Horodecki, R. (1998), Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature?, *Phys. Rev. Lett.* **80**, 5239 (1998).
- Horodecki, R., Horodecki, P., Horodecki, M., and Horodecki, K. (2009), Quantum entanglement, *Rev. Mod. Phys.* **81**, 865.

- Houck, A. A., Türeci, H. E., and Koch, J. (2012), On-chip quantum simulation with superconducting circuits, *Nat. Phys.* **8**, 292.
- Howl, R., Hackermüller, L., Bruschi, D. E., and Fuentes, I. (2018), Gravity in the quantum lab, *Adv. Phys.* **3**, 1.
- Huang, K. (1987), Statistical mechanics (2nd Ed.), John Wiley & Sons, New York.
- Hubbard, J. (1963), Electron correlations in narrow energy bands, *Proc. Roy. Soc. A* **276**, 237.
- Hudson, R. L. (1974), When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.* **6**, 249.
- Imamōglu, A., Schmidt, H., Woods, G., and Deutsch, M. (1997), Strongly interacting photons in a nonlinear cavity, *Phys. Rev. Lett.* **79**, 1467.
- Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. W., and Zoller, P. (1998), Cold bosonic atoms in optical lattices, *Phys. Rev. Lett.* **81**, 3108.
- Johnson, M. W. *et al.* (2011), Quantum annealing with manufactured spins, *Nature* **473**, 194.
- Joshi, S. D. (2018), Space QUEST mission proposal: Experimentally testing decoherence due to gravity, *New. J. Phys.* **20**, 063016.
- Jouguet, P., Kunz-Jacques, S., Leverrier, A., Grangier, P., and Diamanti, E. (2013), Experimental demonstration of long-distance continuous-variable quantum key distribution, *Nature Photon.* **7**, 378.
- Jozsa, R. (1998), Quantum algorithms and the Fourier transform, *Proc. R. Soc. Lond. A* **454**, 323.
- Jozsa, R. and Linden, N. (2003), On the role of entanglement in quantum computational speed-up, *Proc. R. Soc. Lond. A* **459**, 2011.
- Jun, Y., Gavrilov, M., and Bechhoefer, J. (2014), High-precision test of Landauer's principle in a feedback trap, *Phys. Rev. Lett.* **113**, 190601.
- Kane, B. (1998), A silicon-based nuclear spin quantum computer, *Nature* **393**, 133.
- Kattnig, D. R. and Hore, P. J. (2017), The sensitivity of a radical pair compass magnetoreceptor can be significantly amplified by radical scavengers, *Sci. Rep.* **7**, 11640.
- Kelly, J. *et al.* (2015), State preservation by repetitive error detection in a superconducting quantum circuit, *Nature* **519**, 66.
- Kibble, T. W. B. (1976), Topology of cosmic domains and strings, *J. Phys. A: Math. Gen.* **9**, 1387.
- Kieu, T. D. (2004), A reformulation of Hilbert's tenth problem through quantum mechanics, *Proc. R. Soc. Lond. A* **460**, 1535.
- Kitaev, A. Yu. (1995), Quantum measurements and the Abelian stabilizer problem, arXiv:quant-ph/9511026.
- Kitaev, A. Yu. (2001), Unpaired Majorana fermions in quantum wires, *Phys. Usp.* **44**, 131.
- Klappenecker, A. and Rötteler, M. M. (2001), On the irresistible efficiency of signal processing methods in quantum computing, arXiv:quant-ph/0111039, in Proceed-

- ings of the First International Workshop on Spectral Techniques and Logic Design (SPECLOG 2000), p. 483, 2000.
- Kleinjung, T. et al. (2010), Factorization of a 768-Bit RSA Modulus, *Advances in Cryptology – CRYPTO 2010*, Lecture notes in Computer Science **6223**, 333.
- Kloeffel, C. And Loss, D. (2013), Prospects for spin-based quantum computing in quantum dots, *Annu. Rev. Condens. Matter Phys.* **4**, 51.
- Knill, E. and Laflamme, R. (1997), Theory of quantum error-correcting codes, *Phys. Rev. A* **55**, 900.
- Knill, E., Laflamme, R., and Viola, L. (2000), Theory of quantum error correction for general noise, *Phys. Rev. Lett.* **84**, 2525.
- Knill, E., Laflamme, R., and Milburn, G. J. (2001), A scheme for efficient quantum computation with linear optics, *Nature* **409**, 46.
- Knill, E., Laflamme, R., Ashikhmin, A., Barnum, H., Viola, L., and Zurek, W. H. (2002), Introduction to quantum error correction, arXiv:quant-ph/0207170, *Los Alamos Science* **27**, 188.
- Knuth, D. E. (1997–1998), The art of computer programming, vol. I: Fundamental algorithms; vol. II: Seminumerical algorithms; vol. III: Sorting and searching, Addison-Wesley, Reading, Massachusetts.
- Kok, P., Munro, W. J., Nemoto, K., Ralph, T. C., Dowling, J. P., and Milburn, G. J. (2007), Linear optical quantum computing with photonic qubits, *Rev. Mod. Phys.* **79**, 135.
- Konopik, M., Friedenberger, A., Kiesel, N., and Lutz, E. (2018), Nonequilibrium information erasure below  $kT \ln 2$ , arXiv:1806.01034.
- Kornfeld, I. P., Fomin, S. V., and Sinai, Ya. G. (1982), Ergodic theory, Springer-Verlag.
- Koski, J. V., Maisi, V. F., Pekola, J. P., and Averin, D. V. (2014), Experimental realization of a Szilard engine with a single electron, *Proc. Natl. Acad. Sci.* **111**, 13786.
- Kosloff, R. (2013), Quantum thermodynamics: A dynamical viewpoint, *Entropy* **15**, 2100.
- Kraus, K. (1983), States, effects, and operations, Springer-Verlag, Berlin.
- Krenn, M., Malik, M., Scheidl, T., Ursin, R., and Zeilinger, A. (2016), Quantum communication with photons, in “Optics in Our Time”, Al-Amri, M. D., El-Gomati, M. M., and Zubairy, M. S. (Eds.), Springer International Publishing.
- Kretschmann, D. and Werner, R. F. (2005), Quantum channels with memory, *Phys. Rev. A* **72**, 062323.
- Krisnanda, T., Marletto, C., Vedral, V., Paternostro, M., and Paterek, T. (2017), Probing quantum features of photosynthetic organisms, arXiv:1711.06485.
- Laflamme, R., Miquel, C., Paz, J. P., and Zurek, W. H. (1996), Perfect quantum error correcting code, *Phys. Rev. Lett.* **77**, 198.
- Lambert, N., Chen, Y.-N., Cheng, Y.-C., Li, C.-M., Chen, G.-Y., and Nori, F. (2013), Quantum biology, *Nature Phys.* **9**, 10.

- Landauer, R. (1961), Irreversibility and heat generation in the computing process, *IBM J. Res. Dev.* **5**, 183.
- Lang, S. (1996), Linear algebra, Springer–Verlag.
- Lanting, T. *et al.* (2014), Entanglement in a quantum annealing processor, *Phys. Rev. X* **4**, 021041.
- Latora, V. and Baranger, M. (1999), Kolmogorov–Sinai entropy rate versus physical entropy, *Phys. Rev. Lett.* **82**, 520.
- Lavor, C., Manssur, L. R. U., and Portugal, R., (2003), Shor’s algorithm for factoring large integers, arXiv:quant-ph/0303175.
- Lee, J.-S., Chung, Y., Kim, J., and Lee, S. (1999), A practical method of constructing quantum combinatorial logic circuits, arXiv:quant-ph/9911053.
- Le Hur, K., Henriet, L., Petrescu, A., Plekhanov, K., Roux, G., and Schiró, M. (2016), Many-body quantum electrodynamics networks: Non-equilibrium condensed matter physics with light, *C. R. Physique* **17**, 808.
- Leibfried, D., Blatt, R., Monroe, C., and Wineland, D. (2003a), Quantum dynamics of single trapped ions, *Rev. Mod. Phys.* **75**, 281.
- Leibfried, D. *et al.* (2003b), Experimental demonstration of a robust, high-fidelity geometric two ion-qubit phase gate, *Nature* **422**, 412.
- Leibfried, D. *et al.* (2005), Creation of a six-atom ‘Schrödinger cat’ state, *Nature* **438**, 638.
- Leijnse, M. and Flensberg, K. (2012), Introduction to topological superconductivity and Majorana fermions, *Semicond. Sci. Technol.* **27**, 124003.
- Leonhardt, U. (1997), Measuring the quantum state of light, Cambridge University Press, Cambridge.
- Levitin, L. B., Toffoli, T., and Walton, Z. (2002), Operation time of quantum gates, arXiv:quant-ph/0210076, in Proceedings of 6th International Conference on Quantum Communication, Measurement, and Computing (QCMC’2002).
- Lewenstein, M., Kraus, B., Cirac, J. I., and Horodecki, P. (2000), Optimization of entanglement witnesses, *Phys. Rev. A* **62**, 052310.
- Lewenstein, M., Sanpera, A., Ahufinger, V., Damski, B., Sen(De), A., and Sen, U. (2007), Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond, *Adv. Phys.* **56**, 243.
- Lidar, D. A., Bacon, D., Kempe, J., and Whaley, K. B. (2000), Protecting quantum information against exchange errors using decoherence free states, *Phys. Rev. A* **61**, 052307.
- Lidar, D. A. and Whaley, B. (2003), Decoherence-free subspaces and subsystems, arXiv:quant-ph/0301032, in “Irreversible quantum dynamics”, Benatti, F. and Floreanini, R. (Eds.), Lecture Notes in Physics, Vol. 622, p. 83, Springer–Verlag.
- Lieb, E., Schultz, T., and Mattis, D. (1961), Two soluble models of an antiferromagnetic chain, *Ann. Phys.* **16**, 407.
- Lindblad, G. (1976), On the generators of quantum dynamical semigroups, *Commun. Math. Phys.* **48**, 119.

- Lloyd, S. (1995), Almost any quantum logic gate is universal, *Phys. Rev. Lett.* **75**, 346.
- Lloyd, S. (1996), Universal quantum simulators, *Science* **273**, 1073.
- Lomonaco, S. J. (2000), Shor's quantum factoring algorithm, arXiv:quant-ph/0010034.
- Lomonaco, S., J. (2001), A talk on quantum cryptography, or how Alice outwits Eve, arXiv:quant-ph/0102016.
- López-Suárez, M, Neri, I., and Gammaiton, L. (2016), Sub- $k_B T$  micro-electromechanical irreversible logic gate, *Nat. Commun.* **7**, 12068.
- Loss, D. and DiVincenzo, D. P. (1998), Quantum computation with quantum dots, *Phys. Rev. A* **57**, 120.
- Lubkin, E. (1978), Entropy of an  $n$ -system from its correlation with a  $k$ -reservoir, *J. Math. Phys.* **19**, 1028.
- Ma, X. S. et al. (2012), Quantum teleportation over 143 kilometres using active feed-forward, *Nature* **489**, 269.
- Macchiavello, C. and Palma, G. M. (2002), Entanglement-enhanced information transmission over a quantum channel with correlated noise, *Phys. Rev. A* **65**, 050301(R).
- Margolus, N. and Levitin, L. B. (1997), The maximum speed of dynamical evolution *Physica D* **120**, 188.
- Martin, J. and Vennin, V. (2017), Obstructions to Bell CMB experiments, *Phys. Rev. D* **96**, 063501.
- Maruyama, K., Nori, F., and Vedral, V. (2009), The physics of Maxwell's demon and information, *Rev. Mod. Phys.* **81**, 1.
- Matiyasevich, Yu. V. (1993), Hilbert's tenth problem, MIT Press, Cambridge, Massachusetts.
- McCoy, B. M., Barouch, E., and Abraham, D. B. (1971), Statistical mechanics of the XY model. IV. Time-dependent spin-correlation functions, *Phys. Rev. A* **4**, 2331.
- McCulloch, I. P. (2007), From density-matrix renormalization group to matrix product states, *J. Stat. Mech.*, P10014.
- Merli, P. G., Missiroli, G. F., and Pozzi, G. (1974), Electron interferometry with the Elmiskop 101 electron microscope, *J. Phys. E: Sci. Instrum.* **7**, 729.
- Mertens, S. (2000), cond-mat/0012185, in Computational complexity for physicists, *Computing in Science and Engineering* **4**, 31.
- Merzbacher, E. (1997), Quantum mechanics, John Wiley & Sons, New York.
- Meyer, K. R., Hall, G. H., and Offin, D. (2009), Introduction to hamiltonian dynamical systems and the  $N$ -body problem (2nd Ed.), Springer, New York.
- Meystre, P. and Sargent III, M. (2007), Elements of quantum optics, 4th Ed., Springer-Verlag, Berlin, 2007.
- Millen, J. and Xuereb, A. (2016), Perspective on quantum thermodynamics, *New J. Phys.* **18**, 011002.

- Miquel, C., Paz, J. P., and Perazzo, R. (1996), Factoring in a dissipative quantum computer, *Phys. Rev. A* **54**, 2605.
- Miquel, C., Paz, J. P., Saraceno, M., Knill, E., Laflamme, R., and Negrevergne, C. (2002), Interpretation of tomography and spectroscopy as dual forms of quantum computation, *Nature* **418**, 59.
- Modi, K., Brodutch, A., Cable, H., Paterek, T., and Vedral, V. (2012), The classical-quantum boundary for correlations: Discord and related measures, *Rev. Mod. Phys.* **84**, 1655.
- Monasson, R., Zecchina, R., Kirkpatrick, S., Selman, B., and Troyansky, L. (1999), Determining computational complexity from characteristic ‘phase transitions’, *Nature* **400**, 133.
- Monroe, C. (2002), Quantum information processing with atoms and photons, *Nature* **416**, 238.
- Monz, T., Schindler, P., Barreiro, J. T., Chwalla, M., Nigg, D., Coish, W. A., Harlander, M., Hänsel, W., Hennrich, M., and Blatt, R. (2011), 14-qubit entanglement: Creation and coherence, *Phys. Rev. Lett.* **106**, 130506.
- Monz, T., Nigg, D., Martinez, E. A., Brandl, M. F., Schindler, P., Rines, R., Wang, S. X., Chuang, I. L., and Blatt, R. (2016), *Science*, **351**, 1068.
- Movassagh, R. and Shor, P. (2016), Supercritical entanglement in local systems: Counterexample to the area law for quantum matter, *Proc. Natl. Acad. Sci.* **113**, 13278.
- Muhonen, J. T., Meschke, M., and Pekola, J. P. (2012), Micrometre-scale refrigerators, *Rep. Prog. Phys.* **75**, 046501.
- Myers, C. R. and Laflamme, R. (2006), Linear optics quantum computation: An overview, arXiv:quant-ph/0512104, in Proceedings of the “E. Fermi” Varenna School on “Quantum computers, algorithms and chaos”, Casati, G., Shepelyansky, D. L., Zoller, P., and Benenti G. (Eds.), IOS Press and SIF, Bologna.
- Nagaosa, N. (1999), Quantum field theory in condensed matter physics, Springer, Germany.
- Nakamura, Y., Pashkin, Yu. A., and Tsai, J. S. (1999), Coherent control of macroscopic quantum states in a single-Cooper-pair box, *Nature* **398**, 786.
- Namiki, M., Pascazio, S., and Nakazato, H. (1997), Decoherence and quantum measurements, World Scientific, Singapore.
- Nandkishore, R. and Huse, D. A. (2015), Many body localization and thermalization in quantum statistical mechanics, *Annu. Rev. Condens. Matter Phys.* **6**, 15.
- Nation, P. D., Johansson, J. R., Bencowe, M. P., and Nori, F. (2012), Stimulating uncertainty: Amplifying the quantum vacuum with superconducting circuits, *Rev. Mod. Phys.* **84**, 1.
- Nielsen, M. A. and Chuang, I. L. (2000), Quantum computation and quantum information, Cambridge University Press, Cambridge.
- Noh, C. and Angelakis, D. G. (2016), Quantum simulations and many-body physics with light, *Rep. Prog. Phys.* **80**, 016401.

- Ollivier, H. and Zurek, W. H. (2001), Quantum discord: A measure of the quantumness of correlations, *Phys. Rev. Lett.* **88**, 017901.
- Ortiz, G., Gubernatis, J. E., Knill, E., and Laflamme, R. (2001), Quantum algorithms for fermionic simulations *Phys. Rev. A* **64**, 022319; erratum *ibid.* **65**, 029902 (2002).
- Orus, R. (2014), A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117.
- Osborne, T. and Nielsen, M. (2002), Entanglement in a simple quantum phase transition, *Phys. Rev. A* **66**, 032110.
- Osterloh, A., Amico, L., Falci, G., and Fazio, R. (2002), Scaling of entanglement close to a quantum phase transition, *Nature* **416**, 608.
- Ott, E. (2002), Chaos in dynamical systems (2nd. Ed.), Cambridge University Press, Cambridge.
- Pachos, J., Zanardi, P., and Rasetti, M. (1999), Non-Abelian Berry connections for quantum computation, *Phys. Rev. A* **61**, 010305(R).
- Page, D. N. (1993), Average entropy of a subsystem, *Phys. Rev. Lett.* **71**, 1291.
- Papadimitriou, C. H. (1994), Computational complexity, Addison-Wesley, Reading, Massachusetts.
- Pathria, R. K. and Beale, P. D. (2011), Statistical Mechanics (3rd Ed.), Academic Press, Elsevier.
- Peres, A. (1993), Quantum theory: Concepts and methods, Kluwer Academic, Dordrecht.
- Peres, A. (1996), Separability criterion for density matrices, *Phys. Rev. Lett.* **77**, 1413.
- Petta, J. R., Johnson, A. C., Taylor, J. M., Laird, E. A., Yacoby, A., Lukin, M. D., Marcus, C. M., Hanson, M. P., and Gossard, A. C. (2005), Coherent manipulation of coupled electron spins in semiconductor quantum dots, *Science* **309**, 2180.
- Pfeuty, P. (1970), The one-dimensional Ising model with a transverse field, *Ann. Phys.* **57**, 79.
- Planck, M. (1901), Über das gesetz der energieverteilung im normalspektrum, *Ann.Phys.* **4**, 553; english translation (*On the theory of the energy distribution law of the normal spectrum*), in Ter Haar, D. (1967), *The old quantum theory* (Pergamon Press, Oxford, 1967).
- Plenio, M. B. and Knight, P. L. (1998), The quantum-jump approach to dissipative dynamics in quantum optics, *Rev. Mod. Phys.* **70**, 101.
- Plenio, M. B. and Virmani, S. (2007), An introduction to entanglement measures, *Quant. Inf. Comp.* **7**, 1.
- Preskill, J. (1998a), Lecture notes on quantum information and computation, available at: <http://theory.caltech.edu/people/preskill/>.
- Preskill, J. (1998b), Fault-tolerant quantum computation, arXiv:quant-ph/9712048, in “Introduction to quantum computation and information”, Lo, H.-K., Popescu, S., and Spiller, T. (Eds.), World Scientific, Singapore.

- Preskill, J. (1999), Battling decoherence: The fault-tolerant quantum computer, *Phys. Today*, June 1999, p. 24.
- Prokof'ev, N. V. and Stamp, P. C. E. (2000), Theory of the spin bath, *Rep. Prog. Phys.* **63**, 669.
- Raimond, J. M., Brune, M., and Haroche, S. (2001), Colloquium: Manipulating quantum entanglement with atoms and photons in a cavity, *Rev. Mod. Phys.* **73**, 565.
- Raussendorf, R. and Briegel, H. J. (2001), A one-way quantum computer, *Phys. Rev. Lett.* **86**, 5188.
- Reck, M., Zeilinger, A., Bernstein, H. J., and Bertani, P. (1994), Experimental realization of any discrete unitary operator, *Phys. Rev. Lett.* **73**, 58.
- Resta, R. (1994), Macroscopic polarization in crystalline dielectrics: the geometric phase approach, *Rev. Mod. Phys.* **66**, 899.
- Reu, J.-G. *et al.* (2017), Ground-to-satellite quantum teleportation, *Nature* **549**, 70.
- Rideout, D. *et al.* (2012), Fundamental quantum optics experiments conceivable with satellites-reaching relativistic distances and velocities, *Class. Quantum Grav.* **29**, 224011.
- Riebe, M. *et al.* (2004), Deterministic quantum teleportation with atoms, *Nature* **429**, 734.
- Rivas, Á., Huelga, S. F., and Plenio, M. B. (2010), Entanglement and non-Markovianity of quantum evolutions, *Phys. Rev. Lett.* **105**, 050403.
- Rivas, Á., Huelga, S. F., and Plenio, M. B. (2014), Quantum non-Markovianity: characterization, quantification and detection, *Rep. Prog. Phys.* **77**, 094001.
- Robertson, H. P. (1929), The uncertainty principle, *Phys. Rev.* **34**, 163.
- Rocci, A. (2013), On first attempts to reconcile quantum principles with gravity, *J. Phys.: Conf. Ser.* **470**, 012004.
- Roland, J. and Cerf, N. J. (2002), Quantum search by local adiabatic evolution, *Phys. Rev. A* **65**, 042308.
- Rønnow, T. F., Wang, Z., Job, J., Boixo, S., Isakov, S. V., Wecker, D., Martinis, J. M., Lidar, D. A., and Troyer, M. (2014), Defining and detecting quantum speedup, *Science* **25**, 420.
- Rovelli, C. (2004), Quantum gravity, Cambridge University Press, Cambridge.
- Rovelli, C. and Vidotto, F. (2014), Covariant loop quantum gravity: An elementary introduction to quantum gravity and spinfoam theory, Cambridge University Press, Cambridge.
- Sabín, C., del Rey, M., García-Ripoll, J., and León, J. (2011), Fermi problem with artificial atoms in circuit QED, *Phys. Rev. Lett.* **107**, 150402.
- Sachdev, S. (2011), Quantum phase transitions (2nd Ed.), Cambridge University Press, Cambridge.
- Sakurai, J. J. (1994), Modern quantum mechanics (revised ed.), Addison-Wesley, Reading, Massachusetts.

- Sánchez-Ruiz, J. (1995), Simple proof of Page's conjecture on the average entropy of a subsystem, *Phys. Rev. E* **52**, 5653.
- Santoro, G. E. and Tosatti, E. (2006), Optimization using quantum mechanics: quantum annealing through adiabatic evolution, *J. Phys. A: Math. Gen.* **39**, R393.
- Saraceno, M. (1990), Classical structures in the quantized Baker transformation, *Ann. Phys. (N.Y.)* **199**, 37.
- Scarani, V., Ziman, M., Štelmachovič, P., Gisin, N., and Bužek, V. (2002), Thermalizing quantum machines: dissipation and entanglement, *Phys. Rev. Lett.* **88**, 097905.
- Schack, R. (1998), Using a quantum computer to investigate quantum chaos. *Phys. Rev. A* **57**, 1634.
- Schleich, W. P. (2001), Quantum optics in phase space, Wiley-VCH Verlag, Berlin.
- Schlienz, J. and Mahler, G. (1995), Description of entanglement, *Phys. Rev. A* **52**, 4396.
- Schmidhuber, J. (2015), Deep learning in neural networks: An overview, *Neural Netw.* **61**, 85.
- Schmidt-Kaler, F. et al. (2003), Realization of the Cirac-Zoller controlled-NOT quantum gate, *Nature* **422**, 408.
- Schneider, Ch., Porras, D., and Schaetz, T. (2012), Experimental quantum simulations of many-body physics with trapped ions, *Rep. Prog. Phys.* **75**, 024401.
- Schoelkopf, R. J. and Girvin, S. M. (2008), Wiring up quantum systems, *Nature* **451**, 664.
- Schollwöck, U. (2005), The density-matrix renormalization group, *Rev. Mod. Phys.* **77**, 259.
- Schollwöck, U. (2011), The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96.
- Schönhage, A. and Strassen, V. (1971), Schnelle Multiplikation grosser Zahlen, *Computing* **7**, 281.
- Schrödinger, E. (1944), What is life? The physical aspect of the living cell, Cambridge University Press, Cambridge.
- Schuch, N., Wolf, M. M., Verstraete, F., and Cirac, J. I. (2008), Entropy scaling and simulability by matrix product states, *Phys. Rev. Lett.* **100**, 030504.
- Schumacher, B. (1995), Quantum coding, *Phys. Rev. A* **51**, 2738.
- Schumacher, B. (1996), Sending entanglement through noisy quantum channels, *Phys. Rev. A* **54**, 2614.
- Schumacher, B. (1998), Lecture notes on quantum information theory. **SOURCE?**
- Schumacher, B. and Nielsen, M. A. (1996), Quantum data processing and error correction, *Phys. Rev. A* **54**, 2629.
- Schumacher, B. and Westmoreland, M. D. (1997), Sending classical information via noisy quantum channels, *Phys. Rev. A* **56**, 131.

- Schuman, C. D., Potok, T. E., Patton, R. M., Birdwell, D., Dean, M. E., Rose, G. S., and Plank, J. S. (2017), A survey of neuromorphic computing and neural networks in hardware, arXiv:1705.06963.
- Schwartz, J. T. (1980), Fast probabilistic algorithms for verification of polynomial identities, *Journal of the ACM* **27**, 701.
- Scully, M. O. and Zubairy, M. S. (1997), Quantum optics, Cambridge University Press, Cambridge.
- Selman, B. and Kirkpatrick, S. (1996), Critical behavior in the computational cost of satisfiability testing, *Artif. Intell.* **81**, 273.
- Sen, S. (1996), Average entropy of a quantum subsystem, *Phys. Rev. Lett.* **77**, 1.
- Serafini, A., Illuminati, F., and De Siena, S. (2004), Symplectic invariants, entropic measures and correlations of Gaussian states, *J. Phys. B* **37**, L21.
- Shannon, C. E. (1948), A mathematical theory of communication, *Bell System Tech. J.* **27**, 379; 623.
- Shor, P. W. (1994), Algorithms for quantum computation: Discrete logarithm and factoring, in Proc. of the 35th Annual Symposium on the Foundations of Computer Science, Goldwasser, S. (Ed.), p. 124, IEEE Computer Society Press, Los Alamitos, CA.
- Shor, P. W. (1995), Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493.
- Shor, P. W. (1997), Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Sci. Statist. Comput.* **26**, 1484.
- Siegelman, H. T. (1995), Computation beyond the Turing limit, *Science* **268**, 545.
- Silvi, P., Tschirsich, F., Gerster, M., Jünemann, J., Jaschke, D., Rizzi, M., and Montangero, S. (2017), The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems, arXiv:1710.03733.
- Simon, R., Mukunda, N., and Dutta, B. (1994), Quantum-noise matrix for multi-mode systems:  $U(n)$  invariance, squeezing, and normal forms, *Phys. Rev. A* **49**, 1567.
- Smith, W. D. (2006), Three counterexamples refuting Kieu's plan for "quantum adiabatic hypercomputation"; and some uncomputable quantum mechanical tasks, *Applied Mathematics and Computation* **178**, 545.
- Song, G. and Klappenecker, A. (2003), Optimal realizations of controlled unitary gates, *Quantum Inf. Comput.* **3**, 139.
- Sothmann, B., Sánchez, R., and Jordan, A. N. (2015), Thermoelectric energy harvesting with quantum dots, *Nanotechnology* **26**, 032001.
- Steane, A. M. (1996a), Error correcting codes in quantum theory, *Phys. Rev. Lett.* **77**, 793.
- Steane, A. M. (1996b), Multiple particle interference and quantum error correction, *Proc. R. Soc. Lond. A* **452**, 2551.
- Steane, A. M. (2006), A tutorial on quantum error correction, in Proceedings of the "E. Fermi" Varenna School on "Quantum computers, algorithms and chaos",

- Casati, G., Shepelyansky, D. L., Zoller, P., and Benenti G. (Eds.), IOS Press and SIF, Bologna.
- Stone, J. V. (2013), Bayes' Rule: A Tutorial Introduction to Bayesian Analysis, Sebtel Press, England.
- Strini, G., and Pizzi, R. (2009), A proposal to measure the Rabi oscillations in the retinal rod cells, *Eur. Phys. J. D* **54**, 723.
- Summhammer, J. (2006), Quantum cooperation of two insects, arXiv:quant-ph/0503136.
- Suzuki, S., Inoue, J., and Chakrabarti, B. K. (2013), Quantum Ising phases and transitions in transverse Ising models, Lecture Notes in Physics (2nd Ed.), Springer, Berlin.
- Svozil, K., Levitin, L. B., Toffoli, T., and Walton, Z. (2005), Maximum speed of quantum gate operation, *Int. J. Theor. Phys.*, **44**, 965.
- Syropoulos, A. (2008), Hypercomputation: Computing beyond the Church-Turing barrier, Springer Series: Monographs in Computer Science.
- Sørensen, A. and Mølmer, K. (1999), Spin–spin interaction and spin squeezing in an optical lattice, *Phys. Rev. Lett.* **83**, 2274.
- Teichmann, P. (2012), Adiabatic logic: Future trend and system level perspective, Springer Series in Advanced Microelectronics, Vol. 34.
- Terhal, B. M. (2000), Bell inequalities and the separability criterion, *Phys. Lett. A* **271**, 319.
- Terhal, B. M. and DiVincenzo, D. P. (2000), Problem of equilibration and the computation of correlation functions on a quantum computer, *Phys. Rev. A* **61**, 022301.
- Thiemann, T. (2007), Modern canonical quantum general relativity, Cambridge University Press, Cambridge.
- 't Hooft, G. (2017), Free will in the theory of everything, arXiv:1709.02874.
- Ticozzi, F. and Ferrante, A. (2006), Dynamical decoupling in quantum control: A system theoretic approach, *Syst. Control Lett.* **55**, 578.
- Tinkham, M. (1996), Introduction to superconductivity, McGraw-Hill, New York.
- Toda, M., Kubo, R., and Saitô, N. (1983), Statistical physics I, Springer-Verlag.
- Tonomura, A., Endo, J., Matsuda, T., Kawasaki, T., and Ezawa, H. (1989), Demonstration of single-electron buildup of an interference pattern, *Am. J. Phys.* **57**, 117.
- Trabesinger, A. *et al.* (2012), Nature Physics Insight - Quantum simulation, *Nat. Phys.* **8**, No. 4 (pp. 263–299).
- Tucci, R. R. (1999), A rudimentary quantum compiler (2nd Ed.), arXiv:quant-ph/9902062.
- Turchette, Q. A., Wood, C. S., King, B. E., Myatt, C. J., Leibfried, D., Itano, W. M., Monroe, C., and Wineland, D. J. (1998), Deterministic entanglement of two trapped ions, *Phys. Rev. Lett.* **81**, 3631.

- Turing, A. (1936), On computable numbers, with an application to the Entscheidungsproblem, *Proc. Lond. Math. Soc. (2)* **42**, 230; correction *ibid.*, **43**, 544 (1937).
- Ursin, R., Tiefenbacher F., Schmitt-Manderbach, T., Weier, H., Scheidl, T., Lindenthal, M., Blauensteiner, B., Jennewein, T., Perdigues, J., Trojek, P. *et al.* (2007), Entanglement-based quantum communication over 144 km, *Nat. Phys.* **3**, 481.
- Vedral, V., Barenco, A., and Ekert, A. (1996), Quantum networks for elementary arithmetic operations, *Phys. Rev. A* **54**, 147.
- Verstraete, F. and Cirac, J. I. (2004a), Renormalization algorithms for quantum-many body systems in two and higher dimensions, arXiv:cond-mat/0407066.
- Verstraete, F., García-Ripoll, J. J., and Cirac, J. I. (2004b), Matrix product density operators: Simulation of finite- $T$  and dissipative systems, *Phys. Rev. Lett.* **93**, 207204.
- Verstraete, F., Porras, D., and Cirac, J. I. (2004c), Density matrix renormalization group and periodic boundary conditions: A quantum information perspective, *Phys. Rev. Lett.* **93**, 227205.
- Verstraete, F. and Cirac, J. I. (2006), Matrix product states represent ground states faithfully, *Phys. Rev. B*, **73**, 094423.
- Verstraete, F., Murg, V., and Cirac, J. I. (2008), Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143.
- Vidal, G. (2003), Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* **91**, 147902.
- Vidal, G. (2004), Efficient simulation of one-dimensional quantum many-body systems, *Phys. Rev. Lett.* **93**, 040502.
- Vidal, G. (2008), Class of quantum many-body states that can be efficiently simulated, *Phys. Rev. Lett.* **101**, 110501.
- Vidal, G., Latorre, J. I., Rico, E., and Kitaev, A. (2003), Entanglement in quantum critical phenomena, *Phys. Rev. Lett.* **90**, 227902.
- Vinjanampathy, S. and Anders, J. (2016), Quantum thermodynamics, *Contemp. Phys.* **57**, 545.
- Viola, L. and Lloyd, S. (1998), Dynamical suppression of decoherence in two-state quantum systems, *Phys. Rev. A* **58**, 2733.
- Viola, L., Knill, E., and Lloyd, S. (1999), Dynamical decoupling of open quantum systems, *Phys. Rev. Lett.* **82**, 2417.
- Viola, L., Knill, E., and Laflamme, R. (2001), Constructing qubits in physical systems, *J. Phys. A* **34**, 7067.
- Wallraff, A., Schuster, D. I., Blais, A., Frunzio, L., Huang, R.-S., Majer, J., Kumar, S., Girvin, S. M., and Schoelkopf, R. J. (2004), Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics, *Nature* **431**, 162.

- Walther, H., Varcoe, B. T. H., Englert, B.-G., and Becker, T. (2006), Cavity quantum electrodynamics, *Rep. Prog. Phys.* **69**, 1325.
- Weedbrook, C., Pirandola, S., García-Patron, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H., and Lloyd, S. (2012), Gaussian quantum information, *Rev. Mod. Phys.* **84**, 621.
- Weiss, U. (2012), Quantum dissipative systems (4th Ed.), World Scientific, Singapore.
- Welsh, D. (1997), Codes and Cryptography, Oxford University Press.
- Wenzler, J.-S., Dunn, T., Toffoli, T., and Mohanty, P. (2014), A nanomechanical Fredkin gate, *Nano Lett.* **14**, 89.
- White, S. R. (1992), Density matrix formulation for quantum renormalization groups, *Phys. Rev. Lett.* **69**, 2863.
- White, S. R. (1993), Density-matrix algorithms for quantum renormalization groups, *Phys. Rev. B* **48**, 10345.
- Wiesner, S. (1996), Simulation of many-body quantum systems by a quantum computer, arXiv:quant-ph/9603028.
- Wilczek, F. and Zee, A. (1984), Appearance of gauge structure in simple dynamical systems, *Phys. Rev. Lett.* **52**, 2111.
- Wilde, M. M. (2013), Quantum information theory, Cambridge University Press, Cambridge.
- Wootters, W. K. (1998), Entanglement of formation of an arbitrary state of two qubits, *Phys. Rev. Lett.* **80**, 2245.
- Wootters, W. K. and Zurek, W. H. (1982), A single quantum cannot be cloned, *Nature* **299**, 802.
- Xiao, D., Chang, M.-C., and Niu, Q. (2010), Berry phase effects on electronic properties, *Rev. Mod. Phys.* **82**, 1959.
- Yin, J., Ren, J. G., Lu, H., Cao, Y., Yong, H. L., Wu, Y. P., Liu, C., Liao, S. K., Zhou, F., Jiang, Y., et al. (2012), Quantum teleportation and entanglement distribution over 100-kilometre free-space channels, *Nature* **488**, 185.
- Yin, J. et al. (2017), Satellite-based entanglement distribution over 1200 kilometers, *Science* **356**, 1140.
- Yin, Z.-q. and Li, T. (2017), Bringing quantum mechanics to life: from Schrödinger's cat to Schrödinger's microbe, *Contemp. Phys.* **58**, 119.
- Yoshihara, F., Fuse, T., Ashhab, S., Kakuyanagi, K., Saito, S., and Semba, K. (2017), Superconducting qubit–oscillator circuit beyond the ultrastrong-coupling regime, *Nature Phys.* **13**, 44.
- You, J. Q. and Nori, F. (2005), Superconducting circuits and quantum information, *Phys. Today*, November 2005, p. 42.
- Zachos, C. K., Fairlie, D. B., and Curtright, T. L. (Eds.) (2005), Quantum mechanics in phase space, World Scientific, Singapore.
- Zalka, C. (1998), Efficient simulation of quantum systems by quantum computers, *Fortschr. Phys.* **46**, 877.

- Zalka, C. (1999), Grover's quantum searching algorithm is optimal, *Phys. Rev. A* **60**, 2746.
- Zanardi, P. and Rasetti, M. (1999), Holonomic quantum computation, *Phys. Lett. A* **264**, 94.
- Zeng, B., Chen, X., Zhou, D.-L., and Wen X.-G. (2015), Quantum information meets quantum matter, arXiv:1508.02595.
- Zhang, Q., Xu, F., Chen, Y.-A., Peng, C.-Z., and Pan J.-W. (2018), Large scale quantum key distribution: challenges and solutions, *Opt. Expr.* **26**, 24260.
- Zhirnov, V., Cavin, R., and Gammaiton, L. (2014), Minimum energy of computing, fundamental considerations, in "ICT - energy-concepts towards zero-power information and communication technology", Fagas G., Gammaiton L., Paul, D., and Berini G. A. (Eds.), InTech.
- Zippel, R. E. (1979), Probabilistic algorithms for sparse polynomials, in Proceedings of EUROSAM, Springer Lecture Notes in Computer Science **72**, p. 216.
- Zohar, E. and Reznik, B. (2011), The Fermi problem in discrete systems, *New J. Phys.* **13**, 075016.
- Zurek, W. H. (1985), Cosmological experiments in superfluid helium?, *Nature* **317**, 505.
- Zurek, W. H. (1991), Decoherence and the transition from quantum to classical, *Phys. Today*, October 1991, p. 36; see also arXiv:quant-ph/0306072.
- Zurek, W. H. (2003), Decoherence, einselection, and the quantum origins of the classical, *Rev. Mod. Phys.* **75**, 715.
- Zwolak, M. and Vidal, G. (2004), Mixed-state dynamics in one-dimensional quantum systems: A time-dependent superoperator renormalization algorithm, *Phys. Rev. Lett.*, **93**, 207205.
- Życzkowski, K., Horodecki, P., Sanpera A., and Lewenstein, M. (1998), Volume of the set of separable states, *Phys. Rev. A* **58**, 883.
- Życzkowski, K. and Sommers, H.-J. (2001), Induced measures in the space of mixed quantum states, *J. Phys. A: Math. Gen.* **34**, 7111.

# Index

- accessible information, 357  
adiabatic  
    condition, 136  
    connection forms, 143  
    quantum computation, 145  
    theorem, 134  
adjoint operator, 542  
affine map, 294, 323  
Aharonov–Bohm effect, 138  
AKLT state, 491  
algorithmic complexity, 37  
alphabet, 248, 252  
amplitude damping, 302  
amplitudes, 59  
AND gate, 18  
annihilation operator, 223, 313, 403, 422,  
    444  
anti-commutator, 552  
anti-Jaynes–Cummings Hamiltonian, 429  
area law, 484, 486  
Aspect’s experiment, 80  
authentication, 198  
avalanche photodiode, 449  
  
backward sign propagation, 114, 158, 411  
baker’s map, 183  
basis for a vector space, 538, 547, 557  
bath, 311  
BB84 protocol, 207  
BBO crystal, 448  
beam splitter, 444  
Bell  
    basis, 88, 115  
    inequalities, 77  
    measurement, 216  
states, 74, 88, 115, 201  
Berry  
    connection, 137  
    curvature, 138  
    phase, 135, 137  
billiard-ball computer, 46  
binary  
    addition table, 17  
    arithmetics, 17  
    entropy, 249  
    functions, 128  
bipartite system, 86  
bit, 16  
    -flip channel, 298  
    -phase-flip channel, 300  
Bloch  
    ball, 101  
    -Fano representation, 307  
    sphere, 99, 294  
    vector, 99, 101  
blue sideband, 429  
Bogoliubov  
    quasiparticles, 474  
    rotation, 474  
    transformation, 228  
    vacuum, 474  
Boolean functions, 21  
Born approximation, 310, 315  
Bose–Hubbard model, 459  
bound entangled states, 245  
**BPP** class, 32  
**BQP** class, 33  
Bužek–Hillery copying machine, 204  
Cæsar cypher, 195

- canonical ensemble, 282
- capacity of quantum channels, 363
- carrier resonance, 427
- CARRY circuit, 132
- cat states, 330
- Cauchy–Schwarz inequality, 537
- cavity QED, 418, 461
- CCNOT gate, 45
- centre-of-mass mode, 426
- change of basis, 547
- chaotic systems, 35, 184
- characteristic equation, 548
- charge qubits, 438
- Chernoff bound, 34
- Cholesky decomposition, 555
- CHSH inequality, 79
- Church–Turing thesis, 13
- Cirac–Zoller CNOT gate, 430
- circuit
  - model of computation, 16, 105
  - QED, 440
- circuit QED, 461
- classical
  - capacity, 364
  - correlations, 271
  - cryptography, 195
  - fidelity, 355
- Clebsch–Gordan coefficients, 404
- CMINUS gate, 114
- CNOT gate, 44, 111
- coarse graining, 285
- codewords, 381
- coding-decoding, 374
- coherences, 86, 293, 303, 330, 332, 335
- coherent
  - information, 368
  - states, 223
- coin-tossing sequences, 37
- collapse
  - and revival, 630
  - of the wave function, 67
- collective
  - attack, 209
  - decoherence, 400
  - measurements, 382
- commutator, 551
- complementary channel, 369
- complete
  - orthonormal set, 549
  - positivity, 291
- set of vectors, 538
- completely positive map, 245
- completeness relation, 541, 559
- complexity classes, 30
- composite
  - systems, 72, 86
- computational
  - basis, 98
  - complexity, 22
- concatenated codes, 414
- concurrence, 264
- conditional entropy, 250, 271
- conservative systems, 35, 66
- constant of motion, 66
- continuous variable systems, 220
- control qubit, 111
- controlled
  - EXCHANGE gate, 45
  - gates, 111
  - phase shift gate, 113
- Cooper pair, 437
  - box, 438
- COPY gate, 19
- correctable errors, 390
- covariance matrix, 234
- CPHASE gate, 113
- creation operator, 223, 313, 403, 422, 444
- CROSSOVER gate, 19
- cryptography, 195
- cyclic property of the trace, 553
- D-Wave, 464
- damping probability, 302
- data compression, 348
- De Morgan’s identities, 20
- decision problems, 26
- decoherence
  - free dynamics, 401
  - free subspaces, 399
  - models, 293
  - rate, 304
- de-entanglement, 305
- degenerate codes, 390
- degradable channel, 369
- dense coding, 212
- density
  - matrix, 82
  - operator, 82
- depolarizing channel, 301
- deterministic chaos, 35

- Deutsch's algorithm, 157  
Deutsch–Jozsa problem, 158  
diagonal representation, 547  
diagonalizable operators, 549  
dimension of a vector space, 535, 538  
Dirac delta function, 558  
discrete Fourier transform, 168  
discretized  
    Berry phase, 140  
    non-Abelian Berry phase, 144  
displacement operator, 224  
distance of a code, 391  
DMRG, 506  
    finite-system, 508  
    infinite-system, 506  
Doppler cooling, 430  
dual  
    -rail representation, 445  
    vector, 537  
dynamical  
    decoupling, 404  
    Kolmogorov–Sinai entropy, 284  
    systems, 34, 179
- E91 protocol, 210  
eigenvalues, 548  
eigenvectors, 548  
Einstein's local realism, 76  
element of physical reality, 73  
elementary logical gates, 18  
encoding efficiency, 404  
energy dissipation, 40, 44, 50  
entangled states, 73  
entanglement, 72, 107  
    concentration, 256  
    cost, 259  
    distillable, 259  
    -enhanced transmission, 371  
    fidelity, 365  
    generation, 111, 423  
    monogamy, 269  
    of formation, 264  
    of random states, 260  
    revivals, 327  
    witnesses, 246  
entropy, 39  
    exchange, 367  
    of a Gaussian state, 277  
EPR  
    pairs, 74, 88, 201, 210, 213, 216  
paradox, 73  
phenomenon, 88  
states, 74, 88, 210, 424, 448  
error  
    propagation, 411  
    syndrome, 382  
Euler's decomposition, 556  
Eulerian cycle, 26
- factoring problem, 23, 31, 33, 199  
FANOUT gate, 19  
fast Fourier transform, 161, 185  
fault-tolerant  
    quantum computation, 411  
    quantum gates, 413  
feasible problems, 23  
fibre-based quantum cryptography, 450  
fidelity, 104, 188, 201, 204, 353  
five-qubit code, 396  
Fock  
    basis, 223  
    states, 223, 423  
Fourier transform, 221  
Fredkin gate, 45  
free  
    energy, 281  
    particle, 221  
    -space quantum cryptography, 450  
function evaluation, 128
- garbage bits, 46  
Gaussian  
    operation, 234  
    state, 233  
    wave packet, 221  
generalized measurements, 93  
geometric  
    discord, 276  
    phase, 135  
GHZ state, 218  
GKLS master equation, 320, 340  
global phase, 66  
Gram–Schmidt decomposition, 540  
Gray code, 120  
Grover's algorithm, 161, 166
- Hadamard gate, 108  
halting problem, 15  
Hamiltonian, 35, 63  
    cycle, 27

- harmonic oscillator, 222
  - damping, 316
- Heisenberg
  - spin-1 model, 491
  - spin-1/2 model, 469
  - inequality, 69
    - uncertainty principle, 69
- Hermitian
  - conjugate operator, 542
  - operators, 542
- hidden variables, 76
- Hilbert spaces, 64, 72, 106, 537
- Holevo
  - bound, 358
  - information, 358, 373
- holonomic quantum computation, 141, 152
- Hubbard model, 467
- Huffman code, 348, 349
- hypercomputation, 13
- hyperfine
  - coupling, 433
  - qubits, 432
- information reconciliation, 208
- inner product, 536
- integer factoring, 175
- integrable systems, 38
- inter-qubit interactions, 111
- interaction picture, 311
- intercept and resend attack, 209
- interference term, 66
- intractable problems, 23
- intrapolation, 218
- intrusion detection, 207
- inverse operator, 542
- ion-trap quantum computer, 424
- irreversible
  - computation, 43
  - transformations, 281
- Ising chain, 469
- Jaynes–Cummings
  - Hamiltonian, 422, 463
  - Hubbard model, 463
  - model, 421, 423
- joint entropy, 250
- Jordan–Wigner transformation, 470
- Josephson
  - energy, 437
- junction, 437
- Kerr medium, 446
- key storage, 212
- Kitaev chain, 481
- Kraus
  - operators, 288, 294, 319, 371, 402
  - representation, 287, 288, 319, 402
  - representation theorem, 291
- Kronecker symbol, 538
- Lamb–Dicke parameter, 428–430
- Landauer’s principle, 40, 49
- laser cooling, 430
- law of large numbers, 347
- Levi–Civita antisymmetric tensor, 138, 294
- Lindblad operators, 320, 340
- linear
  - independence, 538
  - operators, 540
  - optics, 444
  - vector spaces, 535
    - of infinite dimension, 557
    - of finite dimension, 535
- Liouville superoperator, 518
- locality principle, 73
- LOCC, 241
- logarithmic negativity, 266
- logic functions, 128
- logical operations for stabilizer codes, 395
- logistic map, 36
- Lubkin’s formula, 262
- Lyapunov exponents, 38, 285
- Mach–Zehnder interferometer, 445, 453
- Majorana fermions, 476, 482
- majority voting, 379
- many-body localization, 487
- Markov
  - approximation, 310, 315
- Markov chain, 372
- master equation, 310, 322
  - and quantum operations, 319
  - for a single qubit, 322
- Mathieu equation, 426
- matrix
  - exponential, 544
  - representation, 543
- Maxwell’s demon, 39, 48
- measurement, 97, 98, 103

- postulate, 64
- MERA, 525
- microcanonical ensemble, 282
- micromotion, 632
- minimum spanning tree, 24
- minterms, 21, 129
- mixed states, 81
- MPO, 518
- MPS, 487, 489
  - equivalence with DMRG, 510
  - excited states, 505
  - finite temperature, 515
  - gauge freedom, 498
  - graphical representation, 492
  - normalization, 494
  - observables, 494, 500
  - representation of generic states, 501
  - scaling of correlations, 496
  - Schmidt decomposition, 500
  - TEBD, 514
  - variational optimization, 503
- multipartite entanglement, 266
- mutual information, 251, 270
- NAND gate, 19
- negativity, 265
- Nernst's theorem, 280
- Neumark's theorem, 94
- neuromorphic computing, 53
- nine-qubit Shor code, 385, 392
- no-cloning theorem, 199
- noise
  - channels, 296
  - threshold for quantum computing, 414
- non
  - Abelian geometric phase, 141
  - commuting observables, 70
  - degenerate codes, 391
  - divisible quantum maps, 325
  - linear
    - optics, 446
    - sign shift, 447
  - Markovian quantum dynamics, 324
  - separable states, 73
- NOR gate, 19
- norm of a vector, 537
- normal operators, 550
- NOT gate, 18
- NP class, 30
- NP-complete problems, 31
- NPC** class, 31
- number
  - operator, 223
  - states, 423
- open quantum systems, 310
- operator-sum representation, 288, 290
- optical
  - lattices, 459
  - qubits, 432
- optimization problems, 24
- OR gate, 18
- orthogonality, 538
- orthonormality, 538, 559
- outcome of a measurement, 64
- P** class, 30
- Page's formula, 261
- parametric down-conversion, 448, 449
- partial
  - trace, 87
  - transpose, 244
- participation ratio, 91
- Paul trap, 425
- Pauli
  - matrices, 544, 546, 550, 552
  - operators, 70, 72, 103
- PEPS, 519
  - contraction, 521
  - variational optimization, 523
- Peres separability criterion, 244
- period finding, 175
- Pesin's theorem, 284
- phase
  - damping, 303
  - flip channel, 299
  - kick, 303
  - shift gate, 108
  - shifter, 444
  - space representations, 230
  - transition, 28
- photon blockade, 463
- Planck's constant, 57, 63
- Pockels cell, 452
- pointer states, 337
- polar decomposition, 554
- polarization
  - qubit, 445
  - rotator, 445
- polynomial time, 30

- populations, 85, 293
- position and momentum operators, 563
- representations, 560
- uncertainty relations, 220
- positive map, 245
- post-measurement state, 93
- postulates of quantum mechanics, 63
- POVM, 94, 272, 342
- preferential basis, 337
- preamble, 335
- preparation of the initial state, 97, 124
- prime factorization problem, 178
- privacy amplification, 208
- probabilistic
  - quantum gates, 447
  - Turing machine, 15
- projective measurements, 94
- projectors, 541
- PSPACE** class, 32
- public-key cryptosystems, 197
- pure states, 81
- purification, 91
- QR decomposition, 554, 555
- quantum
  - adder, 132
  - advantage, 466
  - annealing, 465
  - baker's map, 184
  - black box, 294
  - Brownian motion, 326
  - capacity, 365, 368
  - communication with photons, 443
  - copying machine, 204
  - critical point, 475
  - cryptography, 207, 449
    - with continuous variables, 237
  - data compression, 351
  - data processing inequality, 368
  - discord, 269, 272
    - of a Gaussian state, 278
  - dots, 434
  - error correction, 389
  - Fourier transform, 168
  - frustration, 469
  - Hamming bound, 391
  - jump detection, 431
  - jumps, 340
  - key distribution, 207, 449
- measurements, 335
- memory channels, 371
- mutual information, 271
- of energy, 60
- operation, 288, 295
- parallelism, 131
- phase
  - estimation, 171
  - transition, 478
- register, 105
- search, 161
- simulation, 179
- simulators, 458
- state engineering, 425
- teleportation, 215
- to classical
  - correspondence, 330
  - transition, 329
- trajectories, 340
- trajectory method, 341
- qubit, 97, 98
  - damping, 319
  - polarization, 102
- Rabi
  - frequency, 116
  - oscillations, 409, 420, 422, 427
  - pulse, 423
  - pulses, 116
- Raman
  - configuration, 432
  - transition, 430, 432
- reality principle, 73
- red sideband, 428
- reduced density matrix, 87
- relative
  - entropy, 264
  - phases, 66, 108
- reservoir, 311
- reversible
  - computation, 43, 51
  - transformations, 280
- rotation operators, 109
- RSA cryptosystem, 198
- Rydberg atoms, 418, 419
- satisfiability, 28, 148
- sawtooth map, 185
- scalars, 535
- Schmidt

- decomposition theorem, 89  
number, 90  
rank, 90  
Schrödinger's cat, 329  
Schrödinger equation, 63, 179  
Schumacher's quantum noiseless coding theorem, 351  
secular motion, 632  
selective measurement, 408  
self-adjoint operator, 542  
semigroup, 290  
separability criteria, 243  
separable states, 73, 242  
Shannon entropy, 248  
Shannon's noiseless coding theorem, 346  
Shor's algorithm, 175  
sideband cooling, 430  
simultaneous diagonalization theorem, 552  
single  
-electron quantum dot, 435  
-qubit gates, 108  
singlet state, 404  
Singular value decomposition, 553  
solid-state qubits, 433  
spectral decomposition, 549  
spectrum of an operator, 549  
speed  
limit of single-qubit gates, 151  
of quantum gates, 149  
spin, 57  
-boson model, 402  
-singlet state, 74  
spontaneous emission, 430  
square-well potential, 437  
squeezed states, 227  
stabilizer coding, 392, 394  
state vector, 63  
stationary states, 67  
statistical  
entropy, 282  
mixture, 65, 81, 307, 340  
Stern–Gerlach experiment, 56, 70  
Stirling's formula, 347  
stretch mode, 426  
strong  
Church–Turing thesis, 24  
-coupling regime, 418  
SUM circuit, 132  
superconducting qubit circuits, 437, 464  
superoperator, 288  
superposition principle, 59, 106  
Suzuki–Trotter decomposition, 513  
SWAP gate, 19, 113  
symplectic decompositions, 555  
target qubit, 111  
teleportation, 215  
tensor  
maps, 545  
networks, 492, 519  
product, 545  
thermal decomposition, 236  
thermodynamic entropy, 280  
three-tangle, 269  
three-qubit  
bit-flip code, 381  
phase-flip code, 384  
time-evolution operator, 64  
Toffoli gate, 45, 119  
topological quantum computation, 484  
trace, 553  
tractable problems, 23  
transfer matrix, 494  
transistor, 22  
translucent attack, 209  
transpose matrix, 544  
trap-door functions, 197  
travelling salesman problem, 25  
triplet state, 404  
TTN, 524  
tunnel  
effect, 437  
junction, 438  
Turing  
machine, 9  
number, 15  
two  
-dimensional electron gas, 435  
-level atom, 421  
-mode Gaussian states, 237  
typical  
sequence, 347  
subspace, 352  
ultracold atoms, 458  
unfeasible problems, 23  
unit vector, 537, 541  
unitary  
errors, 127  
operators, 542

- representation, 289
- universal
  - classical computation, 21
  - classical gates, 21
  - quantum gates, 117
  - quantum simulation, 190
  - Turing machine, 14
- unstructured database, 161
- vacuum state, 224, 423
- vectorization, 517
- vectors, 535
- Vernam cypher, 196
- vibrational qubit, 431
- von Neumann
  - entropy, 252, 373, 423
  - equation, 83
- wave
- function, 63
- particle duality, 62
- vector, 68
- weak measurements, 337
- Werner state, 244
- Wigner function, 230
- Williamson's theorem, 556
- XOR gate, 19
- Young's double-slit experiment, 59
- Zeeman splitting, 435
- Zeno effect, 407
- zero
  - operator, 541
  - vector, 536
- zero-energy Majorana mode, 483