

LECTURE NOTES
IN PHYSICS

L. Diósi

A Short Course in Quantum Information Theory

An Approach
From Theoretical Physics

 Springer

Lecture Notes in Physics

Editorial Board

R. Beig, Wien, Austria
W. Beiglböck, Heidelberg, Germany
W. Domcke, Garching, Germany
B.-G. Englert, Singapore
U. Frisch, Nice, France
P. Hänggi, Augsburg, Germany
G. Hasinger, Garching, Germany
K. Hepp, Zürich, Switzerland
W. Hillebrandt, Garching, Germany
D. Imboden, Zürich, Switzerland
R. L. Jaffe, Cambridge, MA, USA
R. Lipowsky, Golm, Germany
H. v. Löhneysen, Karlsruhe, Germany
I. Ojima, Kyoto, Japan
D. Sornette, Nice, France, and Zürich, Switzerland
S. Theisen, Golm, Germany
W. Weise, Garching, Germany
J. Wess, München, Germany
J. Zittartz, Köln, Germany

The Lecture Notes in Physics

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

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

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

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

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

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

Lajos Diósi

A Short Course in Quantum Information Theory

An Approach From Theoretical Physics

 Springer

Author

Dr. Lajos Diósi
KFKI Research Institute for
Particle and Nuclear Physics
P.O.Box 49
1525 Budapest
Hungary
E-mail: diosi@rmki.kfki.hu

L. Diósi, *A Short Course in Quantum Information Theory*, Lect. Notes Phys. 713
(Springer, Berlin Heidelberg 2007), DOI 10.1007/b11844914

Library of Congress Control Number: 2006931893

ISSN 0075-8450

ISBN-10 3-540-38994-6 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-38994-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springer.com

© Springer-Verlag Berlin Heidelberg 2007

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: by the author and techbooks using a Springer L^AT_EX macro package
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper SPIN: 11844914 54/techbooks 5 4 3 2 1 0

Preface

Quantum information has become an independent fast growing research field. There are new departments and labs all around the world, devoted to particular or even complex studies of mathematics, physics, and technology of controlling quantum degrees of freedom. The promised advantage of quantum technologies has obviously electrified the field which had been considered a bit marginal until quite recently. Before, many foundational quantum features had never been tested or used on single quantum systems but on ensembles of them. Illustrations of reduction, decay, or recurrence of quantum superposition on single states went to the pages of regular text-books, without being experimentally tested ever. Nowadays, however, a youngest generation of specialists has imbibed quantum theoretical and experimental foundations “from infancy”.

From 2001 on, in spring semesters I gave special courses for under- and post-graduate physicists at Eötvös University. The twelve lectures could not include all standard chapters of quantum information. My guiding principles were those of the theoretical physicist and the believer in the unity of physics. I achieved a decent balance between the core text of quantum information and the chapters that link it to the edifice of theoretical physics. Scholarly experience of the passed five semesters will be utilized in this book.

I suggest this thin book for all physicists, mathematicians and other people interested in universal and integrating aspects of physics. The text does not require special mathematics but the elements of complex vector space and of probability theories. People with prior studies in basic quantum mechanics make the perfect readers. For those who are prepared to spend many times more hours with quantum information studies, there have been exhaustive monographs written by Preskill, by Nielsen and Chuang, or the edited one by Bouwmeester, Ekert, and Zeilinger. And for each of my readers, it is almost compulsory to find and read a second thin book “Short Course in Quantum Information, approach from experiments” . . .

Acknowledgements I benefited from the conversations and/or correspondence with Jürgen Audretsch, András Bodor, Todd Brun, Tova Feldmann, Tamás Geszti, Thomas Konrad, and Tamás Kiss. I am grateful to them all for the generous help and useful remarks that served to improve my manuscript.

It is a pleasure to acknowledge financial support from the Hungarian Scientific Research Fund, Grant No. 49384.

Budapest,
February 2006

Lajos Diósi

Contents

1	Introduction	1
2	Foundations of classical physics	5
2.1	State space	5
2.2	Mixing, selection, operation	5
2.3	Equation of motion	6
2.4	Measurements	6
2.4.1	Projective measurement	7
2.4.2	Non-projective measurement	9
2.5	Composite systems	9
2.6	Collective system	11
2.7	Two-state system (bit)	11
	Problems	12
3	Semiclassical — semi-Q-physics	15
	Problems	16
4	Foundations of q-physics	19
4.1	State space, superposition	19
4.2	Mixing, selection, operation	20
4.3	Equation of motion	20
4.4	Measurements	21
4.4.1	Projective measurement	22
4.4.2	Non-projective measurement	23
4.4.3	Continuous measurement	24
4.4.4	Compatible physical quantities	25
4.4.5	Measurement in pure state	26
4.5	Composite systems	27
4.6	Collective system	29
	Problems	29
5	Two-state q-system: qubit representations	31
5.1	Computational-representation	31
5.2	Pauli representation	32
5.2.1	State space	32

5.2.2	Rotational invariance	33
5.2.3	Density matrix	34
5.2.4	Equation of motion	35
5.2.5	Physical quantities, measurement	35
5.3	The unknown qubit, Alice and Bob	36
5.4	Relationship of computational and Pauli representations	37
	Problems	37
6	One-qubit manipulations	39
6.1	One-qubit operations	39
6.1.1	Logical operations	39
6.1.2	Depolarization, re-polarization, reflection	40
6.2	State preparation, determination	42
6.2.1	Preparation of known state, mixing	42
6.2.2	Ensemble determination of unknown state	43
6.2.3	Single state determination: no-cloning	44
6.2.4	Fidelity of two states	44
6.2.5	Approximate state determination and cloning	45
6.3	Indistinguishability of two non-orthogonal states	45
6.3.1	Distinguishing via projective measurement	46
6.3.2	Distinguishing via non-projective measurement	46
6.4	Applications of no-cloning and indistinguishability	47
6.4.1	Q-banknote	47
6.4.2	Q-key, q-cryptography	48
	Problems	50
7	Composite q-system, pure state	53
7.1	Bipartite composite systems	53
7.1.1	Schmidt decomposition	53
7.1.2	State purification	54
7.1.3	Measure of entanglement	55
7.1.4	Entanglement and local operations	56
7.1.5	Entanglement of two-qubit pure states	57
7.1.6	Interchangeability of maximal entanglements	58
7.2	Q-correlations history	59
7.2.1	EPR, Einstein-nonlocality 1935	59
7.2.2	A non-existing linear operation 1955	60
7.2.3	Bell nonlocality 1964	62
7.3	Applications of Q-correlations	64
7.3.1	Superdense coding	64
7.3.2	Teleportation	65
	Problems	67

8	All q-operations	69
8.1	Completely positive maps	69
8.2	Reduced dynamics	70
8.3	Indirect measurement	71
8.4	Non-projective measurement resulting from indirect measurement	73
8.5	Entanglement and LOCC	74
8.6	Open q-system: master equation	75
8.7	Q-channels	75
	Problems	76
9	Classical information theory	79
9.1	Shannon entropy, mathematical properties	79
9.2	Messages	80
9.3	Data compression	80
9.4	Mutual information	82
9.5	Channel capacity	83
9.6	Optimal codes	83
9.7	Cryptography and information theory	84
9.8	Entropically irreversible operations	84
	Problems	85
10	Q-information theory	87
10.1	Von Neumann entropy, mathematical properties	87
10.2	Messages	88
10.3	Data compression	89
10.4	Accessible q-information	91
10.5	Entanglement: the resource of q-communication	91
10.6	Entanglement concentration (distillation)	93
10.7	Entanglement dilution	94
10.8	Entropically irreversible operations	95
	Problems	96
11	Q-computation	99
11.1	Parallel q-computing	99
11.2	Evaluation of arithmetic functions	100
11.3	Oracle problem: the first q-algorithm	101
11.4	Searching q-algorithm	103
11.5	Fourier algorithm	104
11.6	Q-gates, q-circuits	105
	Problems	106
	Solutions	109
	References	123
	Index	125

Symbols, acronyms, abbreviations

$\{ , \}$	Poisson bracket	$\circ$	composition
$[,]$	commutator	$\times$	Cartesian product
$\langle \rangle$	expectation value	$\otimes$	tensor product
$\hat{O}$	matrix	tr	trace
$\hat{O}^\dagger$	adjoint matrix	tr_A	partial trace
$\oplus$	modulo sum		
$x, y, \dots$	phase space points	w	weight in mixture
Γ	phase space	$ \uparrow\rangle, \downarrow\rangle$	spin-up, spin-down basis
$\rho(x)$	phase space distribution, classical state	$\mathbf{n}, \mathbf{m} \dots$	Bloch unit vectors
$x, y \dots$	binary numbers	$ \mathbf{n}\rangle$	qubit state vector
$x_n \dots x_2 x_1$	binary string	$\mathbf{s}$	qubit polarization vector
$\rho(\mathbf{x})$	discrete classical state	$\hat{\sigma}_x, \hat{\sigma}_y \hat{\sigma}_z$	Pauli matrices
$\mathcal{M}$	operation	$\hat{\boldsymbol{\sigma}}$	vector of Pauli matrices
$\mathcal{T}$	polarization reflection	$\mathbf{a}, \mathbf{b}, \boldsymbol{\alpha}, \dots$	real spatial vectors
$\mathcal{I}$	identity operation	$\mathbf{ab}$	real scalar product
$\mathcal{L}$	Lindblad generator	$\hat{x}$	qubit hermitian matrix
$A(x), A(\mathbf{x})$	classical physical quantities	X, Y, Z	one qubit Pauli gates
$H(x)$	Hamilton function	H	Hadamard gate
P	indicator function	$T(\varphi)$	phase gate
$\Pi(x), \Pi(\mathbf{x})$	classical effect	F	fidelity
$\mathcal{H}$	Hilbert space	E	entanglement measure
d	vector space dimension	$S(\rho), S(p)$	Shannon entropy
$ \psi\rangle, \varphi\rangle, \dots$	state vectors	$S(\hat{\rho})$	von Neumann entropy
$\langle\psi , \langle\varphi , \dots$	adjoint state vectors	$S(\rho' \parallel \rho), S(\hat{\rho}' \parallel \hat{\rho})$	relative entropy
$\langle\psi \varphi\rangle$	complex inner product	$ \psi^\pm\rangle, \Phi^\pm\rangle$	Bell basis vectors
$\langle\psi \hat{O} \varphi\rangle$	matrix element	$ \mathbf{x}\rangle$	computational basis vector
$\hat{\rho}$	density matrix, quantum state	$\hat{M}_n$	Kraus matrices
$\hat{A}$	quantum physical quantity	$ n; E\rangle$	environmental basis vector
$\hat{H}$	Hamiltonian	$X, Y, \dots$	classical message
$\hat{P}$	hermitian projector	$H(X), H(Y)$	Shannon entropy
$\hat{I}$	unit matrix	$H(X Y)$	conditional Shannon entropy
$\hat{U}$	unitary map	$I(X:Y)$	mutual information
$\hat{\Pi}$	quantum effect	C	channel capacity
p	probability	$\rho(\mathbf{x} \mathbf{y})$	conditional state
		$\rho(\mathbf{y} \mathbf{x})$	transfer function
q-	quantum	LO	local operation
cNOT	controlled NOT	LOCC	local operation and classical communication

1 Introduction

Classical physics — the contrary to quantum — means all those fundamental dynamical phenomena and their theories which became known until the end of the 19th century, from our studying the macroscopic world. Galileo's, Newton's, and Maxwell's consecutive achievements, built one on the top of the other, obtained their most compact formulation in terms of the classical canonical dynamics. At the same time, the conjecture of the atomic structure of the microworld was also conceived. By extending the classical dynamics to atomic degrees of freedom, certain microscopic phenomena also appearing at the macroscopic level could be explained correctly. This yielded indirect, yet sufficient, proof of the atomic structure. But other phenomena of the microworld (e.g., the spectral lines of atoms) resisted to the natural extension of the classical theory to the microscopic degrees of freedom. After Planck, Einstein, Bohr, and Sommerfeld, there had formed a simple constrained version of the classical theory. The naively *quantized* classical dynamics was already able to describe the non-continuous (discrete) spectrum of stationary states of the microscopic degrees of freedom. But the detailed dynamics of the transitions between the stationary states was not contained in this theory. Nonetheless, the successes (e.g., the description of spectral lines) shaped already the *dichotomous* physics world concept: the microscopic degrees of freedom obey to other laws than macroscopic ones do. After the achievements of Schrödinger, Heisenberg, Born, and Jordan, the *quantum theory* emerged to give the complete description of the microscopic degrees of freedom in perfect agreement with experience. This quantum theory was not a mere quantized version of the classical theory anymore. Rather it was a totally new formalism of completely different structure than the classical theory, which was applied professedly to the microscopic degrees of freedom. As for the macroscopic degrees of freedom, one continued to insist on the classical theory.

For a sugar cube, the center of mass motion is a macroscopic degree of freedom. For an atom, it is microscopic. We must apply the classical theory to the sugar cube, and the quantum theory to the atom. Yet, there is no sharp boundary of where we must switch from one theory to the other. It is, furthermore, obvious that the center of mass motion of the sugar cube should be derivable from the center of mass motions of its atomic constituents. Hence a specific inter-dependence exists between the classical and the quantum theories, which must give consistent resolution for the above dichotomy. The von Neumann “axiomatic” formulation of the quantum theory represents, in the framework of the dichotomous physics world concept, a

description of the microworld maintaining the perfect harmony with the classical theory of the macroworld.

Let us digress about a natural alternative to the dichotomous concept. According to it, all macroscopic phenomena can be reduced to a multitude of microscopic ones. Thus in this way the basic physical theory of the universe would be the quantum theory, and the classical dynamics of macroscopic phenomena should be deducible from it, as limiting case. But the current quantum theory is not capable of holding its own. It refers to genuine macroscopic systems as well, thus requiring classical physics as well. Despite of the theoretical efforts in the second half of 20th century, there has not so far been consensus regarding the (universal) quantum theory which would in itself be valid for the whole physical world.

This is why we keep the present course of lectures within the framework of the dichotomous world concept. The “axiomatic” quantum theory of von Neumann will be used. Among the bizarre structures and features of this theory, discreteness (quantumness) was the earliest, and the theory also drew its name from it. Yet another odd prediction of quantum theory is the inherent randomness of the microworld. During the decades, further surprising features have come to light. It has become “fashion” to deduce paradoxical properties of quantum theory. There is a particular range of paradoxical predictions (Einstein-Podolski-Rosen, Bell) which exploits such correlations between separate quantum systems which could never exist classically. Another cardinal paradox is the non-clonability of quantum states, meaning the fidelity of possible copies will be limited fundamentally and strongly.

The initial role of the paradoxes was better knowledge of quantum theory. We learned the *differenciae specifica*e of the quantum systems with respect to the classical ones. The consequences of the primarily paradoxical quantumness are understood relatively well and also their advantage is appreciated with respect to classical physics (see, e.g., semiconductors, superconductivity, superfluidity). By the end of the 20th century the paradoxes related to *quantum-correlations* have come to the front. We started to discover their advantage only in the past decade. The keyword is: *information*! Quantum correlations, consequent upon quantum theory, would largely extend the options of classical information manipulation including information storage, coding, transmitting, hiding, protecting, evaluating, as well as algorithms, game strategies. All these represent the field of quantum information theory in a wider sense. Our short course covers the basic components only, at the introductory level.

Chapters 2–4 summarize the classical, the semiclassical, and the quantum physics. The two Chaps. 2 and 4 look almost like mirror images of each other. I intended to exploit the maximum of existing parallelism between the classical and quantum theories, and to isolate only the essential differences in the present context. Chapter 5 introduces the text-book theory of abstract two-state quantum systems. Chapter 6 discusses their quantum informatic manipulations and presents two applications: copy-protection of banknotes and of cryptographic keys. Chapter 7 is devoted to composite quantum systems and quantum correlations (also called entanglement). An insight into three theoretical antecedents is discussed, finally I show

two quantum informatic applications: superdense coding and teleportation. Chapter 8 introduces us to the modern theory of quantum operations. The first parts of Chaps. 9 and 10 are anew mirror images of each other. The foundations of classical and quantum information theories, based respectively on the Shannon and von Neumann entropies, can be displayed in parallel terms. This holds for the classical and quantum theories of data compression as well. There is, however, a separate section in Chap. 10 to deal with the entanglement as a resource, and with its conversions which all make sense only in quantum context. Chapter 11 offers simple introduction into the quintessence of quantum information which is quantum algorithms. I present the concepts that lead to the idea of the quantum computer. Two quantum algorithms will close the Chapter: solution of the oracle and of the searching problems. A short section of divers Problems and Exercises follow each Chapter. This can, to some extent, compensate the reader for the laconic style of the main text. A few number of missing or short-spoken proofs and arguments find themselves as Problems and Exercises. That gives a hint how the knowledge, comprised into the economic main text, could be derived and applied.

For further reading, we suggest the monograph [1] by Nielsen and Chuang which is the basic reference work for the time being, together with [2] by Preskill and [3] edited by Bouwmeester, Ekert and Zeilinger. Certain statements or methods, e.g. in Chaps. 10 and 11, follow [1] or [2] and can be checked from there directly. Our bibliography continues with textbooks [4]–[10] on the traditional fields, like e.g. the classical and quantum physics, which are necessary for the quantum information studies. References to two useful reviews on q-cryptography [11] and on q-computation are also included [12]. The rest of the bibliography consists of a very modest selection of the related original publications.

2 Foundations of classical physics

We choose the classical canonical theory of Liouville because of the best match with the q-theory — a genuine statistical theory. Also this is why we devote the particular Sect. 2.4 to the measurement of the physical quantities. Hence the elements of the present Chapter will most faithfully reappear in Chap. 4 on Foundations of q-physics. Let us observe the similarities and the differences!

2.1 State space

The state space of a system with n degrees of freedom is the phase space:

$$\Gamma = \{(q_k, p_k); k = 1, 2, \dots, n\} \equiv \{x_k; k = 1, 2, \dots, n\} \equiv \{x\}, \quad (2.1)$$

where q_k, p_k are the canonically conjugate coordinates of each degree of freedom in turn. The *pure* state of an individual system is described by the phase point $\bar{x}$. The generic state is *mixed*, described by normalized distribution function:

$$\rho \equiv \rho(x) \geq 0, \quad \int \rho dx = 1. \quad (2.2)$$

The generic state is interpreted on the statistical ensemble of identical systems. The distribution function of a pure state reads:

$$\rho_{\text{pure}}(x) = \delta(x - \bar{x}). \quad (2.3)$$

2.2 Mixing, selection, operation

Random mixing the elements of two ensembles of states ρ_1 and ρ_2 at respective rates $w_1 \geq 0$ and $w_2 \geq 0$ yields the new ensemble of state:

$$\rho = w_1 \rho_1 + w_2 \rho_2; \quad w_1 + w_2 = 1. \quad (2.4)$$

A generic mixed state can always be prepared (i.e. decomposed) as the mixture of two or more other mixed states in infinite many different ways. After mixing, however, it is totally impossible to distinguish which way the mixed state was prepared.

It is crucial, of course, that mixing must be probabilistic. A given mixed state can also be prepared (decomposed) as a mixture of pure states and this mixture is unique.

Let operation $\mathcal{M}$ on a given state ρ mean that we perform the same transformation on each system of the corresponding statistical ensemble. Mathematically, $\mathcal{M}$ is linear norm-preserving map of positive kernel to bring an arbitrary state ρ into a new state $\mathcal{M}\rho$. The operation's categorical linearity follows from the linearity of the procedure of mixing (2.4). Obviously we must arrive at the same state if we mix two states first and then we subject the systems of the resulting ensemble to the operation $\mathcal{M}$ or, alternatively, we perform the operation prior to the mixing the two ensembles together:

$$\mathcal{M}(w_1\rho_1 + w_2\rho_2) = w_1\mathcal{M}\rho_1 + w_2\mathcal{M}\rho_2 . \quad (2.5)$$

This is just the mathematical expression of the operation's linearity.

Selection of a given ensemble into specific sub-ensembles, a contrary process of mixing, will be possible via so-called selective operations. They correspond mathematically to norm-reducing positive maps. The most typical selective operations are called measurements 2.4.

2.3 Equation of motion

Dynamical evolution of a closed system is determined by its real Hamilton function $H(x)$. The Liouville equation of motion takes this form¹:

$$\frac{d}{dt}\rho = \{H, \rho\} , \quad (2.6)$$

where $\{\cdot, \cdot\}$ stands for the Poisson brackets. For pure states, this yields the Hamilton equation of motion:

$$\frac{d\bar{x}}{dt} = \{H, \bar{x}\}; \quad H = H(\bar{x}) . \quad (2.7)$$

Its solution $\bar{x}(t) \equiv U(\bar{x}(0); t)$ represents the time-dependent invertible map $U(t)$ of the state space. The Liouville equation (2.6) implies the reversible operation $\mathcal{M}(t)$, which we can write formally as follows:

$$\rho(t) = \rho(0) \circ U^{-1}(t) \equiv \mathcal{M}(t)\rho(0) . \quad (2.8)$$

2.4 Measurements

Consider a partition $\{P_\lambda\}$ of the phase space. The functions $P_\lambda(x)$ are indicator-functions over the phase space, taking values 0 or 1. They form a complete set of pairwise disjoint functions:

¹ The form $d\rho/dt$ is used to match the tradition of q-theory notations, cf. Chap. 4, it stands for $\partial\rho(x, t)/\partial t$.

$$\sum_{\lambda} P_{\lambda} \equiv 1, \quad P_{\lambda} P_{\mu} = \delta_{\lambda\mu} P_{\lambda} . \quad (2.9)$$

We consider the indicator functions as binary physical quantities. The whole variety of *physical quantities* is represented by real functions $A(x)$ on the phase space. Each physical quantity A possesses, in arbitrary good approximation, the step-function expansion:

$$A(x) = \sum_{\lambda} A_{\lambda} P_{\lambda}(x); \quad \lambda \neq \mu \Rightarrow A_{\lambda} \neq A_{\mu} . \quad (2.10)$$

The real values A_{λ} are step-heights, $\{P_{\lambda}\}$ is a partition of the phase space.

The projective partition (2.9) can be generalized. We define a positive decomposition of the constant function:

$$1 = \sum_n \Pi_n(x) ; \quad \Pi_n(x) \geq 0 . \quad (2.11)$$

The elements of the positive decomposition, also called effects, are non-negative functions $\Pi_n(x)$, they need be neither disjoint functions nor indicator-functions at all. They are, in a sense, the unsharp version of indicator-functions.

2.4.1 Projective measurement

On each system in a statistical ensemble of state ρ , we can measure the simultaneous values of the indicator-functions P_{λ} of a given partition (2.9). The outcomes are random. One of the binary quantities, say P_{λ} , is 1 with probability

$$p_{\lambda} = \int P_{\lambda} \rho dx , \quad (2.12)$$

while the rest of them is 0:

$$\begin{array}{cccccccc} P_1 & P_2 & \dots & P_{\lambda-1} & P_{\lambda} & P_{\lambda+1} & \dots \\ \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots \end{array} . \quad (2.13)$$

The state suffers projection according to Bayes theorem of conditional probabilities:

$$\rho \rightarrow \rho_{\lambda} \equiv \frac{1}{p_{\lambda}} P_{\lambda} \rho . \quad (2.14)$$

The post-measurement state ρ_{λ} is also called conditional state, i.e., conditioned on the random outcome λ . As a result of the above measurement we have randomly selected the original ensemble of state ρ into sub-ensembles of states ρ_{λ} for $\lambda = 1, 2, \dots$

The projective measurement is repeatable. Repeated measurements of the indicator functions P_{μ} on ρ_{λ} yield always the former outcomes $\delta_{\lambda\mu}$. The above selection

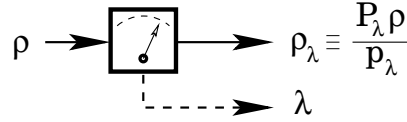


Fig. 2.1. Selective measurement. The ensemble of pre-measurement states ρ is selected into sub-ensembles of conditional post-measurement states ρ_λ according to the obtained measurement outcomes λ . The probability p_λ coincides with the norm of the unnormalized conditional state $P_\lambda \rho$

is also reversible. If we re-unite the obtained sub-ensembles, the post-measurement state becomes the following mixture of the conditional states ρ_λ :

$$\sum_{\lambda} p_{\lambda} \rho_{\lambda} = \sum_{\lambda} p_{\lambda} \frac{1}{p_{\lambda}} P_{\lambda} \rho = \rho . \quad (2.15)$$

This is, of course, identical to the original pre-measurement state.

By the projective measurement of a general physical quantity A we mean the projective measurement of the partition (2.9) generated by its step-function-expansion (2.10). The measured value of A is one of the step-heights:

$$A \rightarrow A_{\lambda} , \quad (2.16)$$

the probability of the particular outcome is given by (2.12). The projective measurement is always repeatable. If a first measurement yielded A_{λ} on a given state then also the repeated measurement yields A_{λ} . We can define the non-selective measured value of A , i.e., the average of A_{λ} taken with the distribution (2.12):

$$\langle A \rangle \equiv \sum_{\lambda} p_{\lambda} A_{\lambda} = \int A \rho dx . \quad (2.17)$$

This is also called the *expectation value* of A in state ρ .

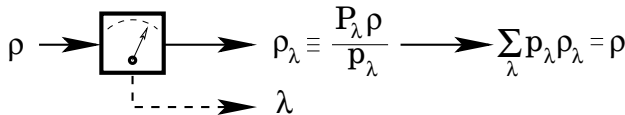


Fig. 2.2. Non-selective measurement. The sub-ensembles of conditional post-measurement states ρ_λ are re-united, contributing to the ensemble of non-selective post-measurement state which is, obviously, identical to the pre-measurement state ρ

2.4.2 Non-projective measurement

Non-projective measurement generalizes the projective one 2.4.1. On each system in a statistical ensemble of state ρ , we can measure the simultaneous values of the effects Π_n of a given positive decomposition (2.11) but we lose repeatability of the measurement. The outcomes are random. One of the effects, say Π_n , is 1 with probability

$$p_n = \int \Pi_n \rho dx, \quad (2.18)$$

while the rest of them is 0:

$$\begin{array}{ccccccc} \Pi_1 & \Pi_2 & \dots & \Pi_{n-1} & \Pi_n & \Pi_{n+1} & \dots \\ \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots \end{array} \quad (2.19)$$

The state suffers a change according to the Bayes theorem of conditional probabilities:

$$\rho \rightarrow \rho_n \equiv \frac{1}{p_n} \Pi_n \rho. \quad (2.20)$$

Contrary to the projective measurements, the repeated non-projective measurements yield different outcomes in general. The effects Π_n are not binary quantities, the individual measurement outcomes 0 or 1 provide unsharp information that can only orient the outcome of subsequent measurements. Still, the selective non-projective measurements are reversible. Re-uniting the obtained sub-ensembles, i.e., averaging the post-measurement conditional states ρ_n , yield the original pre-measurement state.

We can easily generalize the discrete set of effects for continuous sets. This generalization has a merit: one can construct the unsharp measurement of an arbitrarily chosen physical quantity A . One constructs the following set of effects:

$$\Pi_{\bar{A}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(\bar{A} - A(x))^2}{2\sigma^2} \right], \quad -\infty \leq \bar{A} \leq \infty. \quad (2.21)$$

These effects correspond to the unsharp measurement of A . The conditional post-measurement state will be $\rho_{\bar{A}}(x) = p_{\bar{A}}^{-1} \Pi_{\bar{A}}(x) \rho(x)$, cf. eq. (2.20). We interpret $\bar{A}$ as the random outcome representing the measured value of A at the standard measurement error σ . The outcomes' probability (2.18) turns out to be the following distribution function:

$$p_{\bar{A}} = \int \Pi_{\bar{A}}(x) \rho(x) dx, \quad (2.22)$$

normalized obviously by $\int p_{\bar{A}} d\bar{A} = 1$.

2.5 Composite systems

The phase space of the composite system, composed from the (sub)systems A and B , is the Cartesian product of the phase spaces of the subsystems:

$$\Gamma_{AB} = \Gamma_A \times \Gamma_B = \{(x_A, x_B)\} . \quad (2.23)$$

The state of the composite system is described by the normalized distribution function depending on both phase points x_A and x_B :

$$\rho_{AB} = \rho_{AB}(x_A, x_B) . \quad (2.24)$$

The reduced state of subsystem A is obtained by integration of the composite system's state over the phase space of the subsystem B

$$\rho_A = \int \rho_{AB} dx_B \equiv \mathcal{M} \rho_{AB} . \quad (2.25)$$

Our notation indicates that reduction, too, can be considered as an operation: it maps the states of the original system into the states of the subsystem. The state ρ_{AB} of the composite system is the product of the subsystem's states if and only if there is no statistical correlation between the subsystems. But generally there is some:

$$\rho_{AB} = \rho_A \rho_B + \text{cl. corr.} \quad (2.26)$$

Nevertheless, the state of the composite system is always *separable*, i.e., we can prepare it as the statistical mixture of product (uncorrelated) states:

$$\rho_{AB}(x_A, x_B) = \sum_{\lambda} w_{\lambda} \rho_{A\lambda}(x_A) \rho_{B\lambda}(x_B) , \quad w_{\lambda} \geq 0 , \quad \sum_{\lambda} w_{\lambda} = 1 . \quad (2.27)$$

The equation of motion of the composite system reads:

$$\frac{d}{dt} \rho_{AB} = \{H_{AB}, \rho_{AB}\} . \quad (2.28)$$

The composite Hamilton function is the sum of the Hamilton functions of the subsystems themselves plus the interaction Hamilton function:

$$H_{AB}(x_A, x_B) = H_A(x_A) + H_B(x_B) + H_{ABint}(x_A, x_B) . \quad (2.29)$$

If H_{ABint} is zero then the product initial state remains product state, the dynamics does not create correlation between the subsystems. Non-vanishing H_{ABint} does usually create correlation. The motion of the whole system is reversible, of course. But that of the subsystems is not. In case of product initial state $\rho_A(0)\rho_B(0)$, for instance, the *reduced dynamics* of the subsystem A will represent the time-dependent irreversible² operation $\mathcal{M}_A(t)$ which we can formally write as:

$$\rho_A(t) = \int \rho_A(0)\rho_B(0) \circ U_{AB}^{-1}(t) dx_B \equiv \mathcal{M}_A(t)\rho_A(0) . \quad (2.30)$$

The reversibility of the composite state dynamics has become lost by the reduction: the final reduced state $\rho_A(t)$ does not determine a unique initial state $\rho_A(0)$.

² Note that here and henceforth we use the notion of irreversibility as an equivalent to non-invertibility. We discuss the entropic-informatic notion of irreversibility in Sect. 9.8.

2.6 Collective system

The state (2.2) of a system is interpreted on the statistical ensemble of identical systems in the same state. We can form a multiple composite system from a big number n of such identical systems. This we call collective system, its state space is the n -fold Cartesian product of the elementary subsystem's phase spaces:

$$\Gamma \times \Gamma \times \dots \Gamma \equiv \Gamma^{\times n} . \quad (2.31)$$

The collective state reads:

$$\rho(x_1)\rho(x_2)\dots\rho(x_n) \equiv \rho^{\times n}(x_1, x_2, \dots, x_n) . \quad (2.32)$$

If $A(x)$ is a physical quantity of the elementary subsystem then, in a natural way, one can introduce its arithmetic mean, over the n subsystems, as a collective physical quantity:

$$\frac{A(x_1) + A(x_2) + \dots + A(x_n)}{n} . \quad (2.33)$$

Collective physical quantities are not necessarily of such simple form. Their measurement is the collective measurement. It can be reduced to independent measurements on the n subsystems.

2.7 Two-state system (bit)

Consider a system of a single degree of freedom, possessing the following Hamiltonian function:

$$H(q, p) = \frac{1}{2}p^2 + \frac{\omega^2}{8a^2} (q^2 - a^2)^2 . \quad (2.34)$$

The “double-well” potential has two symmetric minima at places $q = \pm a$, and a potential barrier between them. If the energy of the system is smaller than the barrier then the system is localized in one or the other well, moving there periodically “from wall to wall”. If, what is more, the energy is much smaller than the barrier height then the motion is restricted to the narrow parts around $q = a$ or $q = -a$, respectively, whereas the motion “from wall to wall” persists always. In that restricted sense has the system two-states.

One unit of information, i.e. one *bit*, can be stored in it. The localized motional state around $q = -a$ can be associated with the value 0 of a binary digit x , while that around $q = a$ can be associated with the value 1. The information storage is still perfectly reliable if we replace pure localized states and use their mixtures instead. However, the system is more protected against external perturbations if the localized states constituting the mixture are all much lower than the barrier height.

The original continuous phase space (2.1) of the system has thus been restricted to the discrete set $x = \{0, 1\}$ of two elements. Also the states (2.2) have become described by the discrete distribution $\rho(x)$ normalized as $\sum_x \rho(x) = 1$. There are

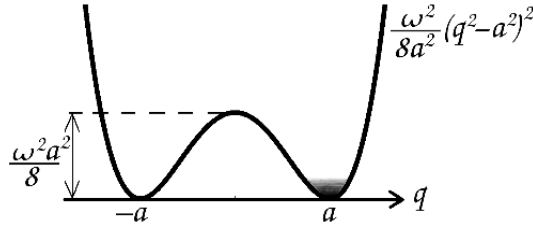


Fig. 2.3. Classical “two-state system” in double-well potential. The picture visualizes the state concentrated in the r.h.s. well. It is a mixture of periodic “from wall to wall” orbits of various energies that are still much smaller than the barrier height $\omega^2 a^2/8$. One can simplify this low energy regime into a discrete two-state system without the dynamics. The state space becomes discrete consisting of two points associated with $x = 0$ and $x = 1$ to store physically what will be called a bit x

only two pure states (2.3), namely δ_{x0} or δ_{x1} . To treat classical information, the concept of *discrete* state space will be essential in Chap. 9. In the general case, we use states $\rho(x)$ where x is an integer of, say, n binary digit. The corresponding system is a composite system of n bits.

Problems, exercises

2.1. Mixture of pure states. Let ρ be a mixed state which we mix from pure states. What are the weights we must take for the pure states, respectively? Let us start the solution with the two-state system.

2.2. Probabilistic or deterministic mixing? What happens if the mixing is not randomly performed? Let the target state of mixing be evenly distributed: $\rho(x) = 1/2$. Let someone mix an *equal* number n of the pure states δ_{x0} and δ_{x1} , respectively. Let us write down the state of this n -fold composite system. Let us compare it with the n -fold composite state corresponding to the proper, i.e. random, mixing.

2.3. Classical separability. Let us prove that a classical composite system is always separable. Method: let the index λ in (2.27) run over the phase space (2.1) of the composite system. Let us choose $\lambda = (\bar{x}_A, \bar{x}_B)$.

2.4. Decorrelating a single state? Does operation $\mathcal{M}$ exist such that it brings an arbitrary correlated state ρ_{AB} into the (uncorrelated) product state $\rho_A \rho_B$ of the reduced states ρ_A and ρ_B ? Remember, the operation $\mathcal{M}$ must be linear.

2.5. Decorrelating an ensemble. Give operation $\mathcal{M}$ such that brings $2n$ correlated states ρ_{AB} into n uncorrelated states $\rho_A \rho_B$: $\mathcal{M} \rho_{AB}^{\times 2n} = (\rho_A \rho_B)^{\times n}$. Method: consider a smart permutation of the $2n$ copies of the subsystem A , followed by a reduction to the suitable subsystem.

2.6. Classical indirect measurement. Let us prove that the non-projective measurement of arbitrarily given effects $\{H_n(x)\}$ can be obtained from projective measurements on a suitably enlarged composite state. Method: Construct the suitable composite state $\rho(x, n)$ to include a hypothetical detector system to count n ; perform projective measurement on the detector's n .

3 Semiclassical — semi-Q-physics

The dynamical laws of classical physics, given in Chap. 2, can approximatively be retained for microscopic systems as well, but with restrictions of the new type. The basic goal is to impose discreteness onto the classical theory. We add discretization q-conditions to the otherwise unchanged classical canonical equations. The corresponding restrictions must be graceful in a sense that they must not modify the dynamics of macroscopic systems and they must not destroy the consistency of the classical equations.

Let us assume that the dynamics of the microsystem is separable in the canonical variables (q_k, p_k) , and the motion is finite in phase space. The canonical action variables are defined as:

$$I_k \equiv \frac{1}{2\pi} \oint p_k dq_k, \quad (3.1)$$

for all degrees of freedom $k = 1, 2, \dots$ in turn. The integral is understood along one period of the finite motion in each degree of freedom. Action variables I_k are the adiabatic invariants¹ of classical motion. In classical physics they can take arbitrary values. To impose discreteness on classical dynamics, the Bohr–Sommerfeld q-condition says that each action I_k must be an integer multiple of the Planck constant (plus $\hbar/2$ in case of oscillatory motion):

$$I_k \equiv \frac{1}{2\pi} \oint p_k dq_k = (n_k + \tfrac{1}{2})\hbar. \quad (3.2)$$

The integer q-numbers n_k will label the discrete sequence of phase space trajectories which are, according to this semiclassical theory, the only possible motions. The state with $n_1 = n_2 = \dots = 0$ is the ground state and the excited states are separated by finite energy gaps from it.

Let us consider the double-well potential (2.34) with suitable parameters such that the lowest states be doubly degenerate, of approximate energies $\hbar\omega$, $2\hbar\omega$, $3\hbar\omega$ etc., localized in either the left- or the right-side well. The parametric condition is that the barrier be much higher than the energy gap $\hbar\omega$.

Let us store 1 bit of information in the two ground states, say the ground state in the left-side well means 0 and that in the right-side means 1. These two states are separated from all other states by a minimum energy $\hbar\omega$. Perturbations of energies smaller than $\hbar\omega$ are not able to excite the two ground states. In this sense the

¹ See, e.g., in Chap. VII. of [4].

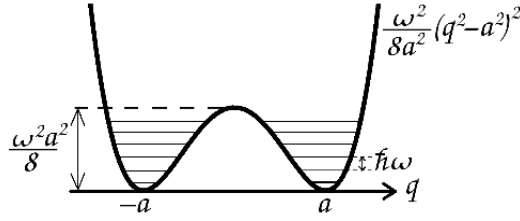


Fig. 3.1. Stationary q -states in double-well potential. The bottoms of the wells can be approximated by quadratic potentials $\frac{1}{2}\omega^2(q \mp a)^2$. Thus we obtain the energy-level structure of two separate harmonic oscillators, one in the l.h.s. well, the other in the r.h.s. well. This approximation breaks down for the upper part of the wells. Perfect two-state q -systems will be realized at low energies where the degenerate ground states do never get excited.

above system is a perfect autonomous two-state system provided the energy of its environment is sufficiently low. This autonomy follows from quantization and is the property of q -systems.

The Bohr–Sommerfeld theory classifies the possible stationary states of dynamically separable microsystems². It remains in debt of capturing non-stationary phenomena. The true q -theory (Chap. 4) will come to the decision that the generic, non-stationary, states emerge from superposition of the stationary states. In case of the above two-state system, the two ground states must be considered as the two orthonormal vectors of a two-dimensional complex vector space. Their normalized complex linear combinations will represent all states of the two-state quantum system. This q -system and its continual number of states will constitute the ultimate notion of q -bit or *qubit*.

Problems, exercises

3.1. Bohr quantization of the harmonic oscillator. Let us derive the Bohr–Sommerfeld q -condition for the one-dimensional harmonic oscillator of mass $m = 1$, bounded by the potential $\frac{1}{2}\omega^2 q^2$.

3.2. The role of adiabatic invariants. Consider the motion of the harmonic oscillator that satisfies the q -conditions with a certain q -number n . Suppose that we are varying the directional force ω^2 adiabatically, i.e., much slower than one period of oscillation. Physical intuition says that the motion of the system should invariably satisfy the q -condition to good approximation, even with the same q -number n . Is that true?

3.3. Classical-like or q -like motion. There is no absolute rule to distinguish between microscopic and macroscopic systems. It is more obvious to ask if a given

² The modern semiclassical theory is more general and powerful, cf. [5].

state (motion) is q-like or classical-like. In semiclassical physics, the state is q-like if the q-condition imposes physically relevant restrictions, and the state is classical-like if the imposed discreteness does not practically restrict the continuum of classical states. Let us argue that, in this sense, small integer q-numbers n mean q-like states and large ones mean classical-like states.

4 Foundations of q-physics

We present the standard q-theory¹ while, at each element, striving for the maximum likeness to Chap. 2 on Foundations of classical physics. We go slightly beyond the traditional treatment and, e.g., we define *non-projective* q-measurements as well as the phenomenon of *entanglement*. Leaf through Chap. 2 again, and compare!

4.1 State space, superposition

The state space of a q-system is a Hilbert space $\mathcal{H}$, in case of d -state q-system it is the d -dimensional complex vector space:

$$\mathcal{H} = C^d = \{c_\lambda; \lambda = 1, 2, \dots, d\}, \quad (4.1)$$

where the c_k 's are the elements of the complex column-vector in the given basis. The pure state of a q-system is described by a complex unit vector, also called state vector. In basis-independent abstract (Dirac-) notation it reads:

$$|\psi\rangle \equiv \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_d \end{bmatrix}, \quad \langle\psi| \equiv [c_1^*, c_2^*, \dots, c_d^*], \quad \sum_{\lambda=1}^d |c_\lambda|^2 = 1. \quad (4.2)$$

The inner product of two vectors is denoted by $\langle\psi|\varphi\rangle$. Matrices are denoted by a “hat” over the symbols, and their matrix elements are written as $\langle\psi|\hat{A}|\varphi\rangle$. In q-theory, the components c_k of the complex vector are called probability amplitudes. Superposition, i.e. normalized complex linear combination of two or more vectors, yields again a possible pure state.

The generic state is mixed, described by trace-one positive semidefinite density matrix:

$$\hat{\rho} = \{\rho_{\lambda\mu}; \lambda, \mu = 1, 2, \dots, d\} \geq 0, \quad \text{tr } \hat{\rho} = 1. \quad (4.3)$$

The generic state is interpreted on the statistical ensemble of identical systems. The density matrix of pure state (4.2) is a special case, it is the one-dimensional hermitian projector onto the subspace given by the state vector:

¹ Cf. [6] by von Neumann.

$$\hat{\rho}_{\text{pure}} = \hat{P} = |\psi\rangle \langle\psi| . \quad (4.4)$$

We can see that multiplying the state vector by a complex phase factor yields the same density matrix, i.e., the same q-state. Hence the phase of the state vector can be deliberately altered, still the same pure q-state is obtained. In the conservative q-theory, contrary to the classical theory, not even the pure state is interpreted on a single system but on the statistical ensemble of identical systems.

4.2 Mixing, selection, operation

Random mixing the elements of two ensembles of q-states $\hat{\rho}_1$ and $\hat{\rho}_2$ at respective rates $w_1 \geq 0$ and $w_2 \geq 0$ yields the new ensemble of the q-state

$$\hat{\rho} = w_1 \hat{\rho}_1 + w_2 \hat{\rho}_2; \quad 0 \leq w_1, w_2 \leq 1; \quad w_1 + w_2 = 1 . \quad (4.5)$$

A generic mixed q-state can always be prepared (i.e. decomposed) as the mixture of two or more other mixed q-states in infinitely many different ways. After mixing, however, it is totally impossible to distinguish which way the mixed q-state was prepared. It is crucial, of course, that mixing must be probabilistic. A given mixed q-state can also be prepared (decomposed) as a mixture of pure q-states and this mixture is, contrary to the classical case, not unique in general.

Let operation $\mathcal{M}$ on a given q-state $\hat{\rho}$ mean that we perform the same transformation on each q-system of the corresponding statistical ensemble. Mathematically, $\mathcal{M}$ is linear trace-preserving completely positive map, cf. Sect. 8.1, to bring an arbitrary state $\hat{\rho}$ into a new state $\mathcal{M}\hat{\rho}$. Contrary to classical operations, not all positive maps correspond to realizable q-operations, but the completely positive ones. The operation's categorical linearity follows from the linearity of the procedure of mixing (4.5). Obviously we must arrive at the same q-state if we mix two states first and then we subject the systems of the resulting ensemble to the operation $\mathcal{M}$ or, alternatively, we perform the operation prior to the mixing the two ensembles together:

$$\mathcal{M}(w_1 \hat{\rho}_1 + w_2 \hat{\rho}_2) = w_1 \mathcal{M}\hat{\rho}_1 + w_2 \mathcal{M}\hat{\rho}_2 . \quad (4.6)$$

This is just the mathematical expression of the operation's linearity.

Selection of a given ensemble into specific sub-ensembles, a contrary process of mixing, will be possible via so-called selective q-operations. They correspond mathematically to trace-reducing completely positive maps, cf. Sect. 8.3. The most typical selective q-operations are called q-measurements 4.4.

4.3 Equation of motion

Dynamical evolution² of a closed q-system is determined by its hermitian Hamilton matrix $\hat{H}$. The von Neumann equation of motion takes this form:

² Our lectures use the Schrödinger-picture: the q-states $\hat{\rho}$ evolve with t , the q-physical quantities $\hat{A}$ do not.

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] . \quad (4.7)$$

For pure states, this is equivalent with the Schrödinger equation of motion:

$$\frac{d|\psi\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi\rangle . \quad (4.8)$$

Its solution $|\psi(t)\rangle \equiv \hat{U}(t)|\psi(0)\rangle$ represents the time-dependent unitary map $\hat{U}(t)$ of the Hilbert space. The von Neumann equation (4.7) implies the reversible q-operation $\mathcal{M}(t)$, which we can write formally as follows:

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t) \equiv \mathcal{M}(t)\hat{\rho}(0) . \quad (4.9)$$

4.4 Measurements

Consider a partition $\{\hat{P}_\lambda\}$ of the state space. The hermitian matrices $\hat{P}_\lambda$ are projectors on the Hilbert space, of eigenvalues 0 and 1. They form a complete orthogonal set:

$$\sum_\lambda \hat{P}_\lambda = \hat{I}, \quad \hat{P}_\lambda \hat{P}_\mu = \delta_{\lambda\mu} \hat{P}_\lambda . \quad (4.10)$$

We consider the projectors as binary q-physical quantities. Hermitian matrices $\hat{A}$ acting on the Hilbert space describe the whole variety of q-physical quantities. Each q-physical quantity $\hat{A}$ possesses the spectral expansion³:

$$\hat{A} = \sum_\lambda A_\lambda \hat{P}_\lambda; \quad \lambda \neq \mu \Rightarrow A_\lambda \neq A_\mu . \quad (4.11)$$

The real values A_λ are eigenvalues of the matrix, $\{\hat{P}_\lambda\}$ is a partition of the state space.

The projective partition (4.10) can be generalized. We define a positive decomposition of the unit matrix:

$$\hat{I} = \sum_n \hat{\Pi}_n; \quad \hat{\Pi}_n \geq 0 . \quad (4.12)$$

The elements of the positive decomposition, also called q-effects, are non-negative matrices $\hat{\Pi}_n$, they need be neither orthogonal nor projectors at all. They are, in a sense, the unsharp version of projectors.

³ Equivalent terminologies, like spectral or diagonal decomposition, or just diagonalization, are in widespread use.

4.4.1 Projective measurement

On each q-system in a statistical ensemble of q-state $\hat{\rho}$, we can measure the simultaneous values of the orthogonal projectors $\hat{P}_\lambda$ of a given partition (4.10). The outcomes are random. One of the binary quantities, say $\hat{P}_\lambda$, is 1 with probability

$$p_\lambda = \text{tr} \left(\hat{P}_\lambda \hat{\rho} \right) , \quad (4.13)$$

while the rest of them is 0:

$$\begin{array}{ccccccc} \hat{P}_1 & \hat{P}_2 & \dots & \hat{P}_{\lambda-1} & \hat{P}_\lambda & \hat{P}_{\lambda+1} & \dots \\ \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots \end{array} . \quad (4.14)$$

The state suffers projection according to the projection postulate of von Neumann and Lüders:

$$\hat{\rho} \rightarrow \hat{\rho}_\lambda \equiv \frac{1}{p_\lambda} \hat{P}_\lambda \hat{\rho} \hat{P}_\lambda . \quad (4.15)$$

The post-measurement q-state $\hat{\rho}_\lambda$ is also called conditional q-state, i.e., conditioned on the random outcome λ . As a result of the above measurement we have randomly selected the original ensemble of q-state $\hat{\rho}$ into sub-ensembles of q-states $\hat{\rho}_\lambda$ for $\lambda = 1, 2, \dots$

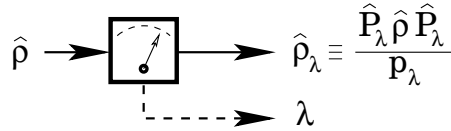


Fig. 4.1. Selective q-measurement. The ensemble of pre-measurement q-states $\hat{\rho}$ is selected into sub-ensembles of conditional post-measurement q-states $\hat{\rho}_\lambda$ according to the obtained measurement outcomes λ . The probability p_λ coincides with the trace of the unnormalized conditional q-state $\hat{P}_\lambda \hat{\rho} \hat{P}_\lambda$. The relative phases of the sub-ensembles $\hat{\rho}_\lambda$ have been irretrievably lost in a mechanism called decoherence.

The projective q-measurement is repeatable. Repeated q-measurements of the orthogonal projectors $\hat{P}_\mu$ on $\hat{\rho}_\lambda$ yield always the former outcomes $\delta_{\lambda\mu}$. The above selection is, contrary to the classical case, not reversible. Let us re-unite the obtained sub-ensembles, the post-measurement q-state becomes the following mixture of the conditional q-states $\hat{\rho}_\lambda$:

$$\sum_{\lambda} p_{\lambda} \hat{\rho}_{\lambda} = \sum_{\lambda} p_{\lambda} \frac{1}{p_{\lambda}} \hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda} = \sum_{\lambda} \hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda} ; \quad \neq \hat{\rho} , \quad (4.16)$$

which is in general not identical to the original pre-measurement q-state. The non-selective measurement realizes an irreversible⁴ q-operation $\mathcal{M}$:

$$\hat{\rho} \rightarrow \sum_{\lambda} \hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda} \equiv \mathcal{M} \hat{\rho} . \quad (4.17)$$

Q-irreversibility has its roots in the mechanism of *decoherence*⁵ which means the general phenomenon when superpositions get destroyed by q-measurement or by environmental interaction, cf. Chap. 8.

By the projective q-measurement of a general q-physical quantity $\hat{A}$ we mean the projective q-measurement of the partition (4.10) generated by its spectral expansion (4.11). The measured value of $\hat{A}$ is one of the eigenvalues:

$$\hat{A} \rightarrow A_{\lambda} , \quad (4.18)$$

the probability of the particular outcome is given by (4.13). The projective q-measurement is always repeatable. If a first measurement yielded A_{λ} on a given state then also the repeated measurement yields A_{λ} . We can define the non-selective measured value of $\hat{A}$, i.e., the average of A_{λ} taken with the distribution (4.13):

$$\langle \hat{A} \rangle \equiv \sum_{\lambda} p_{\lambda} A_{\lambda} = \text{tr} \left(\hat{A} \hat{\rho} \right) . \quad (4.19)$$

This is also called the expectation value of $\hat{A}$ in state $\hat{\rho}$.

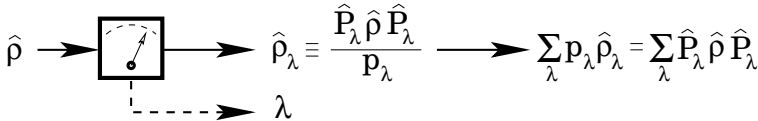


Fig. 4.2. Non-selective q-measurement. The sub-ensembles of conditional post-measurement q-states $\hat{\rho}_{\lambda}$ are re-united, contributing to the ensemble of non-selective post-measurement q-state which is, contrary to the classical case and because of decoherence, irreversibly different from the pre-measurement q-state $\hat{\rho}$.

4.4.2 Non-projective measurement

Non-projective q-measurement generalizes the projective one 4.4.1. On each q-system in a statistical ensemble of state $\hat{\rho}$, we can measure the simultaneous values

⁴ We use the notion of irreversibility as an equivalent to non-invertibility. We discuss the entropic-informatic notion of q-irreversibility in Sect. 10.8.

⁵ Cf., e.g., the monograph [10].

of the q-effects $\hat{I}_n$ of a given positive decomposition (2.11) but we lose repeatability of the measurement. The outcomes are random. One of the q-effects, say $\hat{I}_n$, is 1 with probability

$$p_n = \text{tr} \left(\hat{I}_n \hat{\rho} \right), \quad (4.20)$$

while the rest of them is 0:

$$\begin{array}{ccccccc} \hat{I}_1 & \hat{I}_2 & \dots & \hat{I}_{n-1} & \hat{I}_n & \hat{I}_{n+1} & \dots \\ \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots \end{array} \quad (4.21)$$

The state suffers the following change⁶:

$$\hat{\rho} \rightarrow \hat{\rho}_n \equiv \frac{1}{p_n} \hat{I}_n^{1/2} \hat{\rho} \hat{I}_n^{1/2}. \quad (4.22)$$

Contrary to the projective q-measurements, the repeated non-projective q-measurements yield different outcomes in general. The q-effects $\hat{I}_n$ are not binary quantities, the individual measurement outcomes 0 or 1 provide unsharp information that can only orient the outcome of subsequent measurements. The selective non-projective q-measurements are obviously irreversible: re-uniting the obtained sub-ensembles, i.e., averaging the post-measurement conditional q-states $\hat{\rho}_n$ yields $\sum_n \hat{I}_n^{1/2} \hat{\rho} \hat{I}_n^{1/2}$ which differs from the original pre-measurement q-state $\hat{\rho}$.

We can easily generalize the discrete set of q-effects for continuous sets. This generalization has a merit: one can construct the unsharp measurement of an arbitrarily chosen quantity $\hat{A}$. One constructs the following set of q-effects:

$$\hat{\Pi}_{\bar{A}} = \frac{1}{\sqrt{2\pi}\sigma^2} \exp \left[-\frac{(\bar{A} - \hat{A})^2}{2\sigma^2} \right], \quad -\infty \leq \bar{A} \leq \infty. \quad (4.23)$$

These q-effects correspond to the unsharp q-measurement of $\hat{A}$. The conditional post-measurement q-state will be $\hat{\rho}_{\bar{A}}(x) = p_{\bar{A}}^{-1} \hat{\Pi}_{\bar{A}}^{1/2} \hat{\rho} \hat{\Pi}_{\bar{A}}^{1/2}$, cf. eq. (4.22). We interpret $\bar{A}$ as the random outcome representing the measured value of $\hat{A}$ at the standard measurement error σ . The outcomes' probability (4.20) turns out to be the following distribution function:

$$p_{\bar{A}} = \text{tr} \left(\hat{\Pi}_{\bar{A}} \hat{\rho} \right), \quad (4.24)$$

normalized obviously by $\int p_{\bar{A}} d\bar{A} = 1$.

4.4.3 Continuous measurement

One may be interested in the features of what we can call the time-continuous q-measurement of a given quantity $\hat{A}$, performed by monitoring a *single* q-system.

⁶ Note that, in the q-literature, the post-measurement states are usually specified in a more general form $(1/p_n) \hat{U}_n \hat{I}_n^{1/2} \hat{\rho} \hat{I}_n^{1/2} \hat{U}_n^\dagger$, to include the arbitrary selective post-measurement unitary transformations $\hat{U}_n$.

Intuitively, we can consider q-measurements of $\hat{A}$ repeated at frequency $1/\Delta t$ and then we might take the infinite frequency limit. It is known that such a construction of continuous measurement does really work provided we are applying very unsharp measurements⁷. Their error σ must be proportional to their frequency $1/\Delta t$ of repetition. In the continuous limit, the rate

$$g = \lim_{\Delta t \rightarrow 0, \sigma \rightarrow \infty} \frac{1}{\Delta t \sigma^2} \quad (4.25)$$

is called the strength of the continuous measurement. We are going to show that under such mechanism of continuous measurement of $\hat{A}$ the reversible equation of motion (4.7) turns into the following irreversible q-master equation:

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] - \frac{g}{8}[\hat{A}, [\hat{A}, \hat{\rho}]] . \quad (4.26)$$

The new double-commutator term on the r.h.s. describes the decoherence caused by the continuous q-measurement. This is a special case of the general q-master equations presented in Sect. 8.6.

Let us see how this new term comes about. Note first that, in the asymptotics of the continuous limit (4.25), the error σ in the q-effects (4.23) can be expressed through the measurement strength g and the time-step Δt :

$$\hat{\Pi}_{\bar{A}} = \sqrt{\frac{g\Delta t}{2\pi}} \exp \left[-\frac{g}{2}(\sqrt{\Delta t}\bar{A} - \sqrt{\Delta t}\hat{A})^2 \right] . \quad (4.27)$$

Each time the unsharp measurement of $\hat{A}$ happens we can write the non-selective change of the q-state into the following form:

$$\hat{\rho} \rightarrow \int \hat{\Pi}_{\bar{A}}^{1/2} \hat{\rho} \hat{\Pi}_{\bar{A}}^{1/2} d\bar{A} . \quad (4.28)$$

Then we substitute the q-effects (4.27) and change the integration variable from $\bar{A}$ to $a = \sqrt{\Delta t}\bar{A}$, yielding the following expression for the post-measurement state:

$$\sqrt{\frac{g}{2\pi}} \int \exp \left[-\frac{g}{4}(a - \sqrt{\Delta t}\hat{A})^2 \right] \hat{\rho} \exp \left[-\frac{g}{4}(a - \sqrt{\Delta t}\hat{A})^2 \right] da . \quad (4.29)$$

Since we are interested in the continuous limit, we can restrict the above expression for the leading term in the small Δt , which yields $\hat{\rho} - (g/8)[\hat{A}, [\hat{A}, \hat{\rho}]]\Delta t$. Hence, in the continuous limit $\Delta t \rightarrow 0$, we arrive at the new double-commutator contribution to $d\hat{\rho}/dt$.

4.4.4 Compatible physical quantities

Let $\hat{A}$ and $\hat{B}$ be two arbitrary q-physical quantities. Consider their spectral expansions (4.11):

⁷ Cf. e.g. [13].

$$\hat{A} = \sum_{\lambda} A_{\lambda} \hat{P}_{\lambda} , \quad \hat{B} = \sum_{\mu} B_{\mu} \hat{Q}_{\mu} . \quad (4.30)$$

Let us measure both q-physical quantities, $\hat{A}$ first and then $\hat{B}$, in subsequent projective measurements. Write down the selective change of the q-state (4.16):

$$\hat{\rho} \rightarrow \frac{\hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda}}{p_{\lambda}} \rightarrow \frac{\hat{Q}_{\mu} \hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda} \hat{Q}_{\mu}}{p_{\mu\lambda}} . \quad (4.31)$$

where $p_{\mu\lambda}$ is the probability that $\hat{A} \rightarrow A_{\lambda}$ is the first, then $\hat{B} \rightarrow B_{\mu}$ is the second measurement outcome. Obviously, if we had performed the two measurements in the reversed order then the distribution of the measurement outcomes would in general be different. This we interpret in such a way that $\hat{A}$ and $\hat{B}$ are not simultaneously measurable. Yet, in some important cases, the measurement outcomes are independent of the order of measurements. The corresponding condition is that the matrices of the two q-physical quantities commute:

$$[\hat{P}_{\lambda}, \hat{Q}_{\mu}] = 0 \iff [\hat{A}, \hat{B}] = 0. \quad (4.32)$$

Q-physical quantities with commuting matrices are called compatible. On compatible q-physical quantities one can obtain simultaneous information. Two arbitrarily chosen q-physical quantities are, however, not compatible in general.

The matrices (projectors) of binary q-physical quantities (4.10), contributing to the spectral expansion (4.11) of a given q-physical quantity, will always commute, they are compatible, hence they can be measured simultaneously, cf. Sect. 4.4.1, in projective measurements.

Certain incompatible physical quantities can be measured in non-projective measurements. The q-effects of a positive decomposition (4.12) are not compatible in general, their non-negative matrices $\tilde{I}_n$ may not commute. Yet their simultaneous non-projective measurement 4.4.2 is possible though in restricted sense compared to projective measurements 4.4.1: we lose repeatability of the measurement.

4.4.5 Measurement in pure state

General rules of q-measurement (4.18,4.13) simplify significantly for a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$. The measurement outcome is still one of the eigenvalues A_{λ} . As for the pure state, it jumps into a conditional pure state $|\psi_{\lambda}\rangle$:

$$\hat{A} \rightarrow A_{\lambda} , \quad |\psi\rangle \rightarrow |\psi_{\lambda}\rangle \equiv \frac{1}{\sqrt{p_{\lambda}}} \hat{P}_{\lambda} |\psi\rangle , \quad (4.33)$$

while the probability distribution of the measurement outcome reads:

$$p_{\lambda} = \langle\psi| \hat{P}_{\lambda} |\psi\rangle . \quad (4.34)$$

The description of the process becomes even simpler if the matrix of the measured physical quantity is non-degenerate:

$$\hat{A} = \hat{A}^\dagger = \sum_{\lambda=1}^d A_\lambda \hat{P}_\lambda \iff \hat{P}_\lambda = |\varphi_\lambda\rangle \langle \varphi_\lambda| . \quad (4.35)$$

In this case, the state vector jumps into the unique eigenvector belonging to the measured eigenvalue:

$$\hat{A} \rightarrow A_\lambda, \quad |\psi\rangle \rightarrow |\varphi_\lambda\rangle . \quad (4.36)$$

And the probability distribution of the measurement outcome can be written as the squared modulus of the inner product between the new and the old state vectors:

$$p_\lambda = \langle \psi | \cdot | \varphi_\lambda \rangle \langle \varphi_\lambda | \cdot | \psi \rangle = | \langle \varphi_\lambda | \psi \rangle |^2 . \quad (4.37)$$

Most often, the description of the quantum measurement happens in such a way that we expand the pre-measurement pure state in terms of the eigenvectors of the physical quantity to be measured:

$$|\psi\rangle = \sum_{\lambda} c_\lambda |\varphi_\lambda\rangle . \quad (4.38)$$

In this representation the measurement takes place this way:

$$\hat{A} \rightarrow A_\lambda, \quad \begin{bmatrix} c_1 \\ \vdots \\ c_\lambda \\ \vdots \\ c_d \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} , \quad (4.39)$$

while the probability distribution of the measurement outcome reads:

$$p_\lambda = |c_\lambda|^2 . \quad (4.40)$$

4.5 Composite systems

The state space of the composite q-system, composed from the q-(sub)systems A and B , is the tensor product of the vector spaces of the q-subsystems, cf. (2.23):

$$\mathcal{H}_{AB} = C^{d_A} \otimes C^{d_B} = \{c_{\lambda l}; \lambda = 1, \dots, d_A; l = 1, \dots, d_B\} . \quad (4.41)$$

The pure state of the composite system is described by the normalized state vector of dimension $d_A d_B$:

$$|\psi_{AB}\rangle = \{c_{\lambda l}\} ; \quad \sum_{\lambda, l} |c_{\lambda l}|^2 = 1 . \quad (4.42)$$

If the two subsystems are uncorrelated then the state vector of the composite system is a tensor product:

$$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle = \{c_{A\lambda}c_{B\lambda}\} . \quad (4.43)$$

The mixed state of the composite system is described by the $d_A d_B \times d_A d_B$ -dimensional density matrix (4.3):

$$\hat{\rho}_{AB} = \{\rho_{(\lambda l)(\mu m)}\} . \quad (4.44)$$

The reduced state of subsystem A is obtained by tracing the composite q-system's state over the Hilbert space of subsystem B :

$$\hat{\rho}_A = \text{tr}_B \hat{\rho}_{AB} = \left\{ \sum_l \rho_{(\lambda l)(\mu l)} \right\} \equiv \mathcal{M} \hat{\rho}_{AB} . \quad (4.45)$$

Our notation indicates that a reduction, too, can be considered as an operation $\mathcal{M}$: it maps the states of the original q-system into the states of the q-subsystem. The state $\hat{\rho}_{AB}$ of the composite q-system is the tensor product of the q-subsystem's states if and only if there is no statistical correlation between the subsystems. But generally there is some, and then there can be *q-correlation* as well. Symbolically, we write:

$$\hat{\rho}_{AB} = \hat{\rho}_A \otimes \hat{\rho}_B + \text{cl. corr.} + \mathbf{q\text{-corr.}} , \quad (4.46)$$

which is different from the classical correlations (2.26). The q-correlations are absent if and only if the state of the composite q-system is separable. In other words, if it can be prepared as a statistical mixture of tensor product (uncorrelated) states⁸:

$$\hat{\rho}_{AB} = \sum_{\lambda} w_{\lambda} \hat{\rho}_{A\lambda} \otimes \hat{\rho}_{B\lambda} ; \quad w_{\lambda} \geq 0; \quad \sum_{\lambda} w_{\lambda} = 1 . \quad (4.47)$$

Then and only then there are no q-correlations but classical ones at most. In the contrary case, if $\hat{\rho}_{AB}$ can not be written into the above form, then the subsystems A and B are said to be in entangled composite state. Accordingly, q-correlation and *entanglement* mean exactly the same thing: the lack of classical separability.

The equation of motion of the composite system reads:

$$\frac{d}{dt} \hat{\rho}_{AB} = -\frac{i}{\hbar} [\hat{H}_{AB}, \hat{\rho}_{AB}] . \quad (4.48)$$

The composite Hamilton matrix is the sum of the Hamilton matrices of the subsystems themselves plus the interaction Hamilton matrix:

$$\hat{H}_{AB} = \hat{H}_A \otimes \hat{I}_B + \hat{I}_A \otimes \hat{H}_B + \hat{H}_{ABint} . \quad (4.49)$$

If $\hat{H}_{ABint}$ is zero then the tensor product initial state remains tensor product state, the dynamics does not create correlation between the q-subsystems. Non-vanishing $\hat{H}_{ABint}$ does usually create correlation. The motion of the whole q-system is reversible (unitary), of course. But that of the subsystems is not. In case of tensor

⁸ This definition of q-separability was introduced by Werner [14].

product initial state $\hat{\rho}_A(0) \otimes \hat{\rho}_B(0)$, for instance, the *reduced q-dynamics* of the subsystem A will represent the time-dependent irreversible q-operation $\mathcal{M}_A(t)$ which we can formally write as:

$$\hat{\rho}_A(t) = \text{tr}_B \left[\hat{U}_{AB}(t) \hat{\rho}_A(0) \otimes \hat{\rho}_B(0) \hat{U}_{AB}^\dagger(t) \right] \equiv \mathcal{M}_A(t) \hat{\rho}_A(0) . \quad (4.50)$$

The reversibility of the composite state q-dynamics has become lost by the reduction: the final reduced state $\hat{\rho}_A(t)$ does not determine a unique initial state $\hat{\rho}_A(0)$, cf. Sect. 8.2 for further discussion of reduced q-dynamics.

4.6 Collective system

The state (4.3) of a q-system is interpreted on the statistical ensemble of identical systems in the same state. We can form a multiple composite q-system from a big number n of such identical q-systems. This we call collective q-system, its state space is the n -fold tensor product of the elementary subsystem's vector spaces:

$$\mathcal{H} \otimes \mathcal{H} \otimes \dots \mathcal{H} \equiv \mathcal{H}^{\otimes n} . \quad (4.51)$$

The collective state reads:

$$\hat{\rho} \otimes \hat{\rho} \otimes \dots \hat{\rho} \equiv \rho^{\otimes n} , \quad (4.52)$$

while the state vector of a pure collective state is:

$$|\psi\rangle \otimes |\psi\rangle \otimes \dots |\psi\rangle \equiv |\psi\rangle^{\otimes n} . \quad (4.53)$$

If $\hat{A}$ is a q-physical quantity of the elementary subsystem then, in a natural way, one can introduce its arithmetic mean, over the n subsystems, as a collective q-physical quantity:

$$\frac{\hat{A} \otimes \hat{I}^{\otimes(n-1)} + \hat{I} \otimes \hat{A} \otimes \hat{I}^{\otimes(n-2)} + \dots + \hat{I}^{\otimes(n-1)} \otimes \hat{A}}{n} . \quad (4.54)$$

Collective q-physical quantities are not necessarily of such simple form. Their measurement is the collective q-measurement. Contrary to the classical theory, not all collective q-measurements can be reduced to independent measurements on the n subsystems, cf. Sects. 7.1.4, 10.6.

Problems, exercises

4.1. Decoherence-free projective measurement. There are special conditions to avoid decoherence. Let us prove that the non-selective measurement of a q-physical quantity $\hat{A}$ does not change the measured state $\hat{\rho}$ if and only if $[\hat{A}, \hat{\rho}] = 0$.

4.2. Mixing the eigenstates. Let us prove that a state given by the non-degenerate density matrix $\hat{\rho}$ can be prepared by mixing the pure eigenstates of $\hat{\rho}$. What mixing weights shall we use? How must we generalize the method if $\hat{\rho}$ is degenerate?

4.3. Separability of pure states. Let us prove that the pure state $|\psi_{AB}\rangle$ of a composite q-system is separable if and only if it takes the form $|\psi_A\rangle \otimes |\psi_B\rangle$.

4.4. Unitary cloning? We could try to duplicate the unknown pure state $|\psi\rangle$, cf. Sect. 5.3, of our q-system by cloning it to replace the prepared state $|\psi_0\rangle$ of another q-system with the same dimension of Hilbert space. Let us prove that the map $|\psi\rangle \otimes |\psi_0\rangle \rightarrow |\psi\rangle \otimes |\psi\rangle$ can not be unitary. Method: let us test whether the above transformation preserves the value of inner products.

5 Two-state q-system: qubit representations

Obviously the simplest q-systems are the two-state systems. Typical realization are an atom with its ground state and one of its excited states, a photon with its two polarization states, or an electronic spin with its “up” and “down” states. The smallest unit of q-information, i.e. the qubit, is an abstract two-state q-system. This Chapter is technical: you learn standard mathematics of a single abstract qubit.

5.1 Computational-representation

The Hilbert space of the two-state q-system is the complex 2-dimensional vector space (4.1). The notion of qubit is best realized in the computational basis. We introduce the computational basis vectors $|0\rangle$ and $|1\rangle$:

$$\{|x\rangle; x = 0, 1\}, \quad \sum_{x=0,1} |x\rangle \langle x| = \hat{I}, \quad \langle x'|x\rangle = \delta_{x'x}. \quad (5.1)$$

Also the primitive binary q-physical quantity $\hat{x}$ is defined in the computational basis:

$$\hat{x} = \sum_{x=0,1} x |x\rangle \langle x| = |1\rangle \langle 1|. \quad (5.2)$$

This is the (singular) 2×2 hermitian matrix of the qubit, as q-physical quantity. Its eigenvalues are 0 and 1. Often the q-state, rather than $\hat{x}$, is called the qubit. The generic pure state is a superposition of the basis vectors:

$$c_0 |0\rangle + c_1 |1\rangle \equiv \sum_{x=0,1} c_x |x\rangle; \quad |c_0|^2 + |c_1|^2 = 1. \quad (5.3)$$

This is why we say the qubit carries much richer information than the classical bit does, since the qubit can store the values 0 and 1 in parallel as well.

It is q-logical operations, see Sects. 6.1.1, 11.6, where the computational basis becomes indispensable. To explore the theoretical structure of a qubit, however, the well known physical representation is more convenient, cf. next Section.

5.2 Pauli representation

The mathematical models of all two-state q-systems are isomorphic to each other, regarding the state space and the physical quantities. Also the equations of motion are isomorphic for all closed two-state systems. Therefore the qubit formalism and language can be replaced by the terminology of any other two-state q-system. The electronic spin is the expedient choice. This is a genuine two-state q-system, its formalism is covariant for spatial rotations which guarantees conceptual and calculational advantage.

5.2.1 State space

The generic pure state of a two-state q-system takes the following form in a certain orthogonal basis of “up” and “down” states:

$$c_{\uparrow} |\uparrow\rangle + c_{\downarrow} |\downarrow\rangle ; \quad |c_{\uparrow}|^2 + |c_{\downarrow}|^2 = 1 . \quad (5.4)$$

An obvious parametrization takes normalization and the free choice of the complex phase into the account:

$$\cos \frac{\theta}{2} |\uparrow\rangle + e^{i\varphi} \sin \frac{\theta}{2} |\downarrow\rangle . \quad (5.5)$$

The angular parameters θ, φ can be identified with the standard directional angles of a 3-dimensional real unit vector $\mathbf{n}$. This way the above q-state can be parametrized by that unit vector itself:

$$|\mathbf{n}\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + e^{i\varphi} \sin \frac{\theta}{2} |\downarrow\rangle . \quad (5.6)$$

All pure states of a 2-dimensional q-system have thus been brought to a one-to-one correspondence with the surface points of the 3-dimensional unit-sphere called the Bloch sphere. The two basis q-states $|\uparrow\rangle, |\downarrow\rangle$ correspond to the north and south poles, respectively, on the Bloch sphere. Diametric points of the surface will always correspond to a pair of orthogonal q-states: $\langle -\mathbf{n} | \mathbf{n} \rangle = 0$. Hence $|\pm \mathbf{n}\rangle$ form a basis. Moreover, any basis can be brought to this form upto the phases of basis vectors. The modulus of the inner product of any two pure states $|\mathbf{n}\rangle, |\mathbf{m}\rangle$ is equal to the cosine of the half-angle between the two respective polarization vectors $\mathbf{n}, \mathbf{m}$:

$$|\langle \mathbf{m} | \mathbf{n} \rangle| = \cos \frac{\vartheta}{2} , \quad \cos \vartheta = \mathbf{m} \cdot \mathbf{n} . \quad (5.7)$$

For physical reasons, we call the 3-dimensional real Bloch vector $\mathbf{n}$ the polarization vector of the pure state $|\mathbf{n}\rangle$. As we shall see, the states $|\pm \mathbf{n}\rangle$ can be interpreted as the two eigenstates of the corresponding electronic spin-component matrix. As a q-physical quantity, the spin-vector $\hat{\sigma}/2$ of the electron was introduced by Pauli. Without the factor $1/2$, we call $\hat{\sigma}$ the vector of polarization. Its Cartesian components are the three Pauli matrices:

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (5.8)$$

The generic component of the polarization along the direction $\mathbf{n}$ reads:

$$\hat{\sigma}_n \equiv n_x \hat{\sigma}_x + n_y \hat{\sigma}_y + n_z \hat{\sigma}_z = \mathbf{n} \hat{\boldsymbol{\sigma}}. \quad (5.9)$$

Some basic characteristics of the Pauli matrices are the following:

$$\hat{\sigma}_a = \hat{\sigma}_a^\dagger, \quad \text{tr } \hat{\sigma}_a = 0, \quad [\hat{\sigma}_a, \hat{\sigma}_b] = 2i\epsilon_{abc} \hat{\sigma}_c, \quad \{\hat{\sigma}_a, \hat{\sigma}_b\} = 2\delta_{ab} \hat{I}. \quad (5.10)$$

Let us identify the former up-down basis states (5.4) by the “spin-up” and “spin-down” eigenstates belonging to the $+1$ and, respectively, to the -1 eigenvalues of the polarization component $\hat{\sigma}_z$:

$$|\uparrow\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\downarrow\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (5.11)$$

The polarization $\hat{\sigma}_n$ along the direction $\mathbf{n}$ has just $|\pm\mathbf{n}\rangle$ as eigenstates:

$$\hat{\sigma}_n |\pm\mathbf{n}\rangle = \pm |\pm\mathbf{n}\rangle. \quad (5.12)$$

This directly implies the spectral expansion of $\hat{\sigma}_n$:

$$\hat{\sigma}_n = |\mathbf{n}\rangle \langle \mathbf{n}| - |-\mathbf{n}\rangle \langle -\mathbf{n}|. \quad (5.13)$$

This is why we call and denote $|\mathbf{n}\rangle$ as the n -up state, while the vector $|-\mathbf{n}\rangle$ orthogonal to it we call and denote as the n -down state:

$$|\mathbf{n}\rangle \equiv |\uparrow \mathbf{n}\rangle, \quad |-\mathbf{n}\rangle \equiv |\downarrow \mathbf{n}\rangle. \quad (5.14)$$

When the reference direction is one of the Cartesian axes, we use the notations like $|\uparrow x\rangle, |\downarrow x\rangle, |\uparrow y\rangle, |\downarrow y\rangle$, while the notations of the distinguished z -axis may sometimes be omitted: $|\uparrow z\rangle \equiv |\uparrow\rangle, |\downarrow z\rangle \equiv |\downarrow\rangle$. Typically:

$$|\uparrow x\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{|\uparrow\rangle + |\downarrow\rangle}{\sqrt{2}} \quad (5.15)$$

$$|\uparrow y\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{|\uparrow\rangle + i|\downarrow\rangle}{\sqrt{2}}. \quad (5.16)$$

5.2.2 Rotational invariance

The general form of the 2×2 unitary matrices is, apart from an arbitrary complex phase, the following:

$$\hat{U}(\boldsymbol{\alpha}) \equiv \exp\left(-\frac{i}{2} \boldsymbol{\alpha} \hat{\boldsymbol{\sigma}}\right) = \hat{I} \cos \frac{\alpha}{2} - i \frac{\boldsymbol{\alpha} \hat{\boldsymbol{\sigma}}}{\alpha} \sin \frac{\alpha}{2}, \quad (5.17)$$

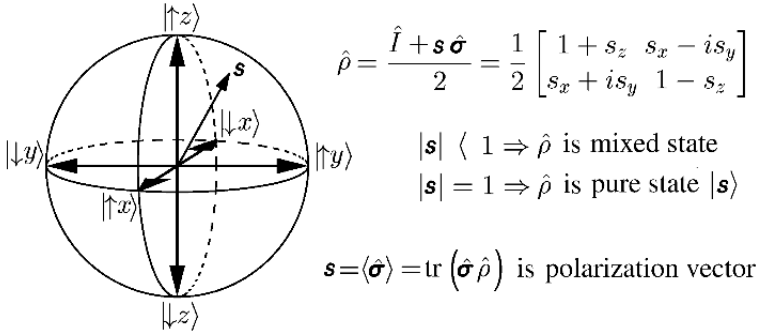


Fig. 5.1. Bloch sphere and density matrix. The set of all possible q-states of a qubit can be visualized by the points of a three-dimensional unit sphere of polarization vectors $\mathbf{s}$. Surface points ($|\mathbf{s}| = 1$) correspond to pure states $|\mathbf{s}\rangle$. North and south poles are conventionally identified with $|\uparrow\rangle, |\downarrow\rangle$, respectively, whereas the polar coordinates θ, φ of the unit vector $\mathbf{s}$ coincide with those in the orthogonal expansion (5.5) of pure states. Internal points ($|\mathbf{s}| < 1$) correspond to mixed states. The closer $\mathbf{s}$ is to the center of the sphere the more mixed is the corresponding state $\hat{\rho}$.

where the real α is called the vector of rotation. To interpret the name, let $R(\alpha)$ denote the orthogonal 3×3 matrix of spatial rotation along the direction α , by the angle $\alpha = |\alpha|$. It can be shown that the influence of the above unitary transformation $\hat{U}(\alpha)$ of the state vector is equivalent with the spatial rotation $R(\alpha)$ of the polarization vector:

$$\hat{U}(\alpha) |\mathbf{n}\rangle = |R^{-1}(\alpha)\mathbf{n}\rangle, \quad (5.18)$$

$$\hat{U}(\alpha) \hat{\boldsymbol{\sigma}} \hat{U}^\dagger(\alpha) = R(\alpha) \hat{\boldsymbol{\sigma}}. \quad (5.19)$$

This feature of rotational covariance makes the quick proof of many structural properties possible in Pauli representation.

5.2.3 Density matrix

The density matrix (4.4) which corresponds to the pure state $|\mathbf{n}\rangle$ of a two-state q-system takes this form:

$$|\mathbf{n}\rangle \langle \mathbf{n}| = \frac{\hat{I} + \mathbf{n} \cdot \hat{\boldsymbol{\sigma}}}{2}. \quad (5.20)$$

In the distinguished case of z -up pure state we write:

$$|\uparrow\rangle \langle \uparrow| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{\hat{I} + \hat{\sigma}_z}{2}. \quad (5.21)$$

From here, by rotation (5.18,5.19), we obtain the general equation (5.20). This can be generalized even further. If we allow the polarization vector $\mathbf{n}$ to have lengths shorter than 1 we can describe any mixed state as well:

$$\hat{\rho} = \frac{\hat{I} + s\hat{\sigma}}{2} ; \quad |s| \leq 1 . \quad (5.22)$$

The parameter s of the state is just the expectation value of the polarization vector $\hat{\sigma}$, as q-physical quantity, in the q-state $\hat{\rho}$:

$$s = \langle \hat{\sigma} \rangle = \text{tr}(\hat{\sigma}\hat{\rho}) . \quad (5.23)$$

5.2.4 Equation of motion

The general Hamilton matrix of a two-state system is the following:

$$\hat{H} = \frac{\hbar}{2} \boldsymbol{\omega} \hat{\sigma} , \quad (5.24)$$

where $\boldsymbol{\omega}$ is the vector of external magnetic field provided we identify the system as the electronic spin (and its giro-magnetic factor has been “absorbed” into the scale of the magnetic field). The von Neumann equation of motion (4.7) takes this form:

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{2} \boldsymbol{\omega} [\hat{\sigma}, \hat{\rho}] . \quad (5.25)$$

It implies the classical magnetic dipole precession for the expectation value (5.23) of the polarization vector:

$$\frac{ds}{dt} = \boldsymbol{\omega} \times \mathbf{s} . \quad (5.26)$$

5.2.5 Physical quantities, measurement

All q-physical quantities in two-state systems are real linear combinations of the unit-matrix $\hat{I}$ and the three Pauli matrices $\hat{\sigma}$:

$$\hat{A} = \hat{A}^\dagger = a_0 \hat{I} + \mathbf{a} \hat{\sigma} , \quad (5.27)$$

where a_0 is a real number, $\mathbf{a}$ is a real vector. Hence the commutator of two such q-physical quantities is:

$$[\hat{A}, \hat{B}] = 2i (\mathbf{a} \times \mathbf{b}) \cdot \hat{\sigma} , \quad (5.28)$$

in obvious notation. This vanishes only when $\mathbf{a}$ and $\mathbf{b}$ are parallel vectors. Then, however, the two physical quantities $\hat{A}$ and $\hat{B}$ are simply functions of each other. It is clear that a two-state system has no non-trivial compatible physical quantities 4.4.4. The maximum compatible set is a single polarization $\hat{\sigma}_n$ in one direction $\mathbf{n}$, or the two orthogonal projectors $\hat{P}_{\pm n} = |\pm \mathbf{n}\rangle \langle \pm \mathbf{n}|$ contributing to its spectral expansion

(5.13). The three q-physical quantities $\hat{\sigma}_n, \hat{P}_n, \hat{P}_{-n}$ are functions of each other. The q-measurement of any of them is in all respects equivalent with the measurement of any other one.

Consider a given pure state (5.6):

$$|\uparrow z'\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + e^{i\varphi} \sin \frac{\theta}{2} |\downarrow\rangle, \quad (5.29)$$

which is a spin in direction z' of polar angle θ . Suppose the selective measurement of the distinguished polarization $\hat{\sigma}_z$ (4.36, 4.37):

$$|\uparrow z'\rangle \rightarrow \begin{cases} |\uparrow\rangle, & \hat{\sigma}_z \rightarrow +1, \quad p_\uparrow = \cos^2(\theta/2) \\ |\downarrow\rangle, & \hat{\sigma}_z \rightarrow -1, \quad p_\downarrow = \sin^2(\theta/2) \end{cases}. \quad (5.30)$$

We can say that the initial spin of polar angle θ has jumped into the vertical direction with probability $\cos^2(\theta/2)$. Exploiting the rotational invariance, we resume that a pure state polarized in direction $\mathbf{n}$ will, when a certain $\hat{\sigma}_m$ is selectively measured, jump into the new direction $\mathbf{m}$ with transition probability

$$p(\mathbf{n} \rightarrow \mathbf{m}) = |\langle \mathbf{m} | \mathbf{n} \rangle|^2 = \cos^2(\vartheta/2); \quad \cos \vartheta = \mathbf{n} \mathbf{m}, \quad (5.31)$$

which is the cosine square of half the angle between the two directions $\mathbf{n}$ and $\mathbf{m}$. This probability itself turns out to be symmetric between the initial and final states. We shall see in Sect. 6.2.4 of the next Chapter that such transition probabilities can be chosen as a measure of likeness — the *fidelity* — between two states in general.

5.3 The unknown qubit, Alice and Bob

Because of the irreversibility of q-measurements an unknown q-state represents an issue much different from that of an unknown classical state. For various purposes in the forthcoming Chapters, we need the precise notion of the unknown qubit. Interestingly, there is no theoretical standard for the totally unknown mixed q-states but for the subset of the pure states. In this sense define we the unknown (random) qubit as a pure state $|\mathbf{n}\rangle$ whose polarization $\mathbf{n}$ is totally random over the 4π solid angle. It should not be confused with the totally mixed state $\hat{\rho} = \hat{I}/2$ which would correspond to the average state of the random qubits.

When we notice it, the concept of unknown qubit may be used in a particular sense. We can suppose more knowledge and less ignorance. To make the definitions more operational, it is common to personalize the one who knows the state and the other one who does not. For instance, we shall assume that Alice prepares and gives Bob a qubit in one of the two non-orthogonal states:

$$|\uparrow z\rangle \quad \text{or} \quad |\uparrow x\rangle, \quad (5.32)$$

between which she has decided upon her tossing a coin. That is all that also Bob knows but he does not know which of the two states he has actually received. We say that Bob has an unknown qubit in this sense. If interested, he must perform a q-measurement.

5.4 Relationship of computational and Pauli representations

The computational 5.1 and the Pauli-representations 5.2 can be transformed into each other by the conventional mapping of the corresponding bases:

$$|0\rangle = |\uparrow\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = |\downarrow\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (5.33)$$

The elemental qubit physical quantity (5.2) is in simple algebraic relationship with the polarization component $\hat{\sigma}_z$:

$$\hat{x} = \frac{\hat{I} - \hat{\sigma}_z}{2}; \quad \hat{\sigma}_z = \hat{I} - 2\hat{x}. \quad (5.34)$$

The density matrices of the computational basis states can also be expressed by the polarization $\hat{\sigma}_z$:

$$|0\rangle\langle 0| = \frac{\hat{I} + \hat{\sigma}_z}{2}; \quad |1\rangle\langle 1| = \frac{\hat{I} - \hat{\sigma}_z}{2}. \quad (5.35)$$

The same thing in compact form reads:

$$|x\rangle\langle x| = \frac{\hat{I} + (1 - 2x)\hat{\sigma}_z}{2}; \quad x = 0, 1. \quad (5.36)$$

Problems, exercises

5.1. Pure state fidelity from density matrices. Let us show that the cosine rule (5.7) for the inner product $\langle m|n \rangle$ of pure states can be derived in a single step starting from the corresponding two density matrices. Method: let us express $|\langle m|n \rangle|^2$.

5.2. Unitary rotation for $|\uparrow\rangle \longrightarrow |\downarrow\rangle$. What can be the rotation vector α that rotates the state $|\uparrow\rangle$ into the orthogonal state $|\downarrow\rangle$? Let us make a simple choice for the rotation axis! Calculate the matrix of the corresponding unitary rotation $\hat{U}(\alpha)$ and verify the result.

5.3. Density matrix eigenvalues and -states in terms of polarization. Let us express the two eigenvalues and the two eigenvectors (i.e. the spectral expansion) of a density matrix in function of the polarization vector s .

5.4. Magnetic rotation for $|\uparrow\rangle \longrightarrow |\downarrow\rangle$. Determine the (constant) magnetic field ω that rotates the electronic spin from state $|\downarrow\rangle$ into state $|\uparrow\rangle$. How long time must the field be switched on?

5.5. Interrelated q-bit physical quantities. Write down all pairwise relationships for the matrices $\hat{\sigma}_n, \hat{P}_n, \hat{P}_{-n}$.

5.6. Mixing non-orthogonal polarizations. Suppose we mix the pure states $|\uparrow z\rangle$ and $|\uparrow x\rangle$ together at rate $1/3 : 2/3$. Write down the polarization vector s of the resulting mixed state.

6 One-qubit manipulations

Numerous principles and methods of q-theory as well as q-information theory can already be demonstrated on a two-state q-system, i.e., on a single qubit. Entanglement can certainly not, we shall present it on composite systems in Chap. 7. The present Chapter teaches the elements of q-state manipulation and various simple instances of the q-informatic approach.

6.1 One-qubit operations

We present the reversible logical operations on single qubits, whose combinations can represent all unitary one-qubit operations. Then we discuss the example of depolarization q-operation generated by polarization measurements. We learn the notorious polarization reflection which is not a q-operation: by classical analogy it were realizable but it turns out not to be so. Let us note that the combinations of one-qubit unitary operations and projective measurements can not represent all possible one-qubit q-operations. These become available if the qubit forms an interacting composite system with an environmental q-system, cf. Chap. 8.

6.1.1 Logical operations

In q-physics, coherent operations are reversible, unitary transformations. If the coherent operation is applied to an *unknown* q-state then the state remains invariably unknown after the operation. For qubits, the unitary transformations have been mapped onto 3-dimensional rotations (5.17). Now we are going to consider those rotations which play a distinguished role for the logical operations on qubits. The Pauli matrices (5.8) themselves, too, are unitary. They are conventionally denoted by $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ whenever qubit-operations are considered, and the corresponding operations are called X -, Y -, Z -operations.

Among the three (non-trivial) classical one-bit logical operations the NOT is the only reversible one. Hence, only the NOT possesses a q-analogue, namely, the X operation. The Z -operation and the H -operation (Hadamard operation) are further one-qubit operations that have no classical counterparts. The reason is that the Z -operation inverts the relative phase between the basis states $|0\rangle$ and $|1\rangle$, while the H -operation brings both $|0\rangle$ and $|1\rangle$ into their superpositions. Phases and superpositions

of $|0\rangle$ and $|1\rangle$ can not be interpreted classically. The elemental one-qubit logical operations and their influence are then the following:

$$\hat{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{Bmatrix} |0\rangle \\ |1\rangle \end{Bmatrix} \rightarrow \boxed{X} \rightarrow \begin{Bmatrix} |1\rangle \\ |0\rangle \end{Bmatrix} \quad (6.1)$$

$$\hat{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{Bmatrix} |0\rangle \\ |1\rangle \end{Bmatrix} \rightarrow \boxed{Z} \rightarrow \begin{Bmatrix} |0\rangle \\ -|1\rangle \end{Bmatrix} \quad (6.2)$$

$$\hat{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \begin{Bmatrix} |0\rangle \\ |1\rangle \end{Bmatrix} \rightarrow \boxed{H} \rightarrow \begin{Bmatrix} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{Bmatrix}. \quad (6.3)$$

Like all coherent (unitary) operations, also the above ones are reversible. All three one-qubit operations are identical to their inverses, respectively.

The X, Z, H operations can be combined, e.g. $\hat{X}\hat{Z}\hat{H}\hat{X}$ is a combined unitary operation. Whether all unitary one-qubit operations can be realized by combining a few basic ones? The above three operations are not sufficient. Extend, nonetheless, the half-turn around the z -axis (Z) by the rotation at an arbitrary angle φ , called the phase-operation:

$$\hat{T}(\varphi) = \begin{bmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{bmatrix}, \quad \begin{Bmatrix} |0\rangle \\ |1\rangle \end{Bmatrix} \rightarrow \boxed{T(\varphi)} \rightarrow \begin{Bmatrix} e^{-i\varphi/2} |0\rangle \\ e^{i\varphi/2} |1\rangle \end{Bmatrix}. \quad (6.4)$$

Now, consider the combined operation $\hat{H}\hat{T}(\theta)\hat{H}$ which is a rotation of the Bloch sphere around the x -axis by the angle θ . Let us follow the basis state $|0\rangle$ under the subsequent transformations $\hat{H}$, $\hat{T}(\theta)$ and $\hat{H}$:

$$|0\rangle \rightarrow \frac{|0\rangle + |1\rangle}{\sqrt{2}} \rightarrow \frac{e^{-i\theta/2} |0\rangle + e^{i\theta/2} |1\rangle}{\sqrt{2}} \rightarrow \cos \frac{\theta}{2} |0\rangle - i \sin \frac{\theta}{2} |1\rangle. \quad (6.5)$$

The resulting state corresponds to a Bloch vector with polar coordinates θ and $-\pi/2$. We do a further rotation by an angle $\pi/2$ plus φ around the z -axis. The cumulative influence of the four logical q-operations is the following:

$$\hat{T}(\tfrac{1}{2}\pi + \varphi)\hat{H}\hat{T}(\theta)\hat{H}|0\rangle = e^{-i(\varphi/2)-i(\pi/4)} \left(\cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \right). \quad (6.6)$$

Apart from an irrelevant complex phase, the r.h.s. shows the generic pure state (5.6) of a qubit. Furthermore, it can be proved that by the combinations of the operations $\hat{H}$ and $\hat{T}(\varphi)$ one can perform all one-qubit unitary operations. Therefore we call the $\hat{H}$ and the $\hat{T}(\varphi)$ a set of *universal* one-qubit operations. This becomes important for the theory of q-computation in Chap. 11.

6.1.2 Depolarization, re-polarization, reflection

The q-measurement is the most important operation among the irreversible ones. Its irreversibility is related to that we have acquired *hidden* q-information from a q-state

and we can not completely re-supply it into the state, because of decoherence. More general irreversible operations can be obtained if we combine q-measurements and unitary operations. It is important that the mapping of the density matrix remains linear and information on the state, cf. Chap. 10 for details, can only be obtained at the expense of irreversibility and decoherence.

We shall see later in Sect. 6.2.5 that there is a naive measurement strategy to determine a single unknown qubit. This naive strategy is just a single measurement of polarization $\hat{\sigma}_n$ along a random direction $\mathbf{n}$. The average post-measurement state of the qubit becomes mixed. For a general state, the post-measurement polarization becomes $1/3$ smaller than the original polarization:

$$\hat{\rho} \equiv \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}}{2} \rightarrow \mathcal{M}\hat{\rho} \equiv \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}/3}{2}. \quad (6.7)$$

The random polarization measurement leads to isotropic depolarization. The corresponding operation is obviously linear.

Consider now the opposite case and suppose that we wish to bring an unknown partially polarized state into a totally polarized one, keeping the direction of polarization invariant:

$$\hat{\rho} \equiv \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}}{2} \rightarrow \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}/|\mathbf{s}|}{2}, \text{ i.e.: } \mathbf{s} \rightarrow \frac{\mathbf{s}}{|\mathbf{s}|}. \quad (6.8)$$

This would mean non-linear transformation for the polarization vector as well as for the density matrix. Hence the above re-polarization of the unknown qubit would contradict the foundations of q-theory and, as a consequence, the corresponding operation does not exist.

Consider finally a map $\mathcal{T}$, also called spin inversion or reflection, which is doubtlessly linear:

$$\hat{\rho} \equiv \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}}{2} \rightarrow \mathcal{T}\hat{\rho} \equiv \frac{\hat{I} - \mathbf{s}\hat{\boldsymbol{\sigma}}}{2}, \text{ i.e.: } \mathbf{s} \rightarrow -\mathbf{s}. \quad (6.9)$$

This map reflects the polarization of an unknown state. On pure states this map is $|\mathbf{n}\rangle \rightarrow |-\mathbf{n}\rangle$. Such apparently innocent operation, which would classically mean the mere inversion of the electron's angular motion, does not exist in q-physics, it can not be realized. The reason is that, in general, an isolated q-system is correlated with the other q-systems of the world by classical *and* q-correlations. These q-correlations, i.e. these entanglements, will cause that not all linear maps are real operations on the space of q-states. For such inhibited operations the above linear map represents a typical instance. However, the impossibility of the operation can only be explained later in Sect. 7.2.2 by the mechanism of entanglement which assumes composite systems.

6.2 State preparation, determination

We start with standard applications and proceed to the paradoxical “no-cloning” theorem of q-theory¹. From this unusual property we derive two applications where q-physics guarantees such security of information processing that classical physics can never do.

6.2.1 Preparation of known state, mixing

A certain demanded state, e.g., the pure state $|\uparrow\rangle$ polarized in the z -direction, can be prepared from an unknown state if we simply measure the polarization $\hat{\sigma}_z$. If the raw state was a totally random pure state then in half of the cases the measurement outcome is $+1$ and the resulting state is the demanded $|\uparrow\rangle$. On average, in the other half of the cases the outcome is -1 . Then we rotate the polarization of $|\downarrow\rangle$ by 180 degrees via the unitary transformation $\hat{\sigma}_x$ (5.17):

$$|\downarrow\rangle \longrightarrow \hat{\sigma}_x |\downarrow\rangle = |\uparrow\rangle, \quad (6.10)$$

so that again we obtain the desired state $|\uparrow\rangle$. The pure states prepared in the above way on demand will serve for preparation of arbitrarily prescribed mixed states.

A generic mixed state (5.22) can be expanded in this form:

$$\hat{\rho} = \frac{\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}}}{2} = \frac{1 + |\mathbf{s}|}{2} |\uparrow s\rangle \langle \uparrow s| + \frac{1 - |\mathbf{s}|}{2} |\downarrow s\rangle \langle \downarrow s|. \quad (6.11)$$

We can thus mix the desired state $\hat{\rho}$ from s -up and s -down² pure states (5.12). But we can do it from other two or more pure states. Even mixed states can be used as raw material for mixing. Let us find the general rule! The basic equation (4.5) of mixing two q-states yields the following form for two-state density matrices (5.22):

$$\hat{I} + \mathbf{s}\hat{\boldsymbol{\sigma}} = w_1(\hat{I} + \mathbf{s}_1\hat{\boldsymbol{\sigma}}) + w_2(\hat{I} + \mathbf{s}_2\hat{\boldsymbol{\sigma}}). \quad (6.12)$$

Thus the relationship of the polarization vectors must be this:

$$\mathbf{s} = w_1\mathbf{s}_1 + w_2\mathbf{s}_2. \quad (6.13)$$

The rule can be generalized for multiple weighted mixing:

$$\hat{\rho} = \sum_{\lambda} w_{\lambda} \hat{\rho}_{\lambda} \iff \mathbf{s} = \sum_{\lambda} w_{\lambda} \mathbf{s}_{\lambda}. \quad (6.14)$$

The polarization vector of the mixture is identical to the weighted mean of polarization of the components. The Bloch vector of the mixture can therefore be anywhere within the convex hull of the components Bloch vectors, depending on the distribution of the weights of mixing. Hence a pure state can never be obtained from mixing

¹ Cf. [15] by Wootters and Zurek.

² The notations $|\uparrow s\rangle, |\downarrow s\rangle$ mean $|\uparrow \mathbf{n}\rangle, |\downarrow \mathbf{n}\rangle$, resp., where $\mathbf{n} = \mathbf{s}/|\mathbf{s}|$.

any other states. True mixed states can be prepared by mixing pure or even mixed states in infinitely many different ways.

The completely depolarized (rotationally invariant) mixed state, i.e., the ensemble behind it, can be mixed from the z -up/down pure states:

$$\hat{\rho} \equiv \frac{1}{2} \hat{I} = \left\{ \begin{array}{l} 50\% \uparrow z \\ 50\% \downarrow z \end{array} \right\}, \quad (6.15)$$

as well as from the x -up/down pure states:

$$\hat{\rho} \equiv \frac{1}{2} \hat{I} = \left\{ \begin{array}{l} 50\% \uparrow x \\ 50\% \downarrow x \end{array} \right\}. \quad (6.16)$$

Recall Sect. 4.2 that after mixing it is totally impossible to distinguish which way the mixed state was prepared. It is crucial, of course, that mixing always means a probabilistic one. To prepare, e.g., the ensemble (6.15), we can use repeated coin tossing to draw the z -up/down states for the mixture.

The mixed state is uniquely defined by its density matrix $\hat{\rho}$.

For a given mixed state (i.e.: for the corresponding ensemble) there no longer exists any test to distinguish which way the mixing was done or, which way the mixed state was altogether created.

All is the same for classical densities ρ . There is a cardinal departure, however. The decomposition of a classical state (ensemble) into the mixture of pure states is always unique. Differently from the classical mixed states, a non-pure q-state can be decomposed into various mixtures of pure q-states.

Table 6.1. Uniqueness of state, non-uniqueness of mixing.

6.2.2 Ensemble determination of unknown state

The determination of a completely unknown (not necessarily pure) qubit state $\hat{\rho}$ is equivalent to the determination of the polarization vector $\mathbf{s}$. The expectation value of the physical quantity $\mathbf{s}$ parametrizes the state $\hat{\rho}$:

$$\hat{\rho} = \frac{\hat{I} + \mathbf{s} \hat{\boldsymbol{\sigma}}}{2}, \quad \mathbf{s} = \text{tr}(\hat{\boldsymbol{\sigma}} \hat{\rho}) = ? \quad (6.17)$$

The three components of $\hat{\boldsymbol{\sigma}}$ are the three Pauli matrices (5.8). One must measure them on the given state $\hat{\rho}$. Since they do not commute, i.e., are not compatible, one must measure them in separate measurements. If the q-ensemble representing the state $\hat{\rho}$ consists of N elements then it is possible to consume cca. $N/3$ elements for the measurement of each Pauli matrix. The outcome of each measurement is

either $+1$ or -1 ; up or down in other words. The obtained statistics allows for the estimation of the polarization components:

$$s_x \approx \frac{N_{\uparrow x} - N_{\downarrow x}}{N_{\uparrow x} + N_{\downarrow x}}, \quad s_y \approx \frac{N_{\uparrow y} - N_{\downarrow y}}{N_{\uparrow y} + N_{\downarrow y}}, \quad s_z \approx \frac{N_{\uparrow z} - N_{\downarrow z}}{N_{\uparrow z} + N_{\downarrow z}}, \quad (6.18)$$

and, from them, we can write down the estimation of $\hat{\rho}$. The statistical error is concomitant of the method, it would only disappear at infinite N .

The above method is far from being optimum. The problem is open because, first of all, there is no natural definition of “completely” unknown q-state. The situation is simpler if we know at least that the state is pure, i.e. it can be described by a certain state vector $|\mathbf{n}\rangle$. The unknown pure state has the natural definition: we suppose it is the random complex unit vector in the Hilbert space. For two-state systems, equivalently, we supposed in Sect. 5.3 that the polarization unit-vector $\mathbf{n}$ is a perfectly random unit vector on the Bloch sphere.

6.2.3 Single state determination: no-cloning

An unknown q-state can only be determined if we have access to a large number of systems in the same state. If we possess but a single system we would first try copying it, multiplying it. If multiplication were successful, we would produce any number of clones and using this ensemble we would determine the state. In two-state case, we would deliberately reduce the statistical error of polarization estimation.

In reality, however, an unknown q-state can not be cloned. In order of an indirect proof, let us consider a completely mixed q-ensemble of two-state systems. Draw an element of it at random and suppose that we can copy and multiply it in N examples. If the ensemble had been mixed from z -up/down states (6.15) then we have N copies either of z -up or z -down states in our hands. For very large N , this method allows us to explore, with high reliability, which states the depolarized ensemble had been mixed from. In particular, the alternative x -up/down states can be excluded. This would, however, contradict to the principle that a q-ensemble is uniquely and exhaustively determined by its density matrix.

In summary, the unknown q-state of a single physical system can never be implemented on the other single system so that it, *too*, bear the same q-state. As a consequence, the q-information hidden in the unknown state of a single system can never be copied to the other single system’s q-state as well. No carbon copy, no emergency copy can be produced ever.

Suppose, for instance, Alice prepares a single q-system in a pure state. If she does not tell Bob of the direction of polarization then she may safely trust the q-system to Bob since he will never be able to make a copy for himself. If Bob makes an attempt, the copy will usually be faint and Alice will notice the fraud when she gets back either her copy or the fake one, see Sect. 6.2.5.

6.2.4 Fidelity of two states

When a given q-state $\hat{\rho}$ is approximated by another state $\hat{\rho}'$ we may need to measure the quality of the approximation. This purpose is served by the fidelity $F(\hat{\rho}, \hat{\rho}')$. If

both states are pure then the definition is this:

$$F = |\langle \psi | \psi' \rangle|^2 = \text{tr}(\hat{\rho} \hat{\rho}') , \quad (6.19)$$

i.e., their fidelity is equal to the squared modulus of their inner product. Fidelity has simple statistical interpretations for pure states. Suppose, for instance, that Bob asks Alice for a certain pure state $|\psi\rangle$, but she passes another pure state $|\psi'\rangle$ instead. When Bob, by the projective measurement of $|\psi\rangle\langle\psi|$, verifies whether he has received the demanded state $|\psi\rangle$ he will find the received state faithful to the demanded one just with the probability F .

For qubits $|\mathbf{n}\rangle, |\mathbf{m}\rangle$, fidelity becomes equal to the squared-cosine of the half-angle between the two polarization vectors:

$$F = |\langle \mathbf{m} | \mathbf{n} \rangle|^2 = \cos^2(\theta/2) ; \quad \cos(\theta) = \mathbf{m} \cdot \mathbf{n} . \quad (6.20)$$

6.2.5 Approximate state determination and cloning

Let us calculate the average fidelity between an unknown qubit $|\mathbf{n}\rangle$ and its faint copy $|\mathbf{m}\rangle$. Suppose Alice passes a state $|\mathbf{n}\rangle = |\uparrow\rangle$ to Bob. Bob needs a copy. He does not know the state and, for lack of a better idea, he measures the polarization $\hat{\sigma}_{z'}$ at a randomly chosen angle z' :

$$|\uparrow\rangle \rightarrow \begin{cases} |\uparrow z'\rangle, & \hat{\sigma}_{z'} \rightarrow +1, \quad p_{\uparrow} = \cos^2(\theta/2) \\ |\downarrow z'\rangle, & \hat{\sigma}_{z'} \rightarrow -1, \quad p_{\downarrow} = \sin^2(\theta/2) \end{cases} \quad (6.21)$$

and he prepares his copy $|\mathbf{m}\rangle$ after the resulting state $|\uparrow z'\rangle$ or $|\downarrow z'\rangle$ according to the measurement outcome ± 1 . The average fidelity of the copy is:

$$p_{\uparrow} |\langle \uparrow \uparrow z' \rangle|^2 + p_{\downarrow} |\langle \uparrow \downarrow z' \rangle|^2 . \quad (6.22)$$

Since Bob could only choose the direction z' at random, the expected fidelity of the copy will be the isotropic average over all directions, that is $2/3$. What does it mean? If Alice gets the copy back to check it, obviously she measures $\hat{\sigma}_z$ and expects the outcome $\hat{\sigma}_z \rightarrow +1$. The probability of this result is just equal to the above calculated fidelity. Hence Alice confirms the copy with 66% probability whereas with 33% probability she obtains the measurement outcome $\hat{\sigma}_z \rightarrow -1$ meaning that $1/3$ of the copies unveils as completely wrong. Prior to Alice's test, however, it makes no sense to say that the copy is completely true or completely flawed.

6.3 Indistinguishability of two non-orthogonal states

In a particular special case of single state determination, one knows apriori that the unknown state is drawn from two known states. The strategy is exceptionally simple when the two states are orthogonal pure states, e.g., we have to decide between

$$|\uparrow n\rangle \text{ or } |\downarrow n\rangle , \quad (6.23)$$

where n is known. Then a single measurement of $\hat{\sigma}_n$ provides the answer. If, on the contrary, the possible states are not orthogonal, e.g.:

$$|\uparrow z\rangle \text{ or } |\uparrow x\rangle, \quad (6.24)$$

then there, of course, does not exist any physical quantity such that its single measurement would provide the safe answer. Between two non-orthogonal states it is impossible to decide with full certainty. This is what q-banknote 6.4.1 and q-cryptography 6.4.2 will be based on. Decision of limited reliability is still possible. In general, the single measurement strategy can be optimized for various aspects like, e.g., a higher ratio of perfect conclusion, a higher fidelity 6.2.4, or increased accessible information 10.4.

6.3.1 Distinguishing via projective measurement

Let our single qubit have either of the two states $|\uparrow z\rangle, |\uparrow x\rangle$ at 50 – 50% apriori likelihoods. We can try a polarization measurement either in direction z or x , can make a choice at equal rate. At 75% likelihood the outcome will be 1 and at 25% the outcome will be -1 . In the former case the measurement is not conclusive. In the latter case the measurement is perfectly conclusive: if we happened to measure $\hat{\sigma}_z = -1$ then the pre-measurement state could not be the $|\uparrow z\rangle$ but $|\uparrow x\rangle$. And if we measured $\hat{\sigma}_x = -1$ then the pre-measurement state must have been the $|\uparrow z\rangle$. The probability of conclusive answer is 25% only. By using non-projective measurement, a higher ratio of conclusive answers can be achieved³, as it is shown in the next Section.

6.3.2 Distinguishing via non-projective measurement

Let us introduce the following positive decomposition (4.12):

$$\hat{\Pi}_z = w |\downarrow x\rangle \langle \downarrow x|, \quad \hat{\Pi}_x = w |\downarrow z\rangle \langle \downarrow z|, \quad \hat{\Pi}_? = \hat{I} - \hat{\Pi}_z - \hat{\Pi}_x. \quad (6.25)$$

If the non-projective measurement yields $\hat{\Pi}_z \rightarrow 1$ or $\hat{\Pi}_x \rightarrow 1$ then the pre-measurement state must have been $|\uparrow z\rangle$ or $|\uparrow x\rangle$, respectively. If the outcome is $\hat{\Pi}_? \rightarrow 1$ then no information is gained. The likelihood of such inconclusive cases follows from (4.20):

$$p_? = \text{tr} \left[\hat{\Pi}_? \frac{|\uparrow z\rangle \langle \uparrow z| + |\uparrow x\rangle \langle \uparrow x|}{2} \right] = 1 - \frac{w}{2}. \quad (6.26)$$

provided the two non-orthogonal states (6.24) had equal apriori probabilities. The non-projective measurement becomes optimal for the maximum weight w :

$$w = \frac{\sqrt{2}}{1 + \sqrt{2}}, \quad (6.27)$$

³ Cf. [16] by Ivanovic.

i.e., when the effect $\hat{\Pi}_?$ becomes semi-definite:

$$\hat{\Pi}_z = \frac{\sqrt{2}}{1 + \sqrt{2}} |\downarrow x\rangle \langle \downarrow x| , \quad \hat{\Pi}_x = \frac{\sqrt{2}}{1 + \sqrt{2}} |\downarrow z\rangle \langle \downarrow z| , \quad \hat{\Pi}_? = \frac{2}{1 + \sqrt{2}} |\uparrow n\rangle \langle \uparrow n| , \quad (6.28)$$

where

$$\mathbf{n} = \frac{(1, 0, 1)}{\sqrt{2}} . \quad (6.29)$$

Now the ratio of conclusive cases is:

$$1 - p_? = \frac{w}{2} = 1 - \frac{1}{\sqrt{2}} \approx 30\% . \quad (6.30)$$

Hence the optimum non-projective measurement has increased the ratio of correct decisions by cca. 5% with respect to projective measurements.

6.4 Applications of no-cloning and indistinguishability

The non-clonability 6.2.3 of an unknown q-state seems a shortcoming literally. Yet it becomes a benefit if we take a new point of view, namely, the security of information. It can be verified in simple terms that at suitable circumstances the hidden q-information may offer absolute security such that is guaranteed in q-theory whereas classically it can by no means be guaranteed in principle either. The q-theory of unforgeable banknote had been the earliest example. The secure q-key distribution is, for the time being, the only *q-protocol* used already in practice.

6.4.1 Q-banknote

Suppose that Alice, on the behalf of the bank of issue, makes a prepared two-state quantum system stick to each issued banknote, each in randomly different q-state which is of course logged by Alice and the bank. The bank keeps the logs private. Civilians like Bob have no access to the data. Suppose Bob attempts to duplicate or even multiply a banknote. Since the maximum fidelity of the copies is $2/3$, cf. Sect. 6.2.5, the bank will, on average, unveil $1/3$ of the forgeries while $2/3$ of the forged money could be perfectly consumed. Just for this reason, as a matter of course, Alice makes more than one independent qubit-marker stick to each banknote. N qubits reduces Bob's chances from $2/3$ to $(2/3)^N$.

For Alice and the bank, a simple q-protocol may be the following. Each banknote is identified by the usual public serial number and by a private q-serial-number. The latter is this. Each banknote contains a sequence of N independent two-state q-systems (qubits). Each qubit is either $|0\rangle$ or $|1\rangle$ at random⁴. Hence, each banknote carries a random sequence of alternative pure states, e.g.:

⁴ Despite identical notations, one should not confuse $|0\rangle, |1\rangle$ in this Section with the computational basis in the rest of the volume.

$$|1\rangle, |0\rangle, |0\rangle, \dots, |0\rangle, |1\rangle, |1\rangle \quad . \quad (6.31)$$

Alice and the bank log the otherwise public serial number together with the encoded N -digit binary “q-serial-number”. The latter as well as the logs are kept private. At the same time, both states $|0\rangle, |1\rangle$ can be published. If they were orthogonal to each other, Bob would make any number of perfect forges even without destroying the original banknotes. The point is that Alice and the bank should use non-orthogonal states⁵, e.g. (6.24):

$$|0\rangle = |\uparrow z\rangle, \quad |1\rangle = |\uparrow x\rangle, \quad (6.32)$$

then Bob can never profit from forgery. Even doing his best, Bob is not able to produce a sufficient number of high fidelity copies. The bulk of the copied and the original banknotes can not go through a test performed by Alice to verify the logged combination of the public and q-serial numbers. The systematic proof needs detailed statistical analysis which must extend to more sophisticated cloning strategies which Bob might invent. Let us add that prior to Alice’s test, it makes no sense to say that a single banknote is bad or good. This is hidden q-information as long as the test is not yet done.

The task of longtime coherent preservation of arbitrary qubits is not yet solved. Therefore q-banknotes have so far not been realized since the qubit states ought to be preserved on the banknote for an unlimited time. The no-cloning principle has nonetheless another application which is realizable easier, as we see in the next Section.

6.4.2 Q-key, q-cryptography

The fact of non-clonability can be utilized in secret, protected from the unauthorized, communication of information. In this case the coherent preservation of a qubit must be guaranteed for not longer than the duration of its transmission. This duration is particularly short when qubits are transmitted by light.

The simplest classical cryptography is based on a secret-key. This may be a sequence of binary digits which is known exclusively by the authorized parties (Alice and Bob) who will utilize the key to encode secret messages to each other. In course of distribution of the key between the authorized parties it may go through the hands of the unauthorized person (Eve) and she might unnoticedly read and copy the key. Even if Eve does not learn the exact key she can still break the secrecy of the communication between Alice and Bob. Q-cryptography excels classical cryptography in that the *unnoticed* eavesdropping is impossible during the distribution of the secret key. The collaborative procedure of Alice and Bob in order to establish the secret key is governed by the so-called secret-key q-protocol. Let us see the simplest two-state q-protocol⁶.

⁵ This idea of Wiesner from cca. 1969 remained unpublished for many years until [17].

⁶ This is Bennett’s version BB92 [18] of the original four-state BB84 protocol [19] by Bennett and Brassard. For a review on q-cryptography, cf. [11].

Alice sends a long sequence of random binary digits (raw-key) to Bob, encoded into N non-orthogonal qubits like in case of q-banknotes (6.31,6.32). Alice also allows Bob (and anybody else) to know the two non-orthogonal states $|0\rangle$ and $|1\rangle$, like e.g. in (6.24), that she has used to encode 0 and 1 respectively. Accordingly, Bob measures each qubit in turn. For lack of a better idea, he alternates randomly between the measurement of the two physical quantities:

$$\hat{\sigma}_z \equiv 2|0\rangle\langle 0| - \hat{I} \text{ or } \hat{\sigma}_x \equiv 2|1\rangle\langle 1| - \hat{I}. \quad (6.33)$$

The potential outcomes of the measurements are always ± 1 . The average frequency of the -1 is 25%, and in all these cases Bob obtains exact information about the qubit received from Alice. When, for instance, the measurement of $\hat{\sigma}_z$ yields -1 then the received state only can have been the $|\uparrow x\rangle$, never the $|\uparrow z\rangle$. Hence Bob can safely establish that Alice's raw-key contains the 0 in the given binary digit. Similarly, if the outcome of the measurement of $\hat{\sigma}_x$ was -1 then Bob establishes that the given digit of the key is 1. In such a way Bob restores 1/4 of the raw-key, this is the sift-key. Then Bob allows Alice (and anybody else) to know which digits of the raw-key have contributed to the sift-key. The above protocol supplies Alice and Bob with the same binary sequence of cca. $N/4$ digits, which they can use as the secret-key of their cryptographic communication.

In the language of communication theory, Alice sends the qubits, encoding the raw-key, through a *q-channel* to Bob. Further communications of the protocol, like that describing the code qubits to Bob and the locations of the bits of the sift-key to Alice, are sent through classical channels. Both the classical and q-channels are public, anybody has the access to the bits or q-bits travelling through them. Yet, the resulting sift-key is protected from unnoticed eavesdropping. Eve can unnoticedly eavesdrop the classical channel but she does not obtain any useful knowledge regarding the sift-key. She must concentrate on the q-channel and intercept some of the qubits. If Eve eavesdrops the q-channel then Alice and Bob will detect it because their sift-keys will not be identical. Surely, Eve can not discriminate two non-orthogonal states without altering the original state at the same time. Eve can certainly bar Alice and Bob from establishing the secret key if, e.g., she eavesdrops the q-channel too aggressively so that she may as well make the sift-keys of Alice and Bob completely uncorrelated. In this case there is nothing else left, Alice and Bob re-start the protocol.

In reality, the sift-keys of Alice and Bob may differ because of the eavesdropping by Eve and/or because of the own transmission noise of the q-channel (we ignore the noise during the classical communications of the protocol). Fortunately, Alice and Bob can afterwards eliminate the above differences. They can apply classical algorithms: *information reconciliation* and *privacy amplification*. The former makes, actually, error correction while the latter reduces the relevancy of the information procured by the unauthorized Eve. At the end of the day, Alice and Bob arrive at shorter keys than the original sift ones, which will, nonetheless, coincide with very high reliability while Eve's related information will not exceed a given small value. This happens when the transmission noise, caused by the q-channel itself and/or by

Eve’s attack on it, is not too high. If it is, then attempts of information reconciliation and privacy amplification will eat up the sift key completely.

A simple method of information reconciliation goes like this. Using classical public communication, Alice and Bob single out randomly the same two locations k, l from their sift-keys. They calculate the parity $x_k \oplus x_l$ of the corresponding bits and compare their results using the public classical channel. If the results are different, Alice and Bob eliminates both the k ’th and the l ’th bits from their sift-keys. If the results are equal then they agree to drop one of the bits, say the k ’th so that Eve can not acquire any further information on the key even if she listened to the public classical communication of Alice and Bob. The above procedure is iterated by Alice and Bob until they see that the frequency of disagreement between their shortened keys is already smaller than a given little threshold. If the length of the key is still big enough, then Alice and Bob moves to privacy amplification. At the price of further shortening the key, it reduces the information that may has earlier been acquired by Eve. Once again, in public classical communication, Alice and Bob single out randomly the same two locations k, l from their keys. They calculate the parity $x_k \oplus x_l$ of the corresponding bits and substitute their original two bits by one new bit which is the parity. By iterating this procedure the information of Eve, obtained earlier, will gradually decrease until Alice and Bob will with high probability have the identical secret keys in their hands, on which Eve (or anyone else) can not have more information than a given (small) limit. We give some hint of the information theoretical proof in Sect. 9.7.

Alice’s raw key:	1	0	0	1	0	1	...
Alice’s qubits to Bob:	$ \uparrow x\rangle$	$ \uparrow z\rangle$	$ \uparrow z\rangle$	$ \uparrow x\rangle$	$ \uparrow z\rangle$	$ \uparrow x\rangle$	...
Bob’s measured quantities:	$\hat{\sigma}_z$	$\hat{\sigma}_x$	$\hat{\sigma}_x$	$\hat{\sigma}_z$	$\hat{\sigma}_z$	$\hat{\sigma}_x$	...
Bob’s measured outcomes:	+1	-1	+1	-1	+1	+1	...
Bob to Alice:		*		*			...
distributed sift key:		0		1			...

Table 6.2. Two-state q-key protocol.

Problems, exercises

6.1. Universality of Hadamard and phase operations. Let us show that all 2×2 unitary matrices (upto a complex phase) can be constructed by the repeated application of the Hadamard- $\hat{H}$ and the phase-operation $\hat{T}(\varphi)$. Method: use Euler angles of rigid body rotation kinematics.

6.2. Statistical error of qubit determination. Suppose Alice hands over N identically prepared qubits $\hat{\rho}$ to Bob but she does not tell Bob what the state $\hat{\rho}$ is. Let Bob estimate the polarization vector s by the simple method 6.2.2. Write down

the estimation errors $\Delta s_x, \Delta s_y, \Delta s_z$ in function of s and the number of single state measurements. Method: determine the mean statistical fluctuation of the counts $N_{\uparrow x}, N_{\uparrow y}, N_{\uparrow z}$.

6.3. Fidelity of qubit determination. Suppose Alice sends Bob a random qubit $|n\rangle$. Bob knows this but he does not know the state itself. Bob measures a polarization $\hat{\sigma}_m$ chosen along a random direction m . Let us determine the best expected fidelity of Bob's state estimate.

6.4. Post-measurement depolarization. Alice prepares for Bob a state of polarization s which is unknown to Bob but he wishes to learn it. On the received state, Bob would measure the polarization along a random direction, for lack of a better idea. Let us prove that after the measurement the polarization of the state reduces to $s/3$.

6.5. Anti-linearity of polarization reflection. The polarization reflection $\mathcal{T}$ is equivalent with time inversion and corresponds to anti-linear transformation $\hat{T}$ on the space of state vectors. Let us verify that, upto a joint complex phase factor, the basis vectors transform like $\hat{T}|\uparrow\rangle = -|\downarrow\rangle$ and $\hat{T}|\downarrow\rangle = |\uparrow\rangle$.

6.6. General qubit effects. Suppose the q-effects:

$$\hat{H}_n = w_n \left(\hat{I} + \mathbf{a}_n \hat{\sigma} \right) ; \quad n = 1, 2, \dots$$

which form a positive decomposition for a qubit, cf. Sect. 4.4. List the necessary conditions on the w_n 's and $\mathbf{a}_n$'s.

7 Composite q-system, pure state

Even the simplest q-systems can not exhaustively be discussed without the concept of composite systems. As we shall see, the reason is q-correlations between otherwise independent q-systems. While classical correlations permit separate treatment of local systems, q-correlations will only permit this with particular limitations. The mathematics of composite q-systems will be introduced from the aspect of q-correlations (entanglements). The reader may learn three historical instances of q-correlation. Two genuine q-informatic applications based on q-correlations will close the Chapter.

7.1 Bipartite composite systems

The simplest composite system has 4-states, consists of two 2-state subsystems, i.e., of two qubits. Peculiar features of the composite q-systems can already be interpreted for such 2×2 -state systems as well. Also the major part of our theoretical knowledge on q-correlations has been obtained for 2×2 -state systems. Yet we start with a summary of the general bipartite composite systems, although we restrict the quantitative theory of entanglement for pure composite states:

$$|\psi_{AB}\rangle = \sum_{\lambda=1}^{d_A} \sum_{l=1}^{d_B} c_{\lambda l} |\lambda; A\rangle \otimes |l; B\rangle, \quad (7.1)$$

where $\{|\lambda; A\rangle\}$ and $\{|l; B\rangle\}$ are certain orthogonal bases in the respective subsystems A and B .

7.1.1 Schmidt decomposition

The state vector $|\psi_{AB}\rangle$ of a bipartite composite q-system consists of probability amplitudes $c_{\lambda l}$ of double indices (4.42) where the first (Greek) one labels the basis of subsystem A while the second (Latin) index labels the basis of subsystem B . Therefore the state vector can formally be considered a matrix and we can apply the theorem of Schmidt diagonalization to it: the above two orthogonal bases can be chosen in such a way that the matrix $c_{\lambda l}$ of amplitudes becomes diagonal, real, and non-negative. Taking normalization into account, we can write this theorem in the following form:

$$\{c_{\lambda l}\} = \sqrt{w_{\lambda}} \delta_{\lambda l}, \quad \lambda, l = 1, 2, \dots, \min\{d_A, d_B\}. \quad (7.2)$$

Observe that we have formally written the normalized non-negative amplitudes as the square-roots of a normalized probability distribution w_{λ} whose interpretation becomes clear below. The Schmidt decomposition generalizes the diagonal expansion of hermitian matrices, cf. Sect. 4.4, for arbitrary, even non-quadratic matrices. The rank of the given matrix coincides with the number of the nonzero diagonal terms, called also the Schmidt number. Hence, according to the Schmidt decomposition, the pure state of a bipartite composite q-system can always be written as the superposition of orthogonal tensor product state vectors:

$$|\psi_{AB}\rangle = \sum_{\lambda} \sqrt{w_{\lambda}} |\lambda; A\rangle \otimes |\lambda; B\rangle, \quad (7.3)$$

where the number of terms is the rank of the matrix of the composite state amplitudes. This is at most the dimension of the “smaller” subsystem and in this way it may be “much” less than the dimension of the composite system.

The reduced states of the two subsystems follow:

$$\begin{aligned} \hat{\rho}_A &\equiv \text{tr}_B (|\psi_{AB}\rangle \langle \psi_{AB}|) = \sum_{\lambda} w_{\lambda} |\lambda; A\rangle \langle \lambda; A|, \\ \hat{\rho}_B &\equiv \text{tr}_A (|\psi_{AB}\rangle \langle \psi_{AB}|) = \sum_l w_l |l; B\rangle \langle l; B|. \end{aligned} \quad (7.4)$$

Observe that the bases and the coefficients of the Schmidt decomposition can be obtained from the eigenvectors and eigenvalues of the reduced density matrices $\hat{\rho}_A, \hat{\rho}_B$. Their spectra $\{w_{\lambda}\}$ and $\{w_l\}$ are identical, in this sense the mixednesses of $\hat{\rho}_A$ and $\hat{\rho}_B$ are identical. (Do not forget, this is only true when the state of the composite system is pure.) The reduced states remain pure if and only if the state vector of the composite state is of the tensor product form $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, i.e., the Schmidt number is 1.

7.1.2 State purification

An arbitrary mixed state $\hat{\rho}$ of a q-system can be derived by reduction from a pure state of a suitably constructed fictitious larger composite q-system. Let us form the composite system consisting of the q-system in question and a fictitious *environmental* q-system E ¹. Consider the spectral expansion (4.11, 4.35) of the state $\hat{\rho}$ of the original q-system:

$$\hat{\rho} = \sum_{\lambda} w_{\lambda} |\lambda\rangle \langle \lambda|. \quad (7.5)$$

Introducing a basis for the environmental q-system E , we form the following pure state of the composite (“big”) system:

¹ Sometimes we call them the principal system and the ancilla, or the system and the meter in case of the indirect measurement 8.3.

$$|\psi_{big}\rangle = \sum_{\lambda} \sqrt{w_{\lambda}} |\lambda\rangle \otimes |\lambda; E\rangle , \quad (7.6)$$

where we assume that the dimension of E is not less than the dimension of the original system. By the way, the above form is a Schmidt decomposition (7.3). If we reduce it, we recover the density matrix (7.5) of the original q-system:

$$\hat{\rho} = \text{tr}_E (|\psi_{big}\rangle \langle \psi_{big}|) . \quad (7.7)$$

Consequently, any mixed state $\hat{\rho}$ can be considered as the reduction of the pure state $|\psi_{big}\rangle$ of an enlarged q-system. The procedure as well as the state $|\psi_{big}\rangle$ are called purification of $\hat{\rho}$. An alternative axiomatic construction of q-theory can therefore be started from pure states instead of mixed ones.

One must avoid a certain terminological inexactitude. Purification of q-states is not a q-operation to be performed on the given q-ensemble. Rather it is a mathematical construction, a nonlinear map.

7.1.3 Measure of entanglement

Separability of composite q-systems has been defined by (4.47): the density matrix must be a mixture of uncorrelated (tensor product) states. For pure composite states this is only possible if the state vector is of the tensor product form:

$$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle . \quad (7.8)$$

Tensor product state vectors yield obviously tensor product density matrices which are separable. But non-product state vectors are entangled. This follows from the fact that, in the separability condition (4.47), the decomposition of a pure composite state $\hat{\rho}_{AB}$ must reduce to a single tensor product density matrix:

$$\hat{\rho}_{AB} \equiv |\psi_{AB}\rangle \langle \psi_{AB}| = \hat{\rho}_A \otimes \hat{\rho}_B , \quad (7.9)$$

and this can only be satisfied if the state vector $|\psi_{AB}\rangle$ has the product form (7.8).

Accordingly, the pure composite q-state is either of the tensor product form and is thus completely uncorrelated (separable) or, alternatively, it is the superposition of tensor products and is thus correlated (entangled). Hence, for a pure composite q-state, correlations are always q-correlations, and their structural source is the superposition of uncorrelated (tensor product) state vectors. Therefore it makes sense to define the measure E of entanglement on the Schmidt decomposition (7.3). We postulate that the entanglement measure be equal to the Shannon entropy 9.1 of the distribution w_{λ} governing the structure of superposition:

$$E (|\psi_{AB}\rangle \langle \psi_{AB}|) = - \sum_{\lambda} w_{\lambda} \log w_{\lambda} , \quad 0 \leq E \leq \log \min\{d_A, d_B\} . \quad (7.10)$$

If $E = 0$ then all w_{λ} vanish except for a single one which is unity; the composite state vector becomes a tensor product which is, indeed, not entangled. In the contrary case, E takes its maximum value when the distribution w_{λ} is flat:

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{\min\{d_A, d_B\}}} \sum_{\lambda} |\lambda; A\rangle \otimes |\lambda; B\rangle, \quad (7.11)$$

and this state we consider maximally entangled. In this extremal case we shall see that the composite q-system is asymptotically equivalent with $E = \log \min\{d_A, d_B\}$ pairs of maximally entangled qubits (7.1.6). The more intrinsic meaning of the postulated relationship (7.10) between entropy and entanglement can be verified later, in the possession of the rudiments of information theory in Sects. 10.5, 10.6, 10.7. Still formally, we introduce the von Neumann entropy 10.1 of an arbitrary q-state:

$$S(\hat{\rho}) = -\text{tr}(\hat{\rho} \log \hat{\rho}) ; \quad 0 \leq S \leq \log d. \quad (7.12)$$

The von Neumann entropy is zero for pure states and may serve as a measure of the mixedness of the state. But we use it here to express the measure E of entanglement of a given composite pure state $|\psi_{AB}\rangle$. The von Neumann entropies of the reduced states $\hat{\rho}_A, \hat{\rho}_B$ (7.4) of the two subsystems coincide with each other and with the previously defined (7.10) measure of entanglement:

$$E(|\psi_{AB}\rangle \langle \psi_{AB}|) = S(\hat{\rho}_A) = S(\hat{\rho}_B). \quad (7.13)$$

In summary, the entanglement of a pure bipartite composite q-state is measured by the (coinciding) von Neumann entropies of the reduced q-states of the two subsystems. For mixed bipartite composite q-states the criterion of separability as well as the measure of entanglement are much more difficult and only partially known. There exist separable but classically correlated states; this can only happen to mixed composite states, never to pure ones.

7.1.4 Entanglement and local operations

How can we create entanglement? How can we bring a product state vector into a superposition of such products? If we choose the mere dynamical way (4.48), then we need a nonzero interaction Hamiltonian (4.49). If there were no interaction then the product initial state vector would remain of the product form. The entanglement E (7.13) of a general initial state vector would not change either! Why, in the lack of their interaction the subsystems A and B would evolve independently of each other, according to their own respective Hamilton matrices $\hat{H}_A, \hat{H}_B$. Their unitary evolution leaves their entropies unchanged.

By means of suitable dynamical interaction, the entanglement of A and B can of course be created, modified, increased, decreased or just eliminated. However, in particular q-informatic situations we may not assume dynamical interaction between the two q-subsystems because, for instance, they are far distant from each other. Commonly, we talk about the q-subsystems A and B as the system of Alice and Bob, respectively, who can only influence their own *local* systems by the given choice of the local Hamilton matrices $\hat{H}_A, \hat{H}_B$. Accordingly, we say that the Hamiltonian is *local* if $\hat{H}_{ABint} = 0$ and is nonlocal otherwise. Usually Alice and Bob are assumed

to use local dynamics so they can not influence the entanglement between their systems in the dynamical way.

Another elemental mean of influencing a given q-state is when we measure a certain q-physical quantity. In case of bipartite composite systems, we call a physical quantity *local* if its measurement can be performed via suitable local measurements by Alice and Bob, respectively. Quantities $\hat{A} \otimes \hat{I}$ or $\hat{I} \otimes \hat{B}$ are typical local ones. The following quantity is typically nonlocal:

$$\hat{A} \otimes \hat{B} + \hat{A}' \otimes \hat{B}' . \quad (7.14)$$

By means of local measurements, Alice and Bob are still able to determine its expectation value but the post-measurement state will be different from what it would be after the standard q-measurement of the nonlocal quantity itself. Its q-measurement is not at all viable by local q-measurements.

It will be shown in Sect. 8.5 of next Chapter that local operations (LO) can never create or increase entanglement. Prior to the proof, we need to determine the notion of the most general operations 8.3. If, however, there is already some entanglement then it can by local methods concentrated onto a given composite subsystem. The protocol, quite unusual in common q-theory, is called distillation, 10.6 .

7.1.5 Entanglement of two-qubit pure states

If both A and B are 2-state q-systems (qubits) in pure states and are uncorrelated then their composite (two-qubit) system is a pure state of the product form:

$$|\psi_{AB}\rangle = |\mathbf{n}_A\rangle \otimes |\mathbf{n}_B\rangle . \quad (7.15)$$

However, the general pure two-qubit state is entangled since it can be written as the superposition of 4 mutually orthogonal uncorrelated states like this, e.g., in Pauli representation (5.11):

$$c_{\uparrow\uparrow} |\uparrow\rangle \otimes |\uparrow\rangle + c_{\downarrow\uparrow} |\downarrow\rangle \otimes |\uparrow\rangle + c_{\uparrow\downarrow} |\uparrow\rangle \otimes |\downarrow\rangle + c_{\downarrow\downarrow} |\downarrow\rangle \otimes |\downarrow\rangle . \quad (7.16)$$

The same thing in computational (5.33) representation reads:

$$\sum_{x_1=0,1} \sum_{x_2=0,1} c_{x_1,x_2} |x_2\rangle \otimes |x_1\rangle \equiv \sum_{x=0}^3 c_x |x\rangle , \quad x \equiv x_2x_1 . \quad (7.17)$$

As we see, introducing the two-digit binary number x , i.e.: $2x_2 + x_1$, assigns integer labels 0, 1, 2, 3 to the 4 basis vectors and this allows a compound notation of the general pure state in 2×2 dimensions.

So far we have considered bases of the two-qubit system constructed as the tensor product of the respective bases of the two single-qubit systems. Non-product bases can equally be useful. In the theory of composite spins, the following four orthogonal basis states are common:

$$\text{singlet: } \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}, \quad (7.18)$$

$$\text{triplet: } |\uparrow\uparrow\rangle, \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}, |\downarrow\downarrow\rangle. \quad (7.19)$$

We used compact notations like $|\uparrow\downarrow\rangle$ for $|\uparrow\rangle \otimes |\downarrow\rangle$. The singlet and the middle one of the triplet states are maximally entangled (7.11) while the other two triplet states are uncorrelated. The singlet state is rotationally invariant. This can be made explicit if we write down its density matrix²:

$$\hat{\rho}(\text{singlet}) = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \frac{\langle\uparrow\downarrow| - \langle\downarrow\uparrow|}{\sqrt{2}} = \frac{\hat{I}_{AB} - \hat{\sigma}_A \otimes \hat{\sigma}_B}{4}. \quad (7.20)$$

Accordingly, the states $|\uparrow\rangle, |\downarrow\rangle$ in the expression of the singlet state vector (7.18) can be equivalently replaced by $|\uparrow n\rangle, |\downarrow n\rangle$ no matter what n is.

In the Bell basis, all 4 orthogonal vectors are maximally entangled:

$$\begin{aligned} |\Phi^\pm\rangle &= \frac{|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle}{\sqrt{2}}, \\ |\Psi^\pm\rangle &= \frac{|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle}{\sqrt{2}}. \end{aligned} \quad (7.21)$$

The polarizations are completely correlated in the first two basis states and completely anti-correlated in the second two. The Bell basis vector $|\Psi^-\rangle$ coincides with the singlet state (7.20). Another Bell basis vector takes a compound form in computational basis (7.17):

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} \sum_{x=0,1} |x\rangle \otimes |x\rangle. \quad (7.22)$$

Each of the four Bell states has entanglement $E = 1$ by definition (7.13). We use two-qubit systems in Bell states as standard components to build or to decompose large entangled bipartite systems, as we shall see in Sects. 7.1.6, 10.6, 10.7.

7.1.6 Interchangeability of maximal entanglements

It is desirable to show that a general bipartite system in pure state of entanglement E is equivalent with E independent copies of two-qubit systems each in Bell states of entanglement 1. We present the proof for the special case of maximum entanglement. For the general pure entangled state, the proof 10.6, 10.7 will need further q-informatic notions.

Suppose a maximally entangled bipartite pure state whose Schmidt decomposition (7.11) reads:

² The factors of two $\hat{\sigma}$'s, when appear as $\hat{\sigma} \otimes \hat{\sigma}$ or $\hat{\sigma} \hat{\rho} \hat{\sigma}$, etc., should be understood as spatial scalar products. E.g.: $\hat{\sigma} \otimes \hat{\sigma} = \hat{\sigma}_x \otimes \hat{\sigma}_x + \hat{\sigma}_y \otimes \hat{\sigma}_y + \hat{\sigma}_z \otimes \hat{\sigma}_z$.

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{\lambda=1}^d |\lambda; A\rangle \otimes |\lambda; B\rangle, \quad (7.23)$$

where, for simplicity, let $d_A = d_B = d$ and let $E = \log(d) \equiv k$ be integer. Alice's d -state system can be built up from k qubits, i.e., in the form of a k -fold composite system of 2-state subsystems. A straightforward correspondence between the two bases can be obtained by the k -digit binary decomposition of the label λ of the original basis vectors:

$$|\lambda; A\rangle = |x_k; A\rangle \otimes \dots \otimes |x_2; A\rangle \otimes |x_1; A\rangle; \quad \lambda = x_k \dots x_2 x_1 \equiv x. \quad (7.24)$$

Similarly, we introduce another basis in the state space of Bob. In the new bases, the state (7.23) can be written into this form:

$$\begin{aligned} & \frac{1}{\sqrt{d}} \sum_{x=1}^d [|x_k; A\rangle \otimes \dots \otimes |x_2; A\rangle \otimes |x_1; A\rangle] \otimes [|x_k; B\rangle \otimes \dots \otimes |x_2; B\rangle \otimes |x_1; B\rangle] = \\ & = \frac{1}{\sqrt{d}} \left(\sum_{x=0,1} |x; A\rangle \otimes |x; B\rangle \right)^{\otimes k} = |\Phi^+\rangle^{\otimes k}. \end{aligned} \quad (7.25)$$

In such a way, we have proved that the pure state (7.23) in $d \times d$ -dimension, when it possesses maximum entanglement $E = \log(d)$, is in a suitable basis equivalent with E copies of 2×2 -dimensional (Bell) states of maximum entanglements $E = 1$.

$$|\psi_{AB}\rangle = |\Phi^+\rangle^{\otimes E}. \quad (7.26)$$

7.2 Q-correlations history

The appreciation of unusual correlations in composite q-systems as well as their theoretical characterization took a lengthy time historically. In each of the forthcoming three Sections, we recall a decisive theoretical discovery.

7.2.1 EPR, Einstein-nonlocality 1935

Suppose that Alice and Bob possess one qubit each. As a result of their earlier interaction, these qubits are being in a maximally entangled state, e.g., in the singlet state of the Bell basis (7.21):

$$|\psi_{AB}\rangle \equiv |\Psi^-\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}, \quad (7.27)$$

where, as usual, the first polarization refers to Alice's and the second refers to Bob's qubit, respectively. Alice and Bob, together with their qubit, live far from each other. Imagine that Alice measures $\hat{\sigma}_z$ on her qubit. This is the random result of the measurement:

$$|\psi_{AB}\rangle \rightarrow \begin{cases} |\uparrow z; \downarrow z\rangle; & \hat{\sigma}_{Az} \rightarrow 1; & p_+ = 1/2 \\ |\downarrow z; \uparrow z\rangle; & \hat{\sigma}_{Az} \rightarrow -1; & p_- = 1/2. \end{cases} \quad (7.28)$$

It is clear that the state of Bob's qubit becomes $|\uparrow\rangle$ or $|\downarrow\rangle$ with 50–50% probabilities, perfectly anti-correlated with Alice's post-measurement state. Alternatively, if Alice measures $\hat{\sigma}_x$ instead of $\hat{\sigma}_z$ then Bob's post-measurement state becomes the mixture of $|\uparrow x\rangle$ and $|\downarrow x\rangle$ instead of $|\uparrow\rangle$ and $|\downarrow\rangle$. At a superficial glance, the post-measurement state of Bob's qubit is completely different if Alice measures $\hat{\sigma}_x$ instead of $\hat{\sigma}_z$. In such a way Alice could transmit information to Bob via action-at-a-distance without any physical interaction, merely exploiting the entanglement of their qubits. We know that this can not be so: Bob has access but his own qubit and its reduced density matrix $\hat{\rho}_B$ is totally independent from whatever measurement made by Alice. In our case $\hat{\rho}_B$ is the maximally mixed state. Its decomposition is never unique. It is really ambiguous, exactly like in (6.15) and (6.16), depending on Alice remote choice $\hat{\sigma}_z$ or $\hat{\sigma}_x$ to measure. Yet Bob can never detect the difference.

There is, however, a crucial lesson. Einstein gave a particular strong formulation to the notion of locality. According to the principle of Einstein locality: In a theory *completely describing* the physical reality, the action on system A can *not* influence the *description* of system B if A and B are spatially separated. As we see from the above EPR paradox³, the q-theory violates Einstein locality. EPR chose the following resolution: the q-theory does not give complete description of the physical reality. Let us discuss the underlying motivations of EPR.

If Alice and Bob shared a correlated *classical* composite system, the local selective measurement (2.16) by Alice would, similarly to the above q-case, influence the state (e.g.: phase-space distribution) of Bob's remote system. This is, however, merely the consequence of the incompleteness of the description: certain parameters of the composite system have remained hidden and treated statistically. In classical physics it is simple to see that the complete description is possible when the system is in pure state, i.e. all canonical variables are exactly specified. Then the description of the classical system becomes deterministic and can not violate Einstein locality anymore. In the q-theory, however, even pure states can violate Einstein locality. This is why EPR consider q-theory incomplete. Whether q-theory, too, can be made complete if we discover its *hidden parameters* and we specify their values? Whether this complete theory will satisfy Einstein locality? The negative answer comes soon in Sect. 7.2.3.

7.2.2 A non-existing linear operation 1955

Consider the maximally entangled singlet state $|\Psi^-\rangle$ of two qubits, whose density matrix reads:

$$\frac{\hat{I} \otimes \hat{I} - \hat{\sigma} \otimes \hat{\sigma}}{4}. \quad (7.29)$$

Like in case of EPR, the first qubit belongs to Alice, the second to Bob, who are far from each other. Define the polarization reflection (6.9) of a generic qubit, which

³ Cf. [20] by Einstein, Podolski and Rosen.

could be a hypothetic linear q-operation $\mathcal{T}$:

$$\mathcal{T} \frac{\hat{I} + s\hat{\sigma}}{2} = \frac{\hat{I} - s\hat{\sigma}}{2}. \quad (7.30)$$

Assume that Alice is able to perform such an operation and she does it as well on her qubit entangled in the singlet state. It is straightforward to see the result. The reflection changes the sign of polarization of Alice's own Pauli matrices while it preserves Bob's ones:

$$(\mathcal{T} \otimes \mathcal{I}) \frac{\hat{I} \otimes \hat{I} - \hat{\sigma} \otimes \hat{\sigma}}{4} = \frac{\hat{I} \otimes \hat{I} + \hat{\sigma} \otimes \hat{\sigma}}{4}. \quad (7.31)$$

Surprisingly, this 4×4 Hermitian matrix has become indefinite! For instance, its quadratic form with the singlet state is negative:

$$\langle \Psi^- | \frac{\hat{I} \otimes \hat{I} + \hat{\sigma} \otimes \hat{\sigma}}{4} | \Psi^- \rangle = \text{tr} \left(\frac{\hat{I} \otimes \hat{I} - \hat{\sigma} \otimes \hat{\sigma}}{4} \frac{\hat{I} \otimes \hat{I} + \hat{\sigma} \otimes \hat{\sigma}}{4} \right) = -\frac{1}{2}. \quad (7.32)$$

Hence the local reflection (7.30) of the singlet state is no more a density matrix⁴. It can be shown that this anomaly arises not just for the singlet state but for all entangled two-qubit states. Therefore the local (also called partial) reflection makes an elegant mathematical test of entanglement, this is the Peres-Horodecki test⁵.

Let $\mathcal{T}$ be the *reflection* of a single qubit:

$$\mathcal{T} \frac{\hat{I} + s\hat{\sigma}}{2} = \frac{\hat{I} - s\hat{\sigma}}{2}.$$

Let $\hat{\rho}$ be a two-qubit state. If its *partial reflection* is non-negative:

$$(\mathcal{T} \otimes \mathcal{I})\hat{\rho} \geq 0,$$

then the state $\hat{\rho}$ is *separable*. If its partial reflection is indefinite:

$$(\mathcal{T} \otimes \mathcal{I})\hat{\rho} \not\geq 0,$$

then the state $\hat{\rho}$ is *entangled*.

Table 7.1. The Peres-Horodecki test of separability.

The polarization reflection $\mathcal{T}$ can not be performed in reality. To interpret the reflection (7.30) of a given local qubit, say of Alice, one should assume that it is not entangled with whatever other qubits in the Universe. Such an assumption, however,

⁴ This happens just because $\mathcal{T}$ belongs to non-completely positive maps, cf. Chap. 8, discovered by Stinespring [21].

⁵ Cf. [22] by Peres and [23] by the Horodecki's.

can not be justified. On the contrary, there are classical and q-correlations between local and various remote q-systems of the Universe. But while the existence of external classical correlations does not influence the structural properties (the theory) of local systems, the existence of external q-correlations (entanglements) imposes serious conditions for (the theory of) local q-systems.

7.2.3 Bell nonlocality 1964

Can we make the q-theory complete if we introduce certain *hidden parameters* so that Einstein nonlocality 7.2.1 disappear from it? John Bell pointed out that, like the classical theory, also the q-theory can be made complete and deterministic via introducing suitable hidden parameters. (From this aspect it is irrelevant that such a theory would be more complicated and less elegant than the standard q-theory.) Bell has, however, also proved that the obtained complete description can not satisfy the locality principle of Einstein. In Bell's formulation: no local hidden parameter model can replace the q-theory⁶.

Like in case of EPR nonlocality, we suppose Alice and Bob share an ensemble of singlet states $|\Psi_{AB}\rangle = |\Psi^-\rangle$ (7.27). Let us consider two couples of local physical quantities on the side of Bob and Alice, respectively:

$$\hat{A} = a\hat{\sigma}_A, \quad \hat{A}' = a'\hat{\sigma}_A, \quad \hat{B} = b\hat{\sigma}_B, \quad \hat{B}' = b'\hat{\sigma}_B, \quad (7.33)$$

where a, a', b, b' are polarization unit vectors. Alice and Bob measure 4 different nonlocal combinations:

$$\hat{A} \otimes \hat{B}, \quad \hat{A}' \otimes \hat{B}, \quad \hat{A} \otimes \hat{B}', \quad \hat{A}' \otimes \hat{B}'. \quad (7.34)$$

The measurement of the expectation values of such nonlocal quantities is possible with local measurements. For instance, the measurement of $\langle \hat{A} \otimes \hat{B} \rangle$ is equivalent with the simultaneous local measurements of $\hat{A}$ and $\hat{B}$ by Alice and Bob, respectively. But the 4 combinations are not compatible therefore they can not be simultaneously measured. Alice and Bob will take 4 independent measurement statistics on 4 randomly chosen sub-ensembles of the q-ensemble of the singlet states. Alice and Bob measure in coincidence. The outcome of each local measurement is ± 1 . Having the four statistics been taken, Alice and Bob can calculate the expectation value of the following nonlocal quantity, cf. also (7.14):

$$\hat{C} = \hat{A} \otimes \hat{B} + \hat{A}' \otimes \hat{B} + \hat{A} \otimes \hat{B}' - \hat{A}' \otimes \hat{B}'. \quad (7.35)$$

Let us now suppose that there exist hidden parameters which completely determine the outcomes of all measurements. We also adopt a further, delicate and much discussed, assumption: let the hidden parameters determine the outcomes of the chosen measurements as well as what the outcomes *would have been*, if we had chosen to do one of the other q-measurements incompatible with the one that

⁶ Cf. [24] by John Bell.

was actually done. Let us then index the configurations of the hidden parameters by r . Each q-system of the singlet-state q-ensemble has its r . The hidden parameters assign definite outcome ± 1 to each local polarization measurement (7.33):

$$\hat{A} = A_r = \pm 1, \quad \hat{A}' = A'_r = \pm 1, \quad \hat{B} = B_r = \pm 1, \quad \hat{B}' = B'_r = \pm 1. \quad (7.36)$$

Note that such a hidden parameter description of the measured reality would trivially satisfy the Einstein-locality: any manipulation by Alice on her qubit of hidden parameter r would not influence the description, i.e. the values B_r, B'_r , of Bob's qubit. Hence we call the r 's local hidden parameters.

The relative frequencies of the ± 1 outcomes as well as their correlations should, of course, satisfy the predictions of q-theory. For instance:

$$\langle \hat{A} \otimes \hat{B} \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{r \in \Omega} A_r B_r; \quad N = |\Omega|, \quad (7.37)$$

and similarly for the other 3 combinations (7.34) as well. This implies the following identity:

$$\langle \hat{C} \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_r A_r B_r + A'_r B_r + A_r B'_r - A'_r B'_r. \quad (7.38)$$

Consider the important relationship:

$$A_r B_r + A'_r B_r + A_r B'_r - A'_r B'_r = \pm 2, \quad (7.39)$$

which is a consequence of (7.36). Therefore the existence of the local hidden parameter expression (7.38) imposes the following restriction on the correlation quantity $\hat{C}$ (7.35):

$$-2 \leq \langle \hat{C} \rangle \leq 2. \quad (7.40)$$

This is the Bell inequality⁷. Note that this inequality itself is purely classical, its derivation has nothing to do with q-theory. We can ask whether it is satisfied by the value $\langle \hat{C} \rangle$ predicted by the q-theory.

We substitute the combination (7.35) of the polarizations (7.33) in place of $\langle \hat{C} \rangle$. In singlet state we obtain:

$$-2 \leq -\mathbf{a}\mathbf{b} - \mathbf{a}'\mathbf{b} - \mathbf{a}\mathbf{b}' + \mathbf{a}'\mathbf{b}' \leq 2. \quad (7.41)$$

This geometrical condition must be satisfied for all possible choices of the 4 directions. But it will not be. Suppose, for instance, the following 4 co-planar directions:

$$\mathbf{a} = \rightarrow, \quad \mathbf{a}' = \uparrow, \quad \mathbf{b} = \swarrow, \quad \mathbf{b}' = \nwarrow. \quad (7.42)$$

Insert the corresponding 4 scalar products into the Bell inequality (7.41):

⁷ It is the version [25] by Clauser, Horne, Shimony and Holt.

$$\begin{aligned}
& -ab - a'b - ab' + a'b' \\
& = -\cos(3\pi/4) - \cos(3\pi/4) - \cos(3\pi/4) + \cos(\pi/4) = 2\sqrt{2}. \quad (7.43)
\end{aligned}$$

This value violates the Bell inequality. It means that in singlet state the values (7.36) of the 4 physical quantities can not be chosen as a function of the hidden parameters r in such a way that the statistical predictions of the q-theory be satisfied. This is called Bell nonlocality. Separable states are always Bell local. But we do not know whether or not all entangled states show Bell nonlocality⁸.

7.3 Applications of Q-correlations

We learn two direct applications of q-correlations: superdense coding, and teleportation. Options of information manipulations become available which could not be possible in classical physics.

7.3.1 Superdense coding

In the basic setup of communication, Alice sends information through a channel to Bob via classical binary data. What can they profit if Alice sends qubits, instead of classical bits, through a q-channel? When Alice and Bob agree that Alice encodes the binary information into orthogonal qubits, e.g. into $|\uparrow\rangle, |\downarrow\rangle$, then Bob measures $\hat{\sigma}_z$ and their q-communication becomes equivalent with the classical communication. But q-communication can be more powerful, at least in solving certain particular tasks.

Suppose then a q-channel from Alice to Bob which, for the time being, they can use without limitation as well as they can communicate through a classical channel without limitation. They know that one year later Alice obtains classical information from somewhere and she must forward it to Bob. They also know that one year later the classical channel is already not available and they can only send qubits through the q-channel. Intuition says that they have to consume one qubit to encode one classical bit. We shall see, however, that Alice and Bob can, well in time, agree upon the protocol of superdense coding which encodes two bits into one qubit⁹.

Alice prepares pairs of qubits entangled in singlet states and sends one qubit of each pair to Bob through the q-channel. Both Alice and Bob preserve their own qubits. Of course, Alice neither could nor wished to send any information to Bob via the entangled qubits. One year later, Alice learns the first two bits that she will forward to Bob via one qubit. She encodes the two bits into the 4 different unitary 1-qubit operations $\hat{I}, \hat{X}, \hat{Y}, \hat{Z}$. She performs the actual operation on the first of her qubits kept with her. Then the whole composite singlet state becomes one of the 4 orthogonal Bell states (7.21), according to the scheme:

⁸ Cf., e.g., [26] by Popescu.

⁹ Cf. [27] by Bennett and Wiesner.

$$\begin{aligned}
(\hat{I} \otimes \hat{I}) |\Psi^-\rangle &= |\Psi^-\rangle \\
(\hat{X} \otimes \hat{I}) |\Psi^-\rangle &= |\Phi^-\rangle \\
(\hat{Y} \otimes \hat{I}) |\Psi^-\rangle &= |\Phi^+\rangle \\
(\hat{Z} \otimes \hat{I}) |\Psi^-\rangle &= |\Psi^+\rangle .
\end{aligned} \tag{7.44}$$

After the unitary operation, Alice sends her qubit through the q-channel to Bob who has thus one of the 4 orthogonal Bell states in his hands, encoding the two bits of information according to the protocol. The information can be decoded if Bob performs a projective measurement of the Bell basis. In such a way, Bob acquires the original two qubits. If Alice and Bob repeatedly perform the above protocol, Alice will be able to send $2n$ bit information via n qubits while they have to consume n shared singlet states.

Superdense coding does not mean that we squeeze 2 bits of information into 1 qubit. In fact it means that we can encode 2 bits of information into 2 qubits in such a way that physically we do not touch but one of the qubits. Therefore the other qubit can be transmitted to the other party in advance, say, when the classical information is not yet even known. In this sense can the coding into the retained single qubit be considered denser than what would be available classically. For superdense coding, entanglement has obviously been instrumental.

7.3.2 Teleportation

Transmission of an unknown qubit requires a q-channel. Yet qubit transmission is possible through classical channel if the parties share a reservoir of previously entangled qubits. Suppose a q-channel from Alice to Bob which, for the time being, they can use without limitation as well as they can communicate through a classical channel without limitation. They know that one year later Alice receives unknown qubits from somewhere and she must forward them to Bob. They also know that one year later the q-channel is already not available and they can only send bits through the classical channel. Intuition says that classical communication is not likely to transmit q-information. We shall see, however, that Alice and Bob can, well in time, agree upon the protocol of teleportation¹⁰ which transmits the unknown qubit from Alice to Bob. More precisely, the original qubit remains with Alice all the time while its state becomes perfectly inherited by a raw qubit of Bob. The protocol destroys the state of the original qubit, otherwise teleportation could make perfect cloning, too, which would be a contradiction, cf. Sect. 6.2.3.

Alice prepares pairs of qubits entangled in singlet states and sends one qubit of each pair to Bob through the q-channel. Both Alice and Bob preserves their own qubits. Of course, Alice neither could nor wished to send any information to Bob via the entangled qubits. One year later, Alice receives the first qubit

$$|\psi\rangle = a|\uparrow\rangle + b|\downarrow\rangle \tag{7.45}$$

¹⁰ Cf. [28] by Bennett, Brassard, Crépeau, Jozsa, Peres and Wootters.

that she will forward to Bob via two classical bits. Consider the 3-qubit composite state consisting of the received qubit and a singlet state shared by Alice and Bob:

$$|\psi\rangle \otimes |\Psi^-\rangle = [a|\uparrow\rangle + b|\downarrow\rangle] \otimes \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}. \quad (7.46)$$

This can be re-written if we introduce the Bell basis (7.21) for the two qubits on Alice's side:

$$\begin{aligned} & -\frac{1}{2} |\Psi^-\rangle \otimes [a|\uparrow\rangle + b|\downarrow\rangle] + \frac{1}{2} |\Phi^-\rangle \otimes [a|\downarrow\rangle + b|\uparrow\rangle] + \\ & + \frac{1}{2} |\Phi^+\rangle \otimes [a|\downarrow\rangle - b|\uparrow\rangle] - \frac{1}{2} |\Psi^+\rangle \otimes [a|\uparrow\rangle - b|\downarrow\rangle]. \end{aligned} \quad (7.47)$$

We see that the singlet state $|\Psi^-\rangle$ on Alice's side is multiplied by the state $a|\uparrow\rangle + b|\downarrow\rangle = |\psi\rangle$ on Bob's side, which is just the state (7.45) received originally by Alice to teleport. In fact all four Bell states on Alice's side are correlated with $|\psi\rangle$ upto the trivial unitary operations $\hat{I}, \hat{X}, \hat{Y}, \hat{Z}$ in turn. Then Alice performs the projective measurement of the Bell basis. There are four possible outcomes $|\Psi^\pm\rangle, |\Phi^\pm\rangle$; Alice transmits 2 bits through the classical channel to inform Bob of the outcome. So Bob learns which transform he must apply to his qubit:

$$\begin{aligned} |\Psi^-\rangle &: \hat{I} [a|\downarrow\rangle + b|\uparrow\rangle] \\ |\Phi^-\rangle &: \hat{X} [a|\downarrow\rangle + b|\uparrow\rangle] \\ |\Phi^+\rangle &: \hat{Y} [a|\downarrow\rangle - b|\uparrow\rangle] \\ |\Psi^+\rangle &: \hat{Z} [a|\uparrow\rangle - b|\downarrow\rangle]. \end{aligned} \quad (7.48)$$

The resulting state of Bob's qubit is precisely $|\psi\rangle$ (7.45) which was the state to teleport. If Alice and Bob repeatedly perform the above protocol, Alice will be able to send n qubits via $2n$ bits while they have to consume n shared singlet states.

The four possible measurement outcomes, obtained by Alice, are always totally random as it follows from the coefficients of the orthogonal decomposition (7.47). Alice does not acquire any information on the qubit that she teleports. Yet she entangles this qubit with the qubit of the shared singlet state, and this entanglement makes then possible that, via 2 bits of classical information, the unknown qubit reappears on Bob's side while it becomes completely smashed on Alice's side. The hidden information on the teleported qubit has been travelled through the chain of entanglements of the 3 qubits involved in the protocol.

The above teleportation protocol applies invariably to mixed states, too. More than that, it teleports all external q-correlations of the given qubit. Suppose that the state to teleport has been entangled with the state of a certain environmental q-system:

$$a|1; E\rangle \otimes |\uparrow\rangle + b|2; E\rangle \otimes |\downarrow\rangle. \quad (7.49)$$

Once the teleportation protocol from Alice to Bob has been done, the environmental state becomes entangled with Bob's state exactly the same way. The proof is algebraically identical with the previous proof of the teleportation of the unentangled state (7.45), if we make the formal substitutions:

$$a \rightarrow a |1; E\rangle \otimes, \quad b \rightarrow b |2; E\rangle \otimes. \quad (7.50)$$

Hence Bob's qubit becomes in all respects perfectly identical with the qubit that Alice received to teleport, including all its external q-correlations. The qubit received by Alice is left behind in the completely mixed state.

Problems, exercises

7.1. Schmidt orthogonalization theorem. Let us prove the Schmidt decomposition theorem (7.2) for a complex rectangular matrix $\hat{c}$, starting from the well known spectral expansion theorem of Hermitian matrices. Method: apply the latter to both matrices $\hat{c}^\dagger \hat{c}$ and $\hat{c} \hat{c}^\dagger$.

7.2. Swap operation. Consider two identical q-systems in arbitrary uncorrelated pure state $|\psi\rangle \otimes |\phi\rangle$. The swap matrix is defined by $\hat{S}(|\psi\rangle \otimes |\phi\rangle) = |\phi\rangle \otimes |\psi\rangle$. $\hat{S}$ is unitary and Hermitian. Let us prove that for two qubits

$$\hat{S} = \frac{\hat{I} \otimes \hat{I} + \hat{\sigma} \otimes \hat{\sigma}}{2}.$$

Method: introduce the basis $|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle$.

7.3. Singlet density matrix. Let us write down the density matrix of the two-qubit singlet state in Pauli representation. Method: exploit rotational invariance or express $\hat{\rho}(\text{singlet})$ through the swap $\hat{S}$.

7.4. Local measurement of expectation values. Let us show that the expectation value of the nonlocal physical quantity $\hat{A} \otimes \hat{B} + \hat{A}' \otimes \hat{B}'$ can be measured locally as well. Method: verify that the expectation values of $\hat{A} \otimes \hat{B}$ and of $\hat{A}' \otimes \hat{B}'$ are locally measurable.

7.5. Local measurement of certain nonlocal quantities. A tensor product q-physical quantity $\hat{A} \otimes \hat{B}$ may not be measured locally although its expectation value is always measurable locally. Let us explain why the nonlocal measurement of $\hat{\sigma}_z \otimes \hat{\sigma}_z$ is not equivalent with the local simultaneous measurement of $\hat{\sigma}_z \otimes \hat{I}$ and $\hat{I} \otimes \hat{\sigma}_z$. Let us prove that for the local measurability of $\hat{A} \otimes \hat{B}$ it is sufficient if the spectrum of $\hat{A} \otimes \hat{B}$ is non-degenerate.

7.6. Nonlocal hidden parameters. Let us show that the q-theoretic correlations can after all be reproduced by nonlocal hidden parameters. Method: we complete the hidden parameter r by a switch $\nu = 1, 2, 3, 4$ marking which of the four measuring setups is active. Then we can already assign the values $A_{r\nu}, B_{r\nu}, A'_{r\nu}, B'_{r\nu}$ suitably to each pair of qubits in the ensemble.

7.7. Does teleportation clone the qubit? Teleportation would be a perfect cloner had the procedure not destroyed the state of the original qubit. Let us show that the original qubit is always left behind in the totally mixed state $\hat{\rho} = \hat{I}/2$.

8 All q-operations

So far we have learned certain particular q-operations like the unitary transformations, the q-measurements, and the reduction of state. The set of generic q-operations is much larger. We learn the elegant mathematical classification. We discuss its various physical interpretations typically involving the temporary enlargement of the system by a certain environmental system, subsequent unitary and measurement operations on the obtained composite system, and an ultimate reduction to the original system.

8.1 Completely positive maps

We call a linear map $\mathcal{M}$ positive if it brings the density matrix $\hat{\rho}$ of a given system into a non-negative matrix whose trace does not exceed 1. In q-theory, we have to define a subset of positive maps which is the completely positive maps¹. A positive map $\mathcal{M}$ is completely positive if its trivial extension $\mathcal{M} \otimes \mathcal{I}$ for an arbitrary composite system remains positive map. The completely positive maps are always of the

The map $\mathcal{M} : \hat{\rho} \rightarrow \mathcal{M}\hat{\rho}$ is:	
positive map if	also completely positive map if
$\mathcal{H}$ is Hilbert space,	$\mathcal{H}_E$ is arbitrary “environmental” Hilbert space,
$\hat{\rho}$ is density matrix on $\mathcal{H}$,	$\hat{\rho}_{big}$ is density matrix on $\mathcal{H} \otimes \mathcal{H}_E$,
$\mathcal{M}\hat{\rho} \geq 0$ and $\text{tr}(\mathcal{M}\hat{\rho}) \leq 1$.	$(\mathcal{M} \otimes \mathcal{I})\hat{\rho}_{big} \geq 0$.

Table 8.1. Positive and completely positive maps.

following Kraus form (we omit the proof):

$$\mathcal{M}\hat{\rho} = \sum_n \hat{M}_n \hat{\rho} \hat{M}_n^\dagger, \quad \sum_n \hat{M}_n^\dagger \hat{M}_n \leq \hat{I}, \quad (8.1)$$

and the Kraus forms are always completely positive. However, given a completely positive map $\mathcal{M}$, its Kraus form is never unique: the Kraus matrices $\hat{M}_n$ — forming the “sandwich” — can be chosen in many different ways.

¹ Completely positive maps were discovered in [21] by Stinespring. For their physical elucidation cf., e.g., [8] by Kraus.

There exist non-completely positive maps. Anti-unitary maps are such, like e.g. the reflection $\mathcal{T}$ (7.30) on the two-state q-systems. If we try to construct the Kraus representation we fail. Actually, we would get negative sign for some of the “sandwich” terms. The reflection $\mathcal{T}$ can be written typically into this form:

$$\mathcal{T}\hat{\rho} = \frac{1}{2}\hat{\sigma}\hat{\rho}\hat{\sigma} - \frac{1}{2}\hat{\rho}. \quad (8.2)$$

By no means can it be written into the Kraus form (8.1). It is obvious that $\mathcal{T}$ is non-completely positive if we recall Sect. 7.2.2 where we saw that the trivial extension $\mathcal{T} \otimes \mathcal{I}$ for 2×2 dimension was not a positive map.

The non-completely positive maps can not be realized. As we shall see in the next two Sections, the completely positive maps can, at least in principle, be realized in many different ways. What we called q-operations 4.2 is nothing else but the completely positive maps.

8.2 Reduced dynamics

Consider the composition of the given q-system and a certain, maybe fictitious, environmental q-system E , and introduce a unitary dynamics $\hat{U}$ on the composite system. Let us then reduce the state to the original system, The resulting q-operation is usually non-unitary:

$$\mathcal{M}\hat{\rho} = \text{tr}_E \left[\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)\hat{U}^\dagger \right]. \quad (8.3)$$

The same q-operation $\mathcal{M}$ can be obtained from different unitary dynamics $\hat{U}$, moreover, the choice of the environmental system can also be varied. All trace-preserving

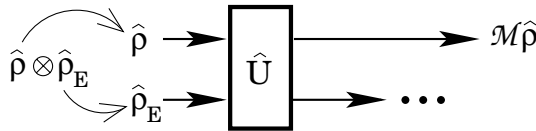


Fig. 8.1. Reduced q-dynamics vs. q-operations. We let our system in question interact unitarily with a (real or fictitious) environmental system. We suppose the uncorrelated initial state $\hat{\rho} \otimes \hat{\rho}_E$ which the interaction transforms into the q-correlated one $\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)$. Finally we ignore the environmental system and we only consider the reduced final state $\text{tr}_E[\hat{U}(\hat{\rho} \otimes \hat{\rho}_E)]$. The unitary interaction with the environment has thus induced a non-unitary dynamics for our system, which we call reduced dynamics. It corresponds to a completely positive map $\mathcal{M}$. The inverse relationship is also true: A q-operation $\mathcal{M}$ (i.e.: a completely positive map) can always be realized by a suitable reduced dynamics.

q-operations can be obtained as reduced dynamics (8.3). The proof can be done in Kraus representation (8.1). Let us introduce a basis for the environmental system E :

$$|n; E\rangle, \quad n = 1, 2, \dots, d_E. \quad (8.4)$$

Suppose the initial state of E is pure:

$$\hat{\rho}_E = |1; E\rangle \langle 1; E|. \quad (8.5)$$

Then the q-operation (8.3) takes the Kraus form:

$$\mathcal{M}\hat{\rho} = \sum_n \hat{M}_n \hat{\rho} \hat{M}_n^\dagger, \quad (8.6)$$

if we insert the following Kraus matrices²:

$$\hat{M}_n = \text{tr}_E \left[(\hat{I} \otimes |1; E\rangle \langle n; E|) \hat{U} \right] \equiv \langle n; E| \hat{U} |1; E\rangle. \quad (8.7)$$

If, the other way around, we know the Kraus form of a trace-preserving q-operation then the equations (8.7) are always solvable for the unitary dynamics $\hat{U}$ if we assume an environmental system E of suitable large dimensions.

8.3 Indirect measurement

Each “sandwich” term of the Kraus form can be interpreted separately. Suppose that we perform a non-selective q-measurement on the environmental system E . If we do it *after* the unitary interaction $\hat{U}$ then it does not influence the reduced state of the original system. If, however, our measurement is selective then we can arrive at the individual “sandwich” terms of the Kraus form. To see this, we start from the basis (8.4) of E , we form the corresponding partition $\hat{P}_{En} = |n; E\rangle \langle n; E|$ and extend it trivially for the whole composite system:

$$\hat{P}_n = \hat{I} \otimes \hat{P}_{En}. \quad (8.8)$$

Perform the projective measurement of these physical quantities, after the unitary interaction $\hat{U}$. The concept is called indirect measurement. We must note of course that the measurement of the projectors (8.8) is equivalent with the measurement of the projectors $\hat{P}_{En}$ on E . Follow the rules of selective measurement (4.18, 4.13). The state of the composite system changes as follows:

$$\hat{\rho} \otimes \hat{\rho}_E \rightarrow \frac{1}{p_n} \hat{P}_n \hat{U} (\hat{\rho} \otimes \hat{\rho}_E) \hat{U}^\dagger \hat{P}_n, \quad (8.9)$$

while the probability distribution of the outcome is:

² To avoid misunderstandings, we note the natural shorthand notation $\langle n; E| \hat{U} |1; E\rangle$ stands for a matrix acting on the states $\hat{\rho}$ of the original system.

$$p_n = \text{tr} \left[\hat{P}_n \hat{U} (\hat{\rho} \otimes \hat{\rho}_E) \hat{U}^\dagger \right] . \quad (8.10)$$

Now we reduce the first equation to the state of the original q-system; the operation tr_E yields:

$$\hat{\rho} \rightarrow \frac{1}{p_n} \hat{M}_n \hat{\rho} \hat{M}_n^\dagger , \quad (8.11)$$

cf. the definition (8.7) of $\hat{M}_n$. The second equation takes the equivalent form:

$$p_n = \text{tr} \left[\hat{M}_n^\dagger \hat{M}_n \hat{\rho} \right] . \quad (8.12)$$

As we see, each “sandwich” term in the Kraus form (8.1) of our q-operation corresponds, apart from a normalizing factor, to what comes out from the selective measurement of the environmental system’s basis (8.4). The inverse of the renormalizing factor is always equal to the probability of the given outcome.

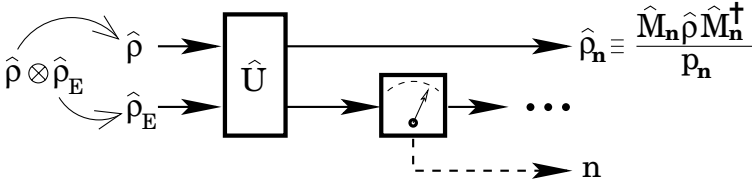


Fig. 8.2. Indirect q-measurement vs. selective q-operation. The setup is a refinement of the reduced dynamics. The difference is that after the unitary interaction we do not ignore the state of the environment. Rather we perform a projective measurement on it. The ensemble of pre-measurement state $\hat{\rho}$ is selected into sub-ensembles of conditional post-measurement states $\hat{\rho}_n$ according to the obtained measurement outcomes n . The probability p_n coincides with the norm of the unnormalized conditional state $\hat{M}_n \hat{\rho} \hat{M}_n^\dagger$. The sub-ensembles of certain conditional post-measurement states $\{\hat{\rho}_n; n \in \Omega\}$ can be re-united contributing to a trace-reducing completely positive map. The inverse relationship is also true: A selective q-operation $\mathcal{M}$ (i.e.: a trace-reducing completely positive map) can always be realized by a suitable indirect measurement.

We got the key to the interpretation of trace-reducing q-operations. Let us make a selection Ω of the measurement outcomes and discard the complementary part of the q-ensemble. Such q-operation can be written into this form:

$$\mathcal{M}_\Omega \hat{\rho} = \sum_{n \in \Omega} \hat{M}_n \hat{\rho} \hat{M}_n^\dagger ; \quad \sum_{n \in \Omega} \hat{M}_n^\dagger \hat{M}_n < \hat{I} , \quad (8.13)$$

where this time the summation over the index n is only partial. The obtained state is not normalized:

$$\text{tr} [\mathcal{M}_\Omega \hat{\rho}] = \sum_{n \in \Omega} p_n . \quad (8.14)$$

The trace of the state is equal to the total probability of the selected measurement outcomes $n \in \Omega$, i.e., it can be less than 1.

The physical interpretation of q-operations has thus been completed. Each trace-preserving q-operation can be realized as the reduced dynamics 8.2 after the suitable unitary interaction with a suitable environmental system. Each trace-compressing q-operation can be realized by a suitable selective indirect measurement via the above environmental system. These realizations are never unique, we can have infinite many choices.

8.4 Non-projective measurement resulting from indirect measurement

Recall the concept of projective 4.4.1 and non-projective 4.4.2 q-measurements. Projective measurements are also called von Neumann, standard, or ideal measurements. Non-projective ones, also called non-ideal or unsharp, were formal extensions of projective ones. Now we can easily prove that they are special cases of indirect measurements. Indeed, (8.11,8.12) will coincide with (4.20,4.22) of non-projective measurement in the special case of hermitian Kraus matrices. Then we can always identify the effects by the square of the Kraus matrices:

$$\hat{H}_n = \hat{M}_n^2, \quad \hat{M}_n = \hat{M}_n^\dagger = \hat{H}_n^{1/2}. \quad (8.15)$$

This way, the mechanism of indirect measurements will underly the concept of non-projective measurements. The non-projective measurement of effects $\hat{H}_n$ can always be realized by indirect measurements which are just projective measurements that happen on a certain, maybe fictitious, environmental system E q-correlated with the system in question.

Note that there exists an equivalent projective measurement on the uncorrelated composition $\hat{\rho} \otimes \hat{\rho}_E$ of the system and the environment. Only we have to relax the special form $\hat{P}_n = \hat{I} \otimes \hat{P}_{nE}$ of the projectors and take the unitarily equivalent set:

$$\hat{P}_n = \hat{U}^\dagger \hat{I} \otimes \hat{P}_{nE} \hat{U}. \quad (8.16)$$

This time the projective measurement happens on the composite state $\hat{\rho} \otimes \hat{\rho}_E$ as a whole, not only on the environment:

$$\hat{\rho} \otimes \hat{\rho}_E \rightarrow \frac{1}{p_n} \hat{P}_n \hat{\rho} \otimes \hat{\rho}_E \hat{P}_n, \quad (8.17)$$

$$p_n = \text{tr} \left[\hat{P}_n (\hat{\rho} \otimes \hat{\rho}_E) \right]. \quad (8.18)$$

Nonetheless, the measurement is equivalent, due to unitary equivalence, with the indirect measurement (8.8,8.9,8.10) defined previously. The projective measurement of $\hat{P}_n$ yields the non-projective measurement of the effect $\hat{H}_n$ on the reduced state $\hat{\rho}$. If $\hat{P}_n$ yields 1 or 0 then it means that $\hat{H}_n$ yields 1 or 0 in coincidence. While the

former expresses sharp information the latter does not. Projective measurement is repeatable, non-projective is not. When we average over the environmental state we get averaged data. The effects, too, are averaged version of the projectors:

$$\hat{\Pi}_n = \text{tr}_E \left[\hat{P}_n (\hat{I} \otimes \hat{\rho}_E) \right] . \quad (8.19)$$

Of course, this expression is equivalent with (8.15).

8.5 Entanglement and LOCC

As we saw earlier in Sect. 7.1.4 of the preceding Chapter, local q-dynamics and q-measurements can not create q-correlation. Having learned the class of general q-operations 8.1, we can now formulate the precise statement. In the postulated situation, Alice and Bob own their separate local systems far from each other and the remote systems form a composite q-system together. Alice and Bob are allowed to perform local operations (LO). When they do so independently then the composite system changes this way:

$$\hat{\rho}_{AB} \longrightarrow (\mathcal{M}_A \otimes \mathcal{M}_B) \hat{\rho}_{AB} . \quad (8.20)$$

Usually we suppose classical communication (CC) as well. Alice and Bob can inform each other about the setups and the outcomes of their local operations so that they can condition their further LOs on the received classical information. Obviously, the resulting q-operation is always the product of independent operations or the statistical mixture of them:

$$\hat{\rho}_{AB} \rightarrow \sum_n w_n (\mathcal{M}_{An} \otimes \mathcal{M}_{Bn}) \hat{\rho}_{AB} ; \quad \sum_n w_n = 1 . \quad (8.21)$$

An independent LO (8.20) does not create any correlation between the systems of Alice and Bob. The statistical mixture (8.21) of independent LOs, which LOCC results in, is able to produce but classical correlation. Q-correlation (entanglement), if it is not already present, can never be produced by LOCC. To see this, suppose the initial state is unentangled, i.e., separable. Then it can be written as a mixture of independent states:

$$\hat{\rho}_{AB} = \sum_{\lambda} w^{\lambda} \hat{\rho}_{A\lambda} \otimes \hat{\rho}_{B\lambda} ; \quad w^{\lambda} \geq 0 ; \quad \sum_{\lambda} w^{\lambda} = 1 . \quad (8.22)$$

When a generic LOCC operation (8.21) is performed on this separable state:

$$\hat{\rho}_{AB} \rightarrow \sum_n w_n (\mathcal{M}_{An} \otimes \mathcal{M}_{Bn}) \hat{\rho}_{AB} = \sum_{n,\lambda} w_n w^{\lambda} (\mathcal{M}_{An} \hat{\rho}_{A\lambda}) \otimes (\mathcal{M}_{Bn} \hat{\rho}_{B\lambda}) , \quad (8.23)$$

it also results in a mixture of independent states which therefore remains separable.

8.6 Open q-system: master equation

When a q-system is continuously interacting with certain environmental q-system or q-systems then it is called open q-system. A subsystem of a closed system makes a typical open system. The open system dynamics is reduced dynamics and is thus irreversible. There exists an idealized (so-called Markovian) class of environmental interactions where the state would still satisfy an equation similar to von Neumann's (4.7). The Markovian master equation takes this form:

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \text{non-Hamiltonian terms} \equiv \mathcal{L}\hat{\rho}. \quad (8.24)$$

It should generate a completely positive map $\mathcal{M}(t) = \exp(t\mathcal{L})$. This condition implies that the master equation can always be written into Lindblad form³ (although we do not prove it here):

$$\frac{d\hat{\rho}}{dt} \equiv \mathcal{L}\hat{\rho} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \sum_n \left(\hat{L}_n \hat{\rho} \hat{L}_n^\dagger - \frac{1}{2} \{ \hat{L}_n^\dagger \hat{L}_n, \hat{\rho} \} \right). \quad (8.25)$$

The supermatrix $\mathcal{L}$ is called the Lindblad generator. Given a master equation by $\mathcal{L}$, its Lindblad form is never unique: the Lindblad matrices $\hat{L}_n$ — forming the non-Hamiltonian term — can be chosen in many different ways. The physical interpretation of Lindblad matrices is related to the form of the interaction between the open system and its environment. Choice of Lindblad matrices of a given open system is non-unique because the same reduced dynamics 8.2 of the open system can be obtained by various interactions with various environments.

8.7 Q-channels

A typical theoretical application of q-operations $\mathcal{M}$ is the concept of a q-channel. A q-channel serves to communicate an arbitrary q-state $\hat{\rho}$ from one location to another, say, from Alice to Bob like in the protocols of the q-banknote 6.4.1, q-key distribution 6.4.2, or superdense coding 7.3.1. In the ideal case, these protocols use noiseless q-channels: the output state coincides with the input one. In general, however, the noisy q-channel will distort the state, it performs a certain q-operation $\mathcal{M}$ on the input. Terminologically, we identify the operation $\mathcal{M}$ and the q-channel. Q-channels are categorized according to the distortions they cause. Elementary q-channels are those communicating single independent qubits. The depolarizing channel is this:

$$\hat{\rho} \rightarrow \mathcal{M}\hat{\rho} = (1 - w)\hat{\rho} + w\frac{\hat{I}}{2}, \quad (8.26)$$

it damps the polarization of a qubit isotropically by a factor $1 - w$. The dephasing channel takes this form:

³ Cf. [29] by Lindblad and [30] by Gorini, Kossakowski and Sudarshan.

$$\hat{\rho} \rightarrow \mathcal{M}\hat{\rho} = (1 - w)\hat{\rho} + w\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z, \quad (8.27)$$

which decreases the off-diagonal elements of $\hat{\rho}$ in the computational basis by a factor $1 - w$.

Problems, exercises

8.1. All q-operations are reductions of unitary dynamics. Consider a trace-preserving q-operation $\mathcal{M}$. It is equivalent with a certain reduced dynamics where the system interacts unitarily with a “fictitious” environmental system of suitably large dimension d_E . Let us prove this equivalence starting from the Kraus representation of $\mathcal{M}$ and show that it suffices if we choose d_E to coincide with the number of the Kraus matrices $\hat{M}_n$. Method: introduce the bases $|\lambda\rangle, |n; E\rangle$ for the system and the environment, respectively, and observe that the states $\sum_n \hat{M}_n |\lambda\rangle \otimes |n; E\rangle$ form an orthonormal set.

8.2. Non-projective effect as averaged projection. Let us write the uncorrelated composite state of the system and the environmental system as $\hat{\rho}\hat{\rho}_E$. We can many times omit the notation of $\otimes$ as well as $\hat{I}$ without risk of misunderstanding. In this example we use such convenience. Suppose we perform a projective measurement of the partition $\{\hat{P}_n\}$ of the composite system. If we ignore the post-measurement state of the environment, we are left with a given non-projective measurement of the positive decomposition $\{\hat{I}_n\}$ on the system alone. Let us prove that $\hat{I}_n = \text{tr}_E(\hat{P}_n\hat{\rho}_E)$.

8.3. Q-operation as supermatrix. Consider an operation $\mathcal{M}$ which transforms density matrices into density matrices. In a given basis, the density matrices are represented by their components $\rho_{\lambda\mu} = \langle\lambda|\hat{\rho}|\mu\rangle$ and the operation is represented by the components of a supermatrix of four indices:

$$\rho_{\lambda\mu} \longrightarrow \sum_{\lambda'\mu'} \mathcal{M}_{\lambda\mu\lambda'\mu'} \rho_{\lambda'\mu'}.$$

Let us express the supermatrix $\mathcal{M}$ by utilizing the Kraus representation.

8.4. Environmental decoherence, time-continuous depolarization. If we are interested in the complete isolation of our q-system then all environmental interactions are considered as parasitic. They tend to destroy superpositions and entanglements within our system. They will paralyze, in particular, q-protocols and q-algorithms that take long time enough to accumulate the above environmental decoherence. A simplest model of an uncontrollable environmental interaction is the time-continuous depolarization of the qubit. Let us show that the master equation

$$\frac{d\hat{\rho}}{dt} = -\frac{1}{4\tau}(\hat{\sigma}\hat{\rho}\hat{\sigma} - 3\hat{\rho})$$

is of the Lindblad form and it describes isotropic depolarization. Derive the equation of motion for the polarization vector $\mathbf{s}$.

8.5. Kraus representation of depolarization. Let us construct Kraus representation for the depolarization channel. Method: exploit the isotropy of the channel.

9 Classical information theory

Shannon entropy is the key-notion of classical information. It provides the statistical measure of information associated with states ρ . Since dynamical aspects shall not be treated at all, we would just talk about probability distributions p instead of physical states ρ . For comparability with Q-information theory of Chap. 10, however, we keep talking about states ρ of classical systems. Typically, we use heuristic proofs though corner stones of the exact derivations will fairly be indicated.

9.1 Shannon entropy, mathematical properties

Shannon entropy in statistical physics used to be introduced for a classical state $\rho(x)$ defined on continuous phase space $\{x\}$. For convenience of classical information theory, however, we shall use states $\rho(x)$ defined on discrete space $\{x\}$, cf. Sect. 2.7.

Shannon entropy of a given state will be defined as follows:

$$S(\rho) = -\langle \log \rho \rangle = -\sum_x \rho(x) \log \rho(x) . \quad (9.1)$$

The Shannon entropy takes its minimum value 0 if the state is pure, say $\rho(x) = \delta_{x0}$. It takes its maximum value if the state is random: $\rho(x) = \text{const}$.

Mixing increases the entropy. The entropy of a weighted mixture is greater than (or equal to) the weighted sum of the constituents' entropies (concavity):

$$S(w_1\rho_1 + w_2\rho_2) \geq w_1S(\rho_1) + w_2S(\rho_2) ; \quad w_1 + w_2 = 1. \quad (9.2)$$

Correlation decreases the entropy. The entropy of a bipartite composite state is less than (or equal to) the sum of the entropies of the two reduced states. The equality only holds when the two subsystems are uncorrelated (subadditivity):

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) . \quad (9.3)$$

Reduction decreases the entropy. The entropy of reduced states is less than (or equal to) the entropy of the composite state:

$$S(\rho_{AB}) \geq S(\rho_A), S(\rho_B) . \quad (9.4)$$

The relative entropy can be defined for two states ρ, ρ' that belong to the same system:

$$S(\rho' \parallel \rho) = -S(\rho) - \sum_x \rho(x) \log \rho'(x) .$$

The relative entropy is zero if $\rho = \rho'$ and it is positive otherwise.

9.2 Messages

According to the abstraction of classical information, a finite sequence of *letters* taken from an *alphabet* constitutes a message. Messages can be used to store or communicate the encoded data. The simplest alphabet is binary, it has two letters usually associated with 0 and 1, respectively. The binary letters are also called bits. Binary messages are sequences of binary digits and can be thought of integers in binary representation. The unit of classical information measure is one *random* bit, or one bit in short (meaning always one random bit). The information of a random n -letter binary message is thus n bits. If the alphabet is not binary the messages can still be faithfully translated into binary messages. One random letter of the K -letter alphabet can, on average, be converted into $\log K$ random binary letters (bits), i.e., into a $\log K$ -letter long random binary message. A shorter binary message could not encode the original letter faithfully. Therefore we say that a random letter of a K -letter alphabet contains $\log K$ bits information.

If the letter is not random it can be compressed faithfully into less than $\log K$ random bits. The original message is redundant and the measure of information is less than $n \log K$ bits. In general, we define the information of a message by the length of the corresponding shortest faithfully compressed binary message. According to Shannon's theory, this shortest length is $nS(\rho)$ in the limit $n \rightarrow \infty$ ¹. We outline the proof.

9.3 Data compression

Consider an n -letter long message of a K -letter alphabet $\{a, b, c, \dots, z\}$:

$$x_1 x_2 \dots x_n , \quad (9.5)$$

and suppose that the letters independently follow the same apriori probability distribution:

$$\rho(a) = p_1, \rho(b) = p_2, \rho(c) = p_3, \dots, \rho(z) = p_K . \quad (9.6)$$

Physically, each letter is represented by a discrete K -state classical system and the statistics of n -letter long messages is represented by the collective state:

$$\rho(x_1) \rho(x_2) \dots \rho(x_n) \equiv \rho^{\times n}(x_1, x_2, \dots x_n) . \quad (9.7)$$

The letters are random if $\rho = \text{const}$, otherwise the data contained in the above collective state can be compressed.

¹ Cf. [31] by Shannon.

Indeed, many combinatorically possible messages are totally unlikely and become negligible if $n \rightarrow \infty$. Only a portion of the combinatorically possible messages will dominate statistically. These are called *typical messages*. In them, the relative frequency of each letter is identical to its apriori probability (9.6)! If in a very long message this rule were violated for any letter then the message would become atypical and thus statistically irrelevant. Therefore in typical messages the letters of the alphabet occur with frequencies

$$np_1, np_2, \dots, np_K. \quad (9.8)$$

The typical messages can only differ from each other by the order of the letters! All typical messages have identical apriori probability. It is thus sufficient if we calculate the number of the typical messages, their probability distribution is even and equals the inverse of their number.

The number of the typical messages equals the number of polynomial combinations of n elements satisfying the frequencies (9.8):

$$\# \text{ of typical messages} = \frac{n!}{(np_1)!(np_2)! \dots (np_K)!} . \quad (9.9)$$

If n is large, Stirling approximation

$$\log n! = n \log n - n \log e + O(\log n) \quad (9.10)$$

applies to all factorials, yielding:

$$\log \frac{n!}{(np_1)!(np_2)! \dots (np_K)!} \approx -n \sum_{k=1}^K p_k \log p_k . \quad (9.11)$$

The r.h.s. is just n -times the Shannon entropy of a single (non-random) letter. Accordingly, the number of typical messages reads:

$$\# \text{ of typical messages} = 2^{nS(\rho)} . \quad (9.12)$$

Their distribution is even. The collective state $\rho^{\times n}$ of the original messages can, information theoretically, be replaced by the even distribution of the typical messages:

$$\rho^{\times n} \longrightarrow \frac{1}{2^{nS(\rho)}} \Omega_n , \quad (9.13)$$

where $\Omega_n(x_1, x_2, \dots, x_n)$ is the indicator function which is 1 for the $2^{nS(\rho)}$ typical messages and zero otherwise. The typical ones of the original redundant messages (9.5) can therefore be numbered by $nS(\rho)$ binary digits constituting $nS(\rho)$ -bit random messages. These binary messages are not redundant, can not be further compressed faithfully. We have in such a way obtained the shortest faithful binary compression. Each letter of the original message has been faithfully compressed from $\log K$ bits into $S(\rho)$ bits.

This data compression theorem is the information theoretic interpretation of the Shannon entropy. The theorem as well as its interpretation are asymptotic, valid for one letter in infinite long redundant messages. This explains the success of the above heuristic proof. It must be obvious that for finite n the frequencies of the letters would fluctuate around the asymptotic values (9.8), the set of typical messages might increase. Nevertheless, the exact proof would confirm that the above heuristic one has yielded the correct asymptotic number of the statistically relevant messages.

9.4 Mutual information

Let us introduce the notation $X = \{x, \rho(x)\}$ for a single-letter message where $\rho(x)$ is the apriori distribution of single letters. Shannon entropy $S(\rho)$ has been attributed to the distribution $\rho(x)$ of single letters. For terminological and notational conveniences we say that this is the entropy of the (single-letter) message X and we introduce the alternative notation $H(X)$ for it:

$$H(X) = S(\rho) . \quad (9.14)$$

The Shannon entropy measures the rate of our apriori ignorance concerning long messages. We must learn $H(X)$ bits per letter to eliminate our ignorance perfectly.

Suppose two separate messages X and Y are letter-wise correlated according to a certain distribution $\rho(x, y)$. Let $H(X, Y)$ stand for the joint Shannon entropy. We define the mutual information of the two messages:

$$I(X : Y) = H(X) + H(Y) - H(X, Y) = \left\langle \log \frac{\rho(x, y)}{\rho(x)\rho(y)} \right\rangle . \quad (9.15)$$

This is zero if the two messages X and Y are uncorrelated. Otherwise it is positive. The mutual information $I(X : Y)$ gives the number of bits, per composite letter (x, y) , that we can spare when we compress the joint message (X, Y) instead of compressing X and Y separately.

The mutual information is related to the *conditional entropy*. While $I(X : Y)$ is symmetric function of the two messages the conditional entropy is not. The conditional entropy of the message X is built upon the conditional probability $\rho(x|y) = \rho(x, y)/\rho(y)$:

$$H(X|Y) = -\langle \log \rho(x|y) \rangle = H(X, Y) - H(Y) . \quad (9.16)$$

Conditional entropy means the minimum number of bits (per letter) needed to encode the message X if the message Y is perfectly known. $H(X|Y) = H(X)$ if the two messages are uncorrelated, and $H(X|Y) = 0$ if they are perfectly correlated. The mutual information and the two conditional entropies are simply related:

$$I(X : Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) . \quad (9.17)$$

9.5 Channel capacity

Let X be the input message of a noisy channel and let Y be its output message. The channel is characterized by the *transfer probability* $\rho(y|x)$. Then the joint distribution of the input and output letters becomes:

$$\rho(x, y) = \rho(y|x)\rho(x) . \quad (9.18)$$

In fact, we can learn the output message Y and, from it, we have to estimate the input message X . According to Bayes rule, we write down the aposteriori distribution of the input letter conditioned on the received output letter:

$$\rho(x|y) = \frac{\rho(x, y)}{\rho(y)} = \frac{\rho(y|x)\rho(x)}{\sum_{x'} \rho(y|x')\rho(x')} . \quad (9.19)$$

This becomes proportional to the transfer probability $\rho(y|x)$ of the channel if $\rho(x) = \text{const}$, i.e., when the input message is random:

$$\rho(x|y) = \frac{\rho(y|x)}{\sum_{x'} \rho(y|x')} . \quad (9.20)$$

The mutual information $I(X : Y)$ (9.17) indicates how many bits of the input message X can faithfully be decoded from the perfect knowledge of one output letter. If the channel is noiseless, $\rho(y|x) = \delta_{xy}$, then $H(X) = H(Y) = H(X, Y) = I(X : Y)$ meaning that the output yields the input perfectly. In the contrary case, the channel noise makes the output message random: $\rho(y|x) = \text{const}$. Then $H(X) = H(X|Y)$ and $I(X : Y) = 0$ so we cannot estimate the input at all.

In the intermediate cases we obtain partial information on the input. Out of the $H(X)$ bits information per input letter, the output still contains $I(X : Y)$ bits information per letter. Those $I(X : Y)$ bits are in general not sufficient to faithfully estimate the input letter itself. Yet, the $I(X : Y)$ bits represent faithful partial information on the input message. The given channel communicates $I(X : Y)$ bits/letter faithful information. This rate depends on the transfer probability $\rho(y|x)$ of the channel and on the distribution $\rho(x)$ of the input letters, see also (9.17, 9.18). By an optimum choice of $\rho(x)$ we can maximize the the rate of faithful communication. This rate is called the channel capacity:

$$C \equiv \max_{\{\rho(x)\}} I(X : Y) , \quad (9.21)$$

which is the function of the transfer probability $\rho(y|x)$ alone.

9.6 Optimal codes

Shannon entropy $H(X)$ of a message X has been interpreted as the bit-length of the shortest faithful code per letter. Our heuristic proof 9.3 could be made exact.

First, one has to construct the block-code of n -letter long messages in such a way that, in the asymptotic limit $n \rightarrow \infty$ the code length is $\sim nH(X)$ bits. Note that the block-code can not be decomposed into letter-wise codes. Second, one has to prove that this code is faithful, i.e., the decoding yields the original message at error rate which goes to zero for $n \rightarrow \infty$. Third, one has to prove that further data compression would necessarily violate faithfulness. Shannon has done the exact proof, that is his noiseless coding theorem.

In a similar way could we make the heuristic proof of channel capacity 9.5 exact. We consider nC -bit random messages to communicate through the channel faithfully. First we encode them properly into the $\sim n$ -letter long input messages X of the channel. Note that $C \leq H(X)$, i.e., the chosen code must be redundant if the channel is noisy. Second, we prove that, in the limit $n \rightarrow \infty$, decoding the output message Y yields the initial nC bits faithfully. Third, we prove that for the given channel any higher rate of faithful communication is impossible. This is Shannon's noisy code theorem.

9.7 Cryptography and information theory

The cryptography situation 6.4.2 can be interpreted in terms of classical information theory. There are three parties: Alice, Bob, and Eve. Alice generates the random key X , Bob estimates a certain key Y which is a noisy version of X , while Eve steals a key Z which is another noisy version of X . The quality of a given key-distribution protocol, whether classical or quantum, depends on the three mutual informations $I(X:Z)$, $I(Y:Z)$ and $I(X:Y)$. The success and faithfulness of key-distribution is measured by $I(X:Y)$, the loss of security is measured by $I(X:Z)$ and $I(Y:Z)$. In the ideal case $I(X:Y) = H(X) = H(Y)$ and $I(X:Z) = I(Y:Z) = 0$. The ultimate security of q-key protocols against unnoticed evedropping is based, among others, on the analysis of these mutual informations².

9.8 Entropically irreversible operations

We introduced the qualitative notion of irreversibility on examples, e.g., of the reduced dynamics $\mathcal{M}(t)$ in Sect. 2.5 where it meant just the non-invertibility of the operation $\mathcal{M}(t)$. The informational concept of irreversibility is more restrictive than the concept of non-invertibility. It assumes information loss or, in alternative terminology: entropy increase. In typical reduced dynamics, e.g., the entropy for some initial states is increasing while for some other ones it is decreasing. Below, we are going to characterize the general class of entropy increasing operations.

A non-invertible operation $\mathcal{M}$ is (entropically) irreversible if it increases or preserves the entropy:

$$S(\mathcal{M}\rho) \geq S(\rho) , \quad (9.22)$$

² Cf., e.g., in [18].

for all states ρ . We shall show that a non-invertible $\mathcal{M}$ is irreversible if it leaves the totally mixed state $\rho_0(x) = \text{const}$ invariant: $\mathcal{M}\rho_0 = \rho_0$.

On a discrete phase space, the class of invertible operations is narrow: it corresponds to the permutations of the discrete points $\{x\}$ of the state space. All other discrete classical operations $\mathcal{M}$ are non-invertible. We are going to show that $\mathcal{M}$ increases (or at least preserves) the entropy for all $\hat{\rho}$ if $\mathcal{M}\hat{\rho}_0 = \hat{\rho}_0$. The general form of an operation $\rho \rightarrow \rho' \equiv \mathcal{M}\rho$ reads: $\rho'(y) = \sum_x \rho(y|x)\rho(x)$ where $\rho(y|x)$ is the non-negative transfer function, satisfying $\sum_y \rho(y|x) = 1$ for all x . From the condition $\mathcal{M}\rho_0 = \rho_0$ it follows that our transfer function is normalized in x as well: $\sum_x \rho(y|x) = 1$ for all y . Such transfer function is also called doubly-stochastic. We can always write:

$$S(\rho') - S(\rho) = \sum_{x,y} \rho(x)\rho(y|x) \log \frac{\rho(x)}{\rho'(y)} . \quad (9.23)$$

We invoke the inequality $\ln \lambda > 1 - \lambda^{-1}$ valid for $\lambda \neq 1$ and apply it to $\lambda = \rho/\rho'$. This yields:

$$S(\rho') - S(\rho) \geq \frac{1}{\ln 2} \sum_{x,y} \rho(y|x)[\rho(x) - \rho'(y)] = 0 . \quad (9.24)$$

In general, the strict inequality holds for $\rho \neq \rho_0$ unless the support of ρ is smaller than the whole state space and the restriction of the non-invertible $\mathcal{M}$ for that support becomes invertible.

Let us see the simplest example. Consider a noisy channel fed by (totally) random input message. If the channel yields totally random output message than it is entropically irreversible, it will increase the entropy for all possible messages apart from the marginal cases mentioned above. As a contrary example, consider a “broken” channel whose output message is a permanent 0 for any input whatsoever. The channel operation is obviously non-invertible but it is not entropically irreversible.

Problems, exercises

9.1. Positivity of relative entropy. Using the inequality $\ln \lambda > 1 - \lambda^{-1}$, let us prove the positivity of the relative entropy.

9.2. Concavity of entropy. Why and how does mixing increase the entropy? Let us argue information theoretically. Method: Start from two messages of different lengths proportional to w_1 and w_2 , and of different single-letter distributions ρ_1, ρ_2 , respectively.

9.3. Subadditivity of entropy. Using the positivity of the relative entropy, let us prove the subadditivity of the entropy.

9.4. Coarse graining increases entropy. Consider a given classical state, i.e. probability distribution $\rho(x)$ defined on the period $[0, 1]$ represented to a precision of k binary digits: $x \equiv 0.x_1x_2 \dots x_k$. Let us introduce the coarse-grained state

$$\tilde{\rho}(\tilde{x}) = \sum_{x_k=0,1} \rho(x) ,$$

defined on the same state space $[0, 1]$ to a precision of one digit less, i.e., we introduced $\tilde{x} \equiv 0.x_1x_2 \dots x_{k-1}$. Let us prove that coarse graining increases the entropy: $S(\tilde{\rho}) \geq S(\rho)$. Method: use the fact that state reduction increases the entropy.

10 Q-information theory

The q-information theory is similar to the classical one. The carrier of information is a q-system and this causes differences as well. A particular difference is that qubits carry information in double sense. First, we can encode classical bits into qubits. Second, unknown qubits carry information which is hidden and protected — as a consequence of the universal limitations of single q-state determination, cloning, or distinguishability.

10.1 Von Neumann entropy, mathematical properties

Von Neumann entropy, or q-entropy, of a given q-state $\hat{\rho}$ will be defined as follows:

$$S(\hat{\rho}) = -\langle \log \hat{\rho} \rangle = -\text{tr}(\hat{\rho} \log \hat{\rho}) . \quad (10.1)$$

Obviously, the q-entropy is unitary invariant:

$$S(\hat{\rho}) = S(\hat{U} \hat{\rho} \hat{U}^\dagger) . \quad (10.2)$$

The von Neumann entropy takes its minimum value zero if the state is pure, i.e.: $S(|\psi\rangle\langle\psi|) = 0$. It takes its maximum value on the maximally mixed state when $S(\hat{I}/d) = \log d$.

Let us construct the spectral expansion of the q-state $\hat{\rho}$:

$$\hat{\rho} = \sum_{\lambda} \rho_{\lambda} |\varphi_{\lambda}\rangle \langle\varphi_{\lambda}| . \quad (10.3)$$

The eigenvalues ρ_{λ} of $\hat{\rho}$ form a probability distribution and its Shannon entropy equals the von Neumann entropy of the q-state $\hat{\rho}$:

$$S(\rho) = -\sum_{\lambda} \rho_{\lambda} \log \rho_{\lambda} = S(\hat{\rho}) . \quad (10.4)$$

Mixing increases the q-entropy. The q-entropy of a weighted mixture is greater than (or equal to) the weighted sum of the constituents' q-entropies (concavity):

$$S(w_1 \hat{\rho}_1 + w_2 \hat{\rho}_2) \geq w_1 S(\hat{\rho}_1) + w_2 S(\hat{\rho}_2) ; \quad w_1 + w_2 = 1. \quad (10.5)$$

Correlation decreases the q-entropy. The q-entropy of a bipartite composite q-system is less than (or equal to) the sum of the q-entropies of the two reduced q-states. The equality only holds when the two q-subsystems are uncorrelated (sub-additivity):

$$S(\hat{\rho}_{AB}) \leq S(\hat{\rho}_A) + S(\hat{\rho}_B) . \quad (10.6)$$

Reduction may change q-entropy in both directions, contrary to classical entropy. The difference between the q-entropies of the two reduced q-states can not be greater than the entropy of the bipartite composite q-state (triangle inequality):

$$S(\hat{\rho}_{AB}) \geq |S(\hat{\rho}_A) - S(\hat{\rho}_B)| . \quad (10.7)$$

This is a genuine q-feature. The Shannon entropy of a classical sub-system is never greater than the entropy of the composite system (9.4):

$$S(\rho_{AB}) \geq S(\rho_A), S(\rho_B) . \quad (10.8)$$

Hence the information content of a classical sub-system is smaller than (or equal to) the composite system's entropy. For q-systems this shall not be true. Typically, a pure entangled q-state has zero q-entropy while its subsystems are in mixed reduced states of positive q-entropy. Contrary to quantum, the classical pure composite states are so trivial that their subsystems remain pure.

The relative q-entropy can be defined for two states $\hat{\rho}, \hat{\rho}'$ that belong to the same q-system:

$$S(\hat{\rho}' || \hat{\rho}) = -S(\hat{\rho}) - \text{tr}(\hat{\rho} \log \hat{\rho}') .$$

The relative entropy is zero if $\hat{\rho} = \hat{\rho}'$ and it is positive otherwise. The inequality $S(\hat{\rho}' || \hat{\rho}) \geq 0$ is also called as the Klein inequality.

10.2 Messages

Consider a classical message $X = \{x, \rho(x)\}$ which we encode letter-wise into certain, not necessarily orthogonal, pure states of a given d-state q-system:

$$\begin{array}{llll} \text{letter:} & x_1 & x_2 & \dots & x_n \\ \text{q-code:} & |x_1\rangle & |x_2\rangle & \dots & |x_n\rangle \\ \text{probability:} & \rho(x_1) & \rho(x_2) & \dots & \rho(x_n) \end{array} \quad (10.9)$$

The sequence $|x_1\rangle, |x_2\rangle, \dots, |x_n\rangle$ is a q-message of length n . Q-messages can be used to store or communicate the classical data encoded into q-states. We shall make an ultimate q-measurement on each q-code $|x\rangle$ in order to infer the encoded classical information to the best possible extent. A single pure state $|x\rangle$ can be unitarily transformed into $\log d$ qubits, and the above q-message can be transformed unitarily into a binary q-message of $n \log d$ qubits. The unit of q-information measure is one *random* qubit, or one qubit in short (meaning always one random qubit). The

information of a random n -qubit q-message is thus n qubits. We can decide if a q-message is random by constructing the density matrix of the 1-letter q-message:

$$\hat{\rho} = \sum_{\mathbf{x}} \rho(\mathbf{x}) |\mathbf{x}\rangle \langle \mathbf{x}| . \quad (10.10)$$

The q-message is random if this density matrix is totally mixed: $\hat{\rho} = \hat{I}/d$. Then the n -letter q-message is unitarily equivalent with $n \log d$ random qubits. If the q-message is not random it can be compressed faithfully into less than $n \log d$ random qubits. The original q-message is redundant and the measure of q-information is less than $n \log d$ qubits. In general, we define the information of a q-message by the length of the corresponding shortest faithfully compressed binary q-message. According to Schumacher's theory¹, this shortest length is $nS(\hat{\rho})$ in the limit $n \rightarrow \infty$. We outline the proof which resembles the proof of the classical case in Sect. 9.3.

10.3 Data compression

From the viewpoint of q-theory, the q-message $|\mathbf{x}_1\rangle, |\mathbf{x}_2\rangle, \dots, |\mathbf{x}_n\rangle$ is a multiple composite pure state in a d^n -dimensional Hilbert space:

$$|\Psi\rangle = |\mathbf{x}_1\rangle \otimes |\mathbf{x}_2\rangle \otimes \dots \otimes |\mathbf{x}_n\rangle . \quad (10.11)$$

This state can be realized by n independent d -state q-systems. The average of this state over all classical n -letter messages $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ will be the following collective (4.52) mixed state:

$$\sum_{\{\mathbf{x}\}} \rho(\mathbf{x}_1) \dots \rho(\mathbf{x}_n) |\Psi\rangle \langle \Psi| = \hat{\rho} \otimes \hat{\rho} \otimes \dots \otimes \hat{\rho} \equiv \hat{\rho}^{\otimes n} . \quad (10.12)$$

If $\hat{\rho} \neq \hat{I}/d$, the q-message $|\Psi\rangle$ is not random; it is redundant and the measure of q-information is less than $n \log d$ qubits. Indeed, many directions $|\Psi\rangle$ in the Hilbert space are totally unlikely to occur and become negligible if $n \rightarrow \infty$. Only a subspace of the whole d^n -dimensional Hilbert space will dominate statistically. This subspace is called the *typical subspace*. We shall represent it by a hermitian projector $\hat{\Omega}_n$ which will be determined for asymptotically large n . It can be shown that the maximum faithful compression of the pure state messages $|\Psi\rangle$ corresponds to the maximum faithful compression of the *average* message $\hat{\rho}^{\otimes n}$. The task reduces to the maximum faithful compression of the collective state $\hat{\rho}^{\otimes n}$.

Of course, classical compression of the initial messages $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ may not lead to an optimum compression of the encoded q-messages. Yet, as we shall argue, this becomes true provided the q-codes $|\mathbf{x}\rangle$ are orthogonal. We shall temporarily assume such orthogonal codes and utilize Shannon's classical data compression theorem. Note then that $S(\hat{\rho}) = S(\rho)$. Furthermore, we can perfectly discriminate the

¹ Cf. [32] by Schumacher.

different q-messages (10.11) because they are orthogonal to each other so that we can simply use the corresponding projective q-measurement. Consequently, the best q-compression turns out to be equivalent with the best classical data compression learnt in Sect. 9.3. We only retain those q-messages (10.11) which correspond to the typical classical messages. We learned that their number is $2^{nS(\rho)}$. The corresponding composite vectors $|\Psi\rangle$ span a $2^{nS(\rho)} = 2^{nS(\hat{\rho})}$ -dimensional typical subspace. Its orthogonal complement in the d^n -dimensional Hilbert space is called the *atypical subspace*. Since the probability of each classical typical message is just $2^{-nS(\hat{\rho})}$, the faithful compression of the average message yields the following new density matrix:

$$\hat{\rho}^{\otimes n} \longrightarrow \frac{1}{2^{nS(\hat{\rho})}} \hat{\Omega}_n, \quad (10.13)$$

where $\hat{\Omega}_n$ is the hermitian projector onto the $2^{nS(\hat{\rho})}$ -dimensional typical subspace. Q-data compression means that the original redundant q-messages (10.11) will be orthogonally projected by $\hat{\Omega}_n$ into the typical subspace. The typical subspace is unitary equivalent with $nS(\hat{\rho})$ qubits. The compressed state corresponds to $nS(\hat{\rho})$ -qubit random q-messages. These binary q-messages are not redundant, can not be further compressed faithfully. We have in such a way obtained the shortest faithful binary compression. Each q-letter $|x\rangle$ of the original q-message has been faithfully compressed from $\log d$ qubits into $S(\hat{\rho})$ qubits.

It is time to come back to the original task (10.9) and allow for non-orthogonal q-codes $|x\rangle$ as well. It can be shown that the desired faithful q-data compression depends on the 1-letter density matrix $\hat{\rho}$ only. Therefore the faithful q-data compression takes always the form (10.13). The projector $\hat{\Omega}_n$ will be calculated as before: for the given $\hat{\rho}$, there can always be constructed a hypothetical q-message (10.9) with orthogonal q-codes. Construction of $\hat{\Omega}_n$ is thus straightforward if we know $\hat{\rho}$. The q-data compression will be realized by the projection $\hat{\Omega}_n |\Psi\rangle$ into the typical subspace of $\hat{\rho}^{\otimes n}$. The compressed q-code becomes a $2^{nS(\hat{\rho})}$ -dimensional statevector and these statevectors occur at random within the typical subspace. Such statevectors are equivalently described by $nS(\hat{\rho})$ qubits which are also random. This way we have outlined the proof of Schumacher's q-data compression theorem: the classical message $x_1, x_2, \dots$, when encoded into the non-orthogonal pure states $|x_1\rangle, |x_2\rangle, \dots$, carries $S(\hat{\rho})$ qubit q-information per letter for asymptotically long messages, where $\hat{\rho}$ is the density matrix of the single letter q-message and $S(\hat{\rho})$ is its von Neumann entropy.

Von Neumann entropy $S(\hat{\rho})$ has thus been interpreted as the qubit-length per letter of the shortest faithful q-code of the message whose 1-letter average density matrix is $\hat{\rho}$. The above heuristic proof could be made exact. First, one has to construct the q-block-code of n -letter long messages in such a way that, in the asymptotic limit $n \rightarrow \infty$ the q-code length is $\sim nS(\hat{\rho})$ qubits. Note that the q-block-code can not be decomposed into letter-wise codes. It must be a multiply entangled pure state of all $\sim nS(\hat{\rho})$ qubits. Second, one has to prove that the compression is faithful, i.e., it does not decrease the accessible information 10.4. Third, one has to prove that further data compression would necessarily violate faithfulness. Schumacher has done the exact proof, that is his noiseless q-coding theorem.

10.4 Accessible q-information

In the previous section, certain classical messages $X = \{x, \rho(x)\}$ have been letter-wise encoded into pure q-states. More generally, we can encode letters into mixed q-states, too:

$$\begin{array}{llll} \text{letter:} & x_1 & x_2 & \dots & x_n \\ \text{q-code:} & \hat{\rho}_{x_1} & \hat{\rho}_{x_2} & \dots & \hat{\rho}_{x_n} \\ \text{probability:} & \rho(x_1) & \rho(x_2) & \dots & \rho(x_n) \end{array} \quad (10.14)$$

The density matrix of the averaged 1-letter code reads:

$$\hat{\rho} = \sum_x \rho(x) \hat{\rho}_x . \quad (10.15)$$

So far we have not discussed how much information is accessible regarding the original classical message X , by measuring the q-code. Now we apply a non-projective q-measurement 4.4.2 to each q-code $\hat{\rho}_x$. Given the set of effects $\{\hat{\Pi}_y\}$, the measurement outcome y will be considered a letter of another classical message Y whereas the probabilities of letters y are conditioned on the original letter x :

$$\rho(y|x) = \text{tr} \left(\hat{\Pi}_y \hat{\rho}_x \right) . \quad (10.16)$$

Formally, this is the transfer function of a noisy classical channel. From the output message Y we can letter-wise estimate the input message X . For statistics of long messages, the faithful average information carried by one letter y is called the *information gain* and measured by the mutual information 9.4:

$$I_{\text{gain}} = I(X : Y) = H(X) - H(X|Y). \quad (10.17)$$

For the noisy classical channel itself, this was considered the communicated amount of information and we searched its maximum by optimizing the input message, recall Sect. 9.5. This time, however, we take both the input message and its q-code granted and search the maximum of $I(X : Y)$ by optimizing the q-effects $\{\hat{\Pi}_y\}$ of our non-projective q-measurement:

$$A \equiv \max_{\{\hat{\Pi}_y\}} I(X : Y) . \quad (10.18)$$

This A is called the accessible information referred to the classical message X , from its q-codes (10.14). In the marginal case, the q-codes are orthogonal pure states. Then the optimum q-effects are orthogonal projectors and their projective q-measurement yields complete information on the original message: $A = H(X)$. The general case is much more difficult.

10.5 Entanglement: the resource of q-communication

Q-communication between Alice and Bob needs a q-channel typically. In the lack of q-channel, Alice and Bob can cope with teleportation 7.3.2 which needs shared

entangled states. One shared Bell state allows teleportation of one qubit, n Bell states allow teleportation of n qubits. We say: entanglement is the resource of q-communication. And vica versa, q-communication is a resource of shared entanglement. Alice and Bob can share n Bell states if they can communicate n qubits through their q-channel.

Accordingly, in the lack of q-communication the amount of (shared) entanglement is constant or at least not increasing, cf. Sect. 8.5. Alice and Bob can transform their entanglement into various forms but the amount of entanglement remains the same all the time. In q-information theory we consider the so-called LOCC situation where Alice and Bob can do any local operation on their respective local systems, and they can also communicate classically: but they can not make q-communication. An important field of q-information theory is studying various q-state manipulations at LOCC constraints.

The prototype of the entangled state is a maximally entangled Bell state:

$$|\Phi_{AB}^+\rangle = \frac{|0; A\rangle \otimes |0; B\rangle + |1; A\rangle \otimes |1; B\rangle}{\sqrt{2}} \equiv \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \quad (10.19)$$

This is the unit of entanglement measure, too. For it: $E = 1$. We say that one qubit is Alice's local qubit, the other qubit is Bob's. When they share k such Bell states we can say that they share entanglement $E = k$. If they share k pairs of qubits that are only partially entangled, e.g.:

$$|\Psi_{AB}\rangle = \cos \theta |00\rangle + \sin \theta |11\rangle, \quad (10.20)$$

then the entanglement per pair is smaller than $E = 1$, Alice and Bob share less entanglement than $E = k$. We made a reasonable choice earlier as to the measure of the partial entanglement. Accordingly, we calculate the reduced states of Alice and/or Bob:

$$\hat{\rho}_A = \hat{\rho}_B = \cos^2 \theta |0\rangle \langle 0| + \sin^2 \theta |1\rangle \langle 1|. \quad (10.21)$$

The partially entangled state (10.20) will, by our heuristic definition (7.13), possess an entanglement measured by the von Neumann entropy of the above reduced states:

$$E = S = -(\cos^2 \theta) \log(\cos^2 \theta) - (\sin^2 \theta) \log(\sin^2 \theta). \quad (10.22)$$

This heuristic definition fits to the minimum ($E = 0$) and maximum ($E = 1$) entanglements. The interpretation of the intermediate values is only possible if we learn the reversible cycle of entanglement distillation—dilution. We are going to prove in the next Section that, at LOCC constraints, a large number n of *partially* entangled states $|\Psi_{AB}\rangle$ can be *distilled* into $k = Sn$ *maximally* entangled states $|\Phi_{AB}^+\rangle$. This LOCC operation is reversible: a large number k of *maximally* entangled states $|\Phi_{AB}^+\rangle$ can be *diluted* into k/S *partially* entangled states $|\Psi_{AB}\rangle$, as we see in the Section after the next. Then it is crucial to understand that neither the efficiency of distillation nor the efficiency of dilution can be further improved. Why? Because we would then multiply the number of Bell states in a single distillation—dilution or dilution—distillation cycle. And this is impossible by LOCC, cf. Sect. 8.5.

In summary: the entanglement E of a composite pure state is interpreted by the rate of the maximum number of distillable Bell states, and this, as we are going to prove, is indeed equivalent to the heuristically found definition $E = S$.

10.6 Entanglement concentration (distillation)

Suppose Alice and Bob share n copies of partially entangled state $|\Psi_{AB}\rangle$ with entanglement $E = S$; let n be large. We show that, using LOCC only, Alice and Bob can distill their states into $k \sim nE = nS$ maximally entangled Bell states:

$$|\Psi_{AB}\rangle^{\otimes n} \longrightarrow |\Phi_{AB}^+\rangle^{\otimes k}. \quad (10.23)$$

To this end, Alice performs a smart collective q-measurement on her n qubits. The measured physical quantity will be the overall polarization in direction z . In the computational representation, this is equivalent with the collective measurement of the sum $\hat{m}$ of the binary physical quantities $\hat{x}$ (5.2) for Alice's n qubits, where:

$$\begin{aligned} \hat{m} &= \hat{x} \otimes \hat{I}^{\otimes(n-1)} + \text{permutations} \\ &\equiv \hat{x} \otimes \hat{I}^{\otimes(n-1)} + \hat{I} \otimes \hat{x} \otimes \hat{I}^{\otimes(n-2)} + \dots + \hat{I}^{\otimes(n-1)} \otimes \hat{x}. \end{aligned} \quad (10.24)$$

If the outcome is a given integer m then the rules of projective q-measurement impose the following state change:

$$|\Psi_{AB}\rangle^{\otimes n} \longrightarrow \frac{|11\rangle^{\otimes m} \otimes |00\rangle^{\otimes(n-m)} + \text{permutations}}{\sqrt{\binom{n}{m}}}. \quad (10.25)$$

where the permutations include all states where m qubits are $x = 1$ and the other $n - m$ qubits are $x = 0$. The most probable value of measurement outcome m is approximately this:

$$n \sin^2 \theta, \quad (10.26)$$

and by that shall we approximate Alice's measurement outcome. If $n \rightarrow \infty$, the Stirling formula (9.10) applies to the average number d of permutations on the r.h.s. of (10.25):

$$d = \binom{n}{m} \approx \binom{n}{n \sin^2 \theta} \approx 2^{nS} = 2^{nE}, \quad (10.27)$$

where S just turns out to be the von Neumann entropy (10.22). It was adopted as the heuristic entanglement E of the state $|\Psi_{AB}\rangle$. Hence Alice's collective measurement (10.25) has led approximately to the following entangled q-state:

$$\frac{|11\rangle^{\otimes m} \otimes |00\rangle^{\otimes(n-m)} + \text{permutations}}{\sqrt{d}}, \quad (10.28)$$

where the number d of the mutually orthogonal terms also yields the dimension of the subspace spanned by themselves. The above expression is just the (7.3) Schmidt

representation of the post-measurement state. The coefficients of the d mutually orthogonal tensor product states are all equal to $1/\sqrt{d}$ and this state is therefore maximally entangled (7.11) on the d -dimensional subspace. Suppose $\log d$ is integer, it can be shown this is of no ultimate restriction. Then, according to Sect. 7.1.6, the state (10.28) is equivalent with

$$k = \log d = nS \quad (10.29)$$

uncorrelated maximally entangled Bell states $|\Phi_{AB}^+\rangle$. This way we have shown that n partially entangled qubit pairs each of entanglement E can, via LOCC, be distilled into at least $k = nE$ maximally entangled pairs.

10.7 Entanglement dilution

Suppose Alice and Bob share k examples of the maximally entangled Bell states $|\Phi_{AB}^+\rangle$; let k be large. We show that, using LOCC only, Alice and Bob can dilute their states into $n \sim k/E = k/S$ partially entangled states $|\Psi_{AB}\rangle$ of entanglement $E = S$:

$$|\Phi_{AB}^+\rangle^{\otimes k} \longrightarrow |\Psi_{AB}\rangle^{\otimes n}. \quad (10.30)$$

To this end, Alice and Bob will combine q-data compression and teleportation.

In addition to her k qubits, Alice prepares $n = k/S = k/E$ local examples of the desired state $|\Psi_{AB}\rangle$ of partial entanglement $E = S$:

$$\left(\cos \theta |0; A'\rangle \otimes |0; A''\rangle + \sin \theta |1; A'\rangle \otimes |1; A''\rangle \right)^{\otimes n} \equiv |\Psi_{A'A''}\rangle^{\otimes n}. \quad (10.31)$$

In addition to his k qubits, also Bob prepares $n = k/S = k/E$ local raw qubits in any states, say in a certain $\hat{\rho}_{B'}^{\otimes n}$. If Alice can, using LOCC only, transmit each qubits from her systems A'' to Bob's systems B' then Alice and Bob would share n example of the desired state $|\Psi_{A'B'}\rangle$ of partial entanglement $E = S$. Indeed, Alice can initiate teleportation which is typical LOCC operation. But Alice and Bob can teleport k qubits only since they share just k Bell pairs $|\Phi_{AB}^+\rangle^{\otimes k}$ initially. Fortunately, q-data compression will map the n qubits into k .

Consider the collective state (10.31) and calculate the reduced collective state of the n qubits on the systems A'' :

$$\left(\cos^2 \theta |0; A''\rangle \langle 0; A''| + \sin^2 \theta |1; A''\rangle \langle 1; A''| \right)^{\otimes n} \equiv \hat{\rho}_{A''}^{\otimes n}. \quad (10.32)$$

This is the state Alice must teleport to Bob. Observe that $\hat{\rho}_A = \hat{\rho}_B = \hat{\rho}_{A'} = \hat{\rho}_{A''}$, cf. (10.21), therefore the corresponding von Neumann entropies S coincide. Alice must compress $\hat{\rho}_{A''}^{\otimes n}$. We learned in Sect. 10.3 that for asymptotically large n the collective state $\hat{\rho}_{A''}^{\otimes n}$ of n qubits can faithfully be compressed into the $2^{nS} = 2^k$ dimensional subspace $\hat{\Omega}_n$ of just k qubits. Therefore Alice teleports these k qubits (the compressed q-codes) into the subspace $\hat{\Omega}_n$ of k qubits of Bob's raw state $\hat{\rho}_{B'}^{\otimes n}$. For $n \rightarrow \infty$, the resulting state of the non-local composite systems $A'B'$ becomes the $n = k/S = k/E$ partially entangled states $|\Psi_{A'B'}\rangle$, identical to the desired ones $|\Psi_{AB}\rangle$.

10.8 Entropically irreversible operations

We introduced the qualitative notion of q-irreversibility on examples, e.g., of q-measurements in Sect. 4.4 and reduced dynamics in Sect. 4.5, where it meant just the non-invertibility of the corresponding q-operations. The informational concept of q-irreversibility is more restrictive than the concept of non-invertibility. It assumes q-information loss or, in alternative terminology: q-entropy increase. In non-selective q-measurements the von Neumann entropy is always increased or preserved. In typical reduced dynamics, however, the q-entropy for some initial q-states is increasing while for some other ones it is decreasing. Below, we are going to characterize the general class of entropy increasing q-operations.

A non-invertible (non-unitary) q-operation $\mathcal{M}$ is (entropically) irreversible if it increases or preserves the von Neumann entropy:

$$S(\mathcal{M}\hat{\rho}) \geq S(\hat{\rho}) , \quad (10.33)$$

for all q-states $\hat{\rho}$. We shall show that a non-invertible $\mathcal{M}$ is irreversible if it leaves the totally mixed q-state $\hat{\rho}_0 = \hat{I}/d$ invariant: $\mathcal{M}\hat{\rho}_0 = \hat{\rho}_0$.

The q-entropy of a state $\hat{\rho}$ is always equal to the Shannon entropy computed from the eigenvalues $\{w_\lambda\}$, provided we take the degenerate eigenvalues twice or more, accordingly. Similarly, the q-entropy of $\hat{\rho}' \equiv \mathcal{M}\hat{\rho}$ coincides with the Shannon entropy computed from the eigenvalues $\{w'_\mu\}$ of $\hat{\rho}'$. Therefore it is sufficient if we prove the classical relationship $S(w') \geq S(w)$ where w, w' are the diagonal (eigenvalue) distributions of the density matrices $\hat{\rho}, \hat{\rho}'$, respectively. Let us introduce the spectral expansions $\hat{\rho} = \sum_\lambda w_\lambda \hat{P}_\lambda$ and $\hat{\rho}' = \sum_\mu w'_\mu \hat{P}'_\mu$. Then we express the eigenvalues of $\hat{\rho}'$ through the following equivalent expressions:

$$w'_\mu = \text{tr} \left(\hat{P}'_\mu \hat{\rho}' \right) = \text{tr} \left(\hat{P}'_\mu \mathcal{M}\hat{\rho} \right) = \sum_\lambda \text{tr} \left(\hat{P}'_\mu \mathcal{M}\hat{P}_\lambda \right) w_\lambda , \quad (10.34)$$

where the last form has been obtained by substituting the spectral expansion of $\hat{\rho}$. For fixed w, w' , the above map looks like a classical operation on w , yielding w' through a transfer function. A transfer function is always normalized in the first variable μ . Our transfer function is normalized in the second variable as well:

$$\sum_\lambda \text{tr} \left(\hat{P}'_\mu \mathcal{M}\hat{P}_\lambda \right) = \text{tr} \left(\hat{P}'_\mu \mathcal{M}\hat{I} \right) = 1 , \quad (10.35)$$

since $\mathcal{M}\hat{I} = \hat{I}$ according to our initial postulation. Our transfer function is double stochastic, therefore the Shannon entropy of w will be greater or equal to the Shannon entropy of w' , cf. Sect. 9.8. In general, the strict inequality holds for $\hat{\rho} \neq \rho_0$ unless $\hat{\rho}$ is singular and the restriction of the non-unitary $\mathcal{M}$ for the support of $\hat{\rho}$ becomes invertible.

The simplest example of (entropically) irreversible q-operations is the non-selective q-measurement. It is non-invertible and it preserve the totally mixed state. As a contrary example, consider a “broken” q-channel whose output message is

a permanent $|0\rangle$ for any input whatsoever: $\hat{\rho}' = \hat{P}_0 \hat{\rho} \hat{P}_0 + \sigma_x \hat{P}_1 \hat{\rho} \hat{P}_1 \sigma_x$ where $\hat{P}_x = |x\rangle \langle x|$, for $x = 0, 1$. The q-channel operation is obviously non-invertible but it is not entropically irreversible.

Problems, exercises

10.1. Subadditivity of q-entropy. Using the Klein inequality, let us prove the subadditivity of the von Neumann entropy.

10.2. Concavity of q-entropy, Holevo entropy. When we prepare a q-state as a mixture $\hat{\rho} = \sum_n w_n \hat{\rho}_n$ then the entropy of the mixture is never smaller than the average entropy of the components:

$$0 \leq S(\hat{\rho}) - \sum_n w_n S(\hat{\rho}_n) .$$

The r.h.s. is called Holevo entropy. It puts an upper bound on the accessible information and it is bounded from above by the so-called entropy of preparation $S(w)$:

$$A \leq S(\hat{\rho}) - \sum_n w_n S(\hat{\rho}_n) \leq S(w) .$$

The first inequality is called the Holevo bound² and it is a most powerful theorem of q-information theory.

Holevo entropy is non-negative because of the concavity of the q-entropy. Let us prove concavity from subadditivity.

10.3. Data compression of the non-orthogonal code. Suppose random bits are encoded into non-orthogonal pure states exactly like in the case of the q-banknote 6.4.1 as well as of q-cryptography 6.4.2. Thus we consider the states $|\uparrow z\rangle, |\uparrow x\rangle$ the q-code of the random classical binary message $X = \{x, \rho(x)\}$:

x letter:	0	1	
q-code:	$ \uparrow z\rangle$	$ \uparrow x\rangle$	
$\rho(x)$ probability:	1/2	1/2	(10.36)

The Shannon entropy of the random classical letter is $H(X) = 1$ bit. Let us determine the maximum faithful q-data compression rate.

10.4. Distinguishing two non-orthogonal qubits: various aspects. We have previously mentioned two strategies 6.3 to distinguish non-orthogonal random states like $|\uparrow z\rangle, |\uparrow x\rangle$. The first strategy 6.3.1 used a single projective measurement of either $\hat{\sigma}_z$ or $\hat{\sigma}_x$ chosen at random. It was perfectly conclusive at rate 25%. The second strategy 6.3.2 used a single non-projective measurement regarding three suitable q-effects

² Cf. [33] by Holevo.

$\hat{\Pi}_z, \hat{\Pi}_x, \hat{\Pi}_?$, and this strategy was perfectly conclusive at rate $1 - 1/\sqrt{2} \approx 30\%$ and perfectly inconclusive in the rest. In the context of Prob. 10.3, the information gain I_{gain} will coincide with the rate 30% of the unambiguous decisions. Yet, there are different projective q-measurements which are never perfectly conclusive but their information gain is bigger. Let us calculate I_{gain} for the single polarization measurement in the direction $(1, 0, -1)$.

10.5. Simple optimum q-code. Assume we have to encode the three colors R,G, and B, occurring at random, into the respective pure states $|R\rangle, |G\rangle, |B\rangle$ of a qubit. Suppose we do not care if the q-code would not guarantee the highest accessible information. Suppose we have limited q-channel capacities therefore we want optimum q-code in a sense that it can not be shortened further, i.e., without the loss of the accessible information. Then we use our q-channel economically. Let us construct the corresponding three pure states $|R\rangle, |G\rangle$, and $|B\rangle$.

11 Q-computation

Classical digital computer performs its algorithm via one- and two-bit operations realized by one- and two-bit “gates”. The (yet hypothetic) q-digital computer would do the same. The classical logical gates become q-gates, their “circuits” are called q-circuits. They can perform what is called q-parallel computation. Q-algorithms are always reversible. Therefore we learn classical reversible computation first then we discuss the earliest q-algorithms that might overcome all classical algorithms targeting a similar task¹.

11.1 Parallel q-computing

A classical computer works like this: the input *storage* stores an n -digit binary number x upon which the *algorithm* is performed and the result will be read out from the same storage. Now, let the computer be a q-system, operating coherently on its q-storage which is an n -qubit statevector:

$$|x_1\rangle \otimes \dots \otimes |x_n\rangle \equiv |x\rangle ; \quad x \equiv x_n \dots x_2 x_1 , \quad (11.1)$$

Unprecedented in classical computation, we can store all possible $N = 2^n$ initial values in the q-storage if we *superpose* them, say, with the same amplitudes:

$$|S\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle . \quad (11.2)$$

We call this state the totally superposed state of the storage. What about a coherent q-algorithm? It is a unitary $N \times N$ matrix $\hat{U}$. Given a classical algorithm, we would naively think that its q-counterpart corresponds to a certain matrix $\hat{U}$. This is the case, indeed, provided the classical algorithm is reversible. If it is not then we have to replace it by a reversible version. Fortunately, it had been known in classical theory of computation that any computational task can be realized reversibly if some redundant data are drawn along, cf. Sect. 11.2. This way we can realize any classical computational task via the corresponding unitary transformation $\hat{U}$ on a q-storage. Parallel computation becomes possible on all possible initial data captured

¹ For a short review on q-computation, see [12].

by the totally superposed state $|S\rangle$. Moreover, we could perform various different reversible algorithms running simultaneously on the same q-computer.

Contrary to early expectations², it is not so easy to find a concrete computational task whose coherent q-algorithm outperforms all possible incoherent classical algorithms. The total superposition of the initial states $|S\rangle$ is instrumental but, surprisingly, it does not in itself guarantee an advantage. We need to invent genuine q-algorithms, like those of Sects. 11.3, 11.4, differing radically from the classical ones.

11.2 Evaluation of arithmetic functions

We are going to discuss a basic classical computational task. The algorithm should compute functions $y = f(x)$ that map n -bit integers x onto m -bit integers y :

$$x \longrightarrow \boxed{f} \longrightarrow f(x) \quad (11.3)$$

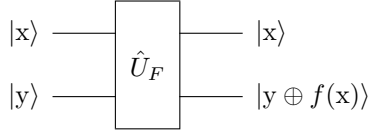
The map f is not necessarily invertible, therefore the corresponding algorithm is not necessarily reversible. For the purpose of q-parallel computation, we extend the map f in such a way that the new map F be reversible even if f were not. Accordingly, we set up independent input and output storages, i.e., the output storage exists already at input time. Let the new invertible map F be the following:

$$\begin{array}{ccc} x & \longrightarrow & \boxed{F} & \longrightarrow & x \\ y & \longrightarrow & & & y \oplus f(x) \end{array} \quad (11.4)$$

The summation is understood mod M where $m = \log M$ is the bit-length of the output storage. How does this algorithm calculate $f(x)$? That's simple. We set the output storage at initial time to zero, $y = 0$, so that the same storage at final time will contain $f(x)$.

The new function $F(x, y)$ is always invertible. Therefore the computation of the arithmetic function $f(x)$ has been realized by a reversible classical algorithm which can directly be converted into the following q-algorithm:

² Cf. [34] by Feynman.



(11.5)

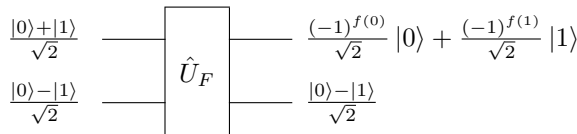
Instead of this diagrammatic definition, the $2^{n+m} \times 2^{n+m}$ unitary matrix of the q-algorithm can be defined in the usual way as well:

$$\hat{U}_F(|x\rangle \otimes |y\rangle) = |x\rangle \otimes |y \oplus f(x)\rangle. \quad (11.6)$$

11.3 Oracle problem: the first q-algorithm

Consider a black-box (11.3) to evaluate a certain arithmetic function $f(x)$, and call it the (classical) oracle. If we ask the oracle a “question” x , it provides the immediate “answer” $y = f(x)$. This time it is irrelevant how it evaluates the function. Suppose the point is that we do not know the function $f(x)$, only the oracle knows it. The hypothetic task will be this: we want to learn the function $f(x)$ therefore we are going to ask questions and evaluating the answers. How many questions should we ask?

It is this class of problems where the first ever task has been invented³ for which a smart q-algorithm is clearly more efficient than any classical algorithm. In the simplest case, let f map a single bit x into a single bit y . Suppose we have to learn whether $f(x)$ is a constant function or it is not. If $f(x)$ is evaluated by a classical oracle then we must ask twice: we submit both initial values $x = 0$ and $x = 1$ in separate “questions”. More questions do not exist, less questions are not sufficient to decide whether or not $f(0) = f(1)$. If, however, the function f is evaluated by a q-oracle (11.5) then a single “question” will be sufficient:



(11.7)

The input storage should be set to the total superposition $|S\rangle$ of all possible states (i.e.: $x = 0, 1$) in accordance with the concept of q-parallel computation (11.2). We do not set the output storage to zero, rather we set it to the superposition of the two possible states $x = 0, 1$ — with the opposite phases. We show that to such a single “question” the q-oracle’s answer is decisive regarding the constancy of f . As we see from (11.7), the output storage never changes. The input storage, however, does:

³ Cf. [35] by Deutsch and [36] by Deutsch and Jozsa.

$$\begin{aligned} & \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \text{if } f(0) = f(1) , \\ & \frac{|0\rangle - |1\rangle}{\sqrt{2}} \quad \text{if } f(0) \neq f(1) . \end{aligned} \quad (11.8)$$

These two input storage states are orthogonal. A single ultimate projective q-measurement can discriminate them. After the measurement we know with certainty whether the function f is constant or not.

We have yet to prove that the q-oracle provides the “answer” (11.7) for the smart “question” there. Obviously we can re-write the initial state this way:

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \sum_{x=0,1} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} . \quad (11.9)$$

The unitary map $\hat{U}_F$ (11.6) applies to this, and yields the following state:

$$\frac{1}{\sqrt{2}} \sum_{x=0,1} |x\rangle \otimes \frac{|f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} . \quad (11.10)$$

We can write the state of the output storage into an equivalent form:

$$|f(x)\rangle - |1 \oplus f(x)\rangle = (-1)^{f(x)} (|0\rangle - |1\rangle) . \quad (11.11)$$

Indeed, the state of the output storage has, upto its phase, remained unchanged. Using the above form, the final composite state produced by the q-oracle becomes:

$$\frac{1}{\sqrt{2}} \sum_{x=0,1} (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} . \quad (11.12)$$

It coincides with the “answer” displayed on the diagram (11.7).

The advantage of q-tests over classical ones becomes more spectacular when the oracle f maps a large number of n bits into one. Suppose the arithmetic function f is either constant or balanced. This latter means that $f(x)$ returns 0 and 1 equal times when x runs over the $N = 2^n$ different values. Suppose that we are interested to learn whether the oracle is constant or balanced. Classically we must ask $N/2 + 1$ “questions”. In case of the corresponding q-oracle, a single “question” is enough. The construction is very similar to the former simplest one. We keep setting the input storage to the totally superposed state $|S\rangle$ and we keep setting the output storage to $(|0\rangle - |1\rangle)/\sqrt{2}$. We analyze the “answer” of the oracle by an ultimate projective q-measurement of $|S\rangle \langle S|$. We are going to prove that f is constant if the outcome is 1 and f is balanced if the outcome is 0. The proof consists of the same steps as before, with the only change that x runs from 0 to $N - 1$. Accordingly, the initial state reads:

$$|S\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} , \quad (11.13)$$

while the final state, i.e., the “answer” will be:

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (11.14)$$

Obviously, the final state of the input storage is $|S\rangle$ when f is constant and it is orthogonal to $|S\rangle$ when f is balanced.

11.4 Searching q-algorithm

Suppose that the function $f(x)$, mapping n bits into one bit, is always zero except for $x = x_0$:

$$f(x) = \delta_{xx_0}. \quad (11.15)$$

Let a classical oracle evaluate the function. Assume that x_0 is unknown to us and we have to learn it by asking the oracle. This is the basic searching problem formulated in the language of the oracle-problem. The unknown integer x_0 can take any values between 0 and $N - 1$ where $N = 2^n$. Obviously, we must ask $N - 1$ questions in the worst case, and $O(N)$ questions in general. Grover showed that we can ask $O(\sqrt{N})$ questions of the corresponding q-oracle, which is much less than $O(N)$. The corresponding q-algorithm⁴ will be much faster than any classical algorithm to solve the same searching problem.

Grover’s q-algorithm is iterative, of $R \sim \sqrt{N}$ iterations. In each cycle, we ask the q-oracle one question and perform a given unitary transformation on its answer. This state becomes the question in the forthcoming cycle, etc. The output storage of the oracle stays always in the state $(|0\rangle - |1\rangle)/\sqrt{2}$, hence the “question” $|x\rangle$ submitted into the input storage will produce this answer:

$$(-1)^{f(x)} |x\rangle = (1 - 2\delta_{xx_0}) |x\rangle. \quad (11.16)$$

This “question-to-answer” map can be described by the following $N \times N$ unitary matrix:

$$\hat{I} - 2 |x_0\rangle \langle x_0|. \quad (11.17)$$

Thus we use the q-oracle to mark for us the searched direction $|x_0\rangle$ in the state-space. Let our first “question” be the totally superposed state $|S\rangle$. The “answer” will be $(\hat{I} - 2 |x_0\rangle \langle x_0|) |S\rangle$ and, according to Grover, we transform it by another unitary matrix:

$$\hat{I} - 2 |S\rangle \langle S|. \quad (11.18)$$

The result will be utilized as the “question” asked of the q-oracle in the second iteration cycle, etc. We are going to prove that after R iterations the resulting state is approximately $|x_0\rangle$:

⁴ Cf. [37] by Grover.

$$\left[(\hat{I} - 2|S\rangle\langle S|)(\hat{I} - 2|x_0\rangle\langle x_0|) \right]^R |S\rangle \approx |x_0\rangle, \quad (11.19)$$

so that the projective q-measurement of $\hat{x}$ on the resulting state yields the searched value x_0 .

The proof is elementary. Why? The iteration (11.19) does nothing else than a rotation of the initial vector $|S\rangle$ toward the searched vector $|x_0\rangle$, by the same angle in each cycle. The process can be described in the real 2-dimensional plane spanned by $|S\rangle$ and $|x_0\rangle$. In this plane, rotations (11.17, 11.18) are considered modulo π since the sign of the state vector is irrelevant. For simplicity's sake, we assume $N \gg 1$. Then $|x_0\rangle$ is almost orthogonal to $|S\rangle$:

$$\langle x_0 | S \rangle = \frac{1}{\sqrt{N}}. \quad (11.20)$$

Their angle $(\pi/2) - \epsilon$ is almost the rectangle. The defect is very small:

$$\epsilon \approx \frac{1}{\sqrt{N}}. \quad (11.21)$$

In the first iteration, the matrix (11.17) reflects the initial state $|S\rangle$ w.r.t. the direction $|x_0\rangle$; the resulting vector will be reflected by the matrix (11.18) w.r.t. the direction of $|S\rangle$. The outcome vector is such that it encloses a smaller angle $(\pi/2) - 3\epsilon$ with the searched vector $|x_0\rangle$. It is easy to see that the two reflections in each further cycle of Grover iteration will always rotate toward $|x_0\rangle$ by the same angle 2ϵ . Since the initial vector $|S\rangle$ was away from $|x_0\rangle$ by an angle $(\pi/2) - \epsilon$, the necessary number of iterations will be:

$$R \approx \frac{\pi}{4\epsilon} \approx \frac{\pi}{4} \sqrt{N}. \quad (11.22)$$

The R -times iterated state approaches $|x_0\rangle$ to within a small angle $\sim \epsilon$. Hence its projective measurement in the computational basis yields x_0 with an error probability as little as $\sim 1/N$.

11.5 Fourier algorithm

The well-known discrete classical Fourier transformation is equivalent with a unitary transformation on the given complex vector space. We can use the q-physics notations. If $\{|x\rangle\}$ stand for the initial basis vectors and $\{|x; F\rangle\}$ are their Fourier transforms, respectively, then:

$$|y; F\rangle = \frac{1}{\sqrt{N}} \sum_x e^{2\pi i y x / N} |x\rangle, \quad (11.23)$$

where x, y are binary numbers of $n = \log N$ bits. Since this classical map is reversible we can directly turn it into a q-algorithm. Yet, as we shall see, the entanglement of the computational basis on the r.h.s. is only apparent, the r.h.s. is a product state in fact. This factorization makes the classical fast-Fourier algorithm possible.

We are going to prove that, using the notation $y = y_n \dots y_2 y_1$, the Fourier transform is a product state of n different qubits. We substitute the digital expansion $x = \sum_{k=1}^n x_k 2^{k-1}$ in order to factorize the elements of the unitary matrix of the Fourier transform:

$$e^{2\pi i y x / N} = \prod_{k=1}^n \exp(2\pi i y x_k 2^{k-1} / N) = e^{2\pi i y_1 x_n / N} e^{2\pi i y_2 y_1 x_{n-1} / N} \dots e^{2\pi i y_n x_1 / N} . \quad (11.24)$$

We can also factorize the sum over the computational basis states:

$$\sum_x e^{2\pi i \dots} |x\rangle = \sum_{x_n=0}^1 \sum_{x_{n-1}=0}^1 \dots \sum_{x_1=0}^1 e^{2\pi i \dots} |x_n\rangle \otimes |x_{n-1}\rangle \otimes \dots \otimes |x_1\rangle . \quad (11.25)$$

Substituting these expressions into that of Fourier transform (11.23), we obtain the following:

$$|y; F\rangle = \frac{|0\rangle + e^{2\pi i y_1 / N} |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + e^{2\pi i y_2 y_1 / N} |1\rangle}{\sqrt{2}} \otimes \dots \otimes \frac{|0\rangle + e^{2\pi i y_n / N} |1\rangle}{\sqrt{2}} . \quad (11.26)$$

This shows that the q-algorithm of Fourier transform can be built up from single qubit operations controlled by the state of other qubits. Thus, in fact, these operations are multi-qubit operations. Although the number of control qubits is more than one, two-qubit controlled gates (cf. Sect. 11.6) are universally sufficient when organized in suitable q-circuits.

The q-Fourier algorithm becomes instrumental for the Shor algorithm⁵ to factorize large numbers, which is the most promising prediction of q-computation theory. The Shor q-algorithm is exponentially faster than all known classical algorithms to factorize large integers.

11.6 Q-gates, q-circuits

Any given classical algorithm can be decomposed into a series of Boolean operations. It is known that a small selection of 1– and 2–bit operations suffice although this selection is not unique. The selected ones are called universal logical operations (or gates, referring to their technical realization). To perform a given algorithm, the classical logical gates are organized into logical circuits. Similarly, we define q-logical gates and circuits to perform q-algorithms. We have already learned about the 1–qubit logical operations 6.1.1 and now we associate the notion of q-gate to them. We also mentioned that the Hadamard gate (H) and phase-gate (T), i.e.:

$$\hat{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad |x\rangle \rightarrow \boxed{H} \rightarrow \frac{(-1)^x}{\sqrt{2}} |x\rangle + \frac{1}{\sqrt{2}} |1-x\rangle \quad (11.27)$$

$$\hat{T}(\varphi) = \begin{bmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{bmatrix} \quad |x\rangle \rightarrow \boxed{\varphi} \rightarrow \exp[i(x - \tfrac{1}{2})\varphi] |x\rangle \quad (11.28)$$

⁵ Cf. [39] by Shor.

are universal 1-qubit gates: all 1-qubit unitary operation can be realized by them.

The 2-qubit operations (gates) can be derived from the so-called controlled operations (gates). Then the first qubit is the control-qubit and the second is the target-qubit. If the control-qubit is 1 then a given unitary operation $\hat{U}$ is performed on the target-qubit. If the control qubit is 0 then nothing will happen:

$$\begin{array}{c}
 |x\rangle \text{---} \bullet \text{---} |x\rangle \\
 |y\rangle \text{---} \boxed{\hat{U}} \text{---} \hat{U}^x |y\rangle
 \end{array} \quad (11.29)$$

This is, for instance, the controlled NOT (cNOT, XOR):

$$\hat{C} = \begin{bmatrix} \hat{I} & 0 \\ 0 & \hat{X} \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{array}{c} |x\rangle \text{---} \bullet \text{---} |x\rangle \\ |y\rangle \text{---} \oplus \text{---} |x \oplus y\rangle \end{array} \quad (11.30)$$

The 4×4 matrix corresponds to the order $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ in obvious notation $|xy\rangle = |x\rangle \otimes |y\rangle$. of the bipartite basis-vectors. The chosen symbol $\oplus$ on the diagram reminds us that in computational basis the cNOT sets the target-qubit to the modulo 2 sum of the two initial qubits.

It is known that the Hadamard-, the phase-, and the cNOT-gates are universal q-gates. These three q-gates allow constructing all unitary operations on a q-storage of any number n of qubits. Design of q-logical circuits, similarly to design of classical logical circuits, has its own standards and tricks⁶. Talking about q-algorithms, standard diagrams of q-circuits are used extensively. Moreover, the language and formalism of q-information theory is penetrating our understanding — and teaching — the traditional q-theory as well.

Problems, exercises

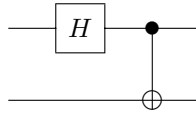
11.1. Creating the totally symmetric state. Let us prove that n independent qubits, in state $|0\rangle$ each, can be transformed into the totally symmetric state $|S\rangle$ by using n Hadamard gates.

11.2. Constructing Z-gate from X-gate. Let us confirm that the 1-qubit q-gates X and Z are related via Hadamard gate sandwiches: $HXH = Z$ and $HZH = X$.

11.3. Constructing controlled Z-gate from cNOT-gate. Adding 1-qubit gates to a cNOT-gate in a suitable way, let us build a q-circuit to realize a controlled-Z operation.

⁶ Cf. e.g. [38] by Tucci.

11.4. Q-circuit to produce Bell states. Let us show that the following q-circuit produces the four Bell states from the computational basis:



11.5. Q-circuit to measure Bell states. Let us construct the q-circuit of projective measurement in the Bell basis.

Solutions

Problems of Chap. 2

2.1 Mixture of pure states. It is easy to see that $\rho(x) = \rho(0)\delta_{x0} + \rho(1)\delta_{x1}$ which suggests that the distribution of the mixing weights must coincide with the distribution ρ . Accordingly, in the case of continuous phase space, we can choose $w(\bar{x}) = \rho(\bar{x})$ to mix the pure states $\delta(x - \bar{x})$:

$$\rho(x) = \int w(\bar{x})\delta(x - \bar{x})d\bar{x} .$$

2.2 Probabilistic or non-probabilistic mixing? Mixing n zeros and n ones is realized by their random permutation that amounts to the following $2n$ -partite composite state:

$$\rho(x_1, \dots, x_{2n}) = \frac{\delta_{x_1 0} \dots \delta_{x_n 0} \delta_{x_{n+1} 1} \dots \delta_{x_{2n} 1} + \text{permutations of } x_1, \dots, x_{2n}}{\# \text{ of permutations}} .$$

In case of the probabilistic mixing, however, we make $2n$ random decisions as to choose a zero or a one and then we mix the $2n$ elements. The probabilistic mixing amounts to $\rho(x_1) \dots \rho(x_{2n})$ which is obviously different from the above composite state.

2.3 Classical separability. If we choose $w(\bar{x}_A, \bar{x}_B) = \rho_{AB}(\bar{x}_A, \bar{x}_B)$ and replace summation over the weights by integration, it works:

$$\rho_{AB}(x_A, x_B) = \int w(\bar{x}_A, \bar{x}_B)\delta(x_A - \bar{x}_A)\delta(x_B - \bar{x}_B)d\bar{x}_A d\bar{x}_B .$$

Thus all ρ_{AB} are mixtures of the uncorrelated pure states $\delta(x_A - \bar{x}_A)\delta(x_B - \bar{x}_B)$.

2.4 Decorrelating a single state? The map $\rho_{AB} \rightarrow \rho_A \rho_B$ is nonlinear:

$$\rho_{AB}(x_A, x_B) \longrightarrow \int \rho_{AB}(x_A, x'_B)dx'_B \int \rho_{AB}(x'_A, x_B)dx'_A .$$

Hence the map is not a real operation.

2.5 Decorrelating an ensemble. The collective state $\rho_{AB}^{\times 2n}$ is granted to start with. We interchange the first n and the second n subsystems A. Then we trace over

the second n composite systems AB. We get $(\rho_A \rho_B)^{\times n}$. This can be verified if, e.g., we calculate the expectation values of observables like $f(x_{A_1})g(x_{B_1})$, and $f_1(x_{A_1})f_2(x_{A_2})g_1(x_{B_1})g_2(x_{B_2})$, etc.

2.6 Classical indirect measurement. Imagine a detector of discrete state space $\{n\}$. Let the composite state of the system and the detector be $\rho(x, n) = \rho(x) \Pi_n(x)$ where $\rho(x)$ stands for the reduced state of the system. Observe that $\Pi_n(x)$ becomes the conditional state $\rho(n|x)$ of the detector provided the system is in the pure state x . Now we perform a projective measurement on the detector quantity n . Formally, the partition $\{P_m\}$ of the detector state space must be defined as $P_m(n) = \delta_{mn}$ for all m . According to the rules of projective measurement, the state change will be:

$$\rho(x, n) \rightarrow \rho_m(x, n) \equiv \frac{1}{p_m} \delta_{mn} \rho(x, n) ,$$

with probability $p_m = \int \rho(x, m) dx$. For the reduced state $\rho(x) = \sum_m \rho(x, m)$, the above projective measurement induces the desired non-projective measurement of the effects $\{\Pi_n(x)\}$.

Problems of Chap. 3

3.1 Bohr quantization of the harmonic oscillator. The sum of the kinetic and potential energies yields the total energy $E = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$ which is constant during the motion. Therefore the phase space point (q, p) moves on the ellipse of surface $2\pi E/\omega$. The surface plays a role in the Bohr-Sommerfeld q-condition because the contour-integral of $p dq$ for one period is equal to the surface of the enclosed ellipse. Hence the canonical action takes the form E/ω and we get:

$$E/\omega = \hbar(n + \frac{1}{2}) .$$

3.2 The role of adiabatic invariants. The canonical actions I_k are adiabatic invariants of the classical motion. This means that they remain approximately constant against whatever large variations of the external parameters of the Hamilton-function provided the variations are slow with respect to the motion. So, the canonical action I of the oscillator will be invariant against the variation of ω in the Hamiltonian $\frac{1}{2}p^2 + \frac{1}{2}\omega^2(t)q^2$ provided $|\dot{\omega}| \ll \omega^2$. The q-condition remains satisfied with the same q-number n .

3.3 Classical-like or q-like motion. The Bohr-Sommerfeld q-condition restricts the continuum of classical motions to a discrete infinite sequence. For small q-numbers this restriction is relevant since the allowed phase space trajectories are well separated. For large q-numbers, typically, the allowed trajectories become quite dense in phase space and might fairly approximate any classical trajectory which does otherwise not satisfy the q-conditions.

Problems of Chap. 4

4.1 Decoherence-free projective measurement. Let us construct the spectral expansion $\hat{A} = \sum_{\lambda} A_{\lambda} \hat{P}_{\lambda}$ and the post-measurement state $\hat{\rho}' = \sum_{\lambda} \hat{P}_{\lambda} \hat{\rho} \hat{P}_{\lambda}$. If $[\hat{A}, \hat{\rho}] = 0$ then $\hat{\rho}' = \hat{\rho}$ since $[\hat{A}, \hat{\rho}] = 0$ is equivalent with $[\hat{P}_{\lambda}, \hat{\rho}] = 0$ for all λ . To prove the inverse statement, we consider the following identity:

$$[\hat{\rho}' - \hat{\rho}, \hat{P}_{\lambda}] = [\hat{P}_{\lambda}, \hat{\rho}] .$$

If $\hat{\rho}' = \hat{\rho}$ then the l.h.s. is zero for all λ which implies that the r.h.s. is zero for all λ which implies $[\hat{A}, \hat{\rho}] = 0$.

4.2 Mixing the eigenstates. Let us consider the spectral expansion of the matrix $\hat{\rho}$:

$$\hat{\rho} = \sum_{\lambda} \rho_{\lambda} \hat{P}_{\lambda} .$$

If $\hat{\rho}$ is non-degenerate then the $\hat{P}_{\lambda}$'s correspond to the pure eigenstates of $\hat{\rho}$ and their mixture yields the state $\hat{\rho}$ if the corresponding eigenvalues make the mixing weights: $w_{\lambda} = \rho_{\lambda}$. In the general case, the spectral expansion implies the mixture $\hat{\rho} = \sum_{\lambda} w_{\lambda} \hat{\rho}_{\lambda}$ with $w_{\lambda} = d_{\lambda} \rho_{\lambda}$ and $\hat{\rho}_{\lambda} = \hat{P}_{\lambda} / d_{\lambda}$ where d_{λ} is the dimension of $\hat{P}_{\lambda}$.

4.3 Separability of pure states. If $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ then the composite density matrix $\hat{\rho}_{AB}$ is a single tensor product and it is trivially separable. The other way around, when the pure state satisfies the separability condition (4.47):

$$|\psi_{AB}\rangle \langle \psi_{AB}| = \sum_{\lambda} w_{\lambda} \hat{\rho}_{A\lambda} \otimes \hat{\rho}_{B\lambda} ,$$

then it follows that the matrices on both sides have rank 1. Accordingly, the r.h.s. must be equivalent to the tensor product of rank-one (i.e.: pure state) density matrices:

$$|\psi_{AB}\rangle \langle \psi_{AB}| = |\psi_A\rangle \langle \psi_A| \otimes |\psi_B\rangle \langle \psi_B| ,$$

which implies the form $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$.

4.4 Unitary cloning? Let us suppose that we have duplicated two states $|\psi\rangle$ and $|\psi'\rangle$:

$$|\psi\rangle \otimes |\psi_0\rangle \longrightarrow |\psi\rangle \otimes |\psi\rangle ; \quad |\psi'\rangle \otimes |\psi_0\rangle \longrightarrow |\psi'\rangle \otimes |\psi'\rangle .$$

The inner product of the two initial composite states is $\langle \psi | \psi' \rangle$ while the inner product of the two final composite states is $\langle \psi | \psi' \rangle^2$. Therefore the above process of state duplication can not be unitary.

Problems of Chap. 5

5.1 Pure state fidelity from density matrices. Observe that $\langle \mathbf{m} | \mathbf{n} \rangle^2$ equals the trace of the product $|\mathbf{n}\rangle \langle \mathbf{n}|$ times $|\mathbf{m}\rangle \langle \mathbf{m}|$. Let us invoke the Pauli-representation of these two density matrices and evaluate the trace of their product:

$$|\langle m|n\rangle|^2 = \text{tr} \left(\frac{\hat{I} + n\hat{\sigma}}{2} \frac{\hat{I} + m\hat{\sigma}}{2} \right) = \frac{1 + nm}{2},$$

which yields $\cos^2(\vartheta/2)$.

5.2 Unitary rotation for $|\uparrow\rangle \longrightarrow |\downarrow\rangle$. Since $|\uparrow\rangle$ corresponds to the north pole and $|\downarrow\rangle$ corresponds to the south pole on the Bloch-sphere, we need a π -rotation around, e.g., the x -axis. The rotation vector is $\alpha = (\pi, 0, 0)$ and the corresponding unitary transformation becomes:

$$\hat{U}(\alpha) \equiv \exp \left(-\frac{i}{2} \alpha \hat{\sigma} \right) = -i \hat{\sigma}_x.$$

We can check the result directly:

$$-i \hat{\sigma}_x |\uparrow\rangle = -i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -i \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -i |\downarrow\rangle.$$

5.3 Density matrix eigenvalues and -states in terms of polarization. Consider the density matrix $\frac{1}{2}(\hat{I} + s\hat{\sigma})$ and find the spectral expansion of $s\hat{\sigma}$. We learned that if s is a unit vector then $s\hat{\sigma} |\uparrow s\rangle = |\uparrow s\rangle$ and $s\hat{\sigma} |\downarrow s\rangle = -|\downarrow s\rangle$. If $s \leq 1$, the two eigenstates remain the same and we keep the simple notations $|\uparrow s\rangle, |\downarrow s\rangle$ to denote qubits polarized along or, respectively, opposite to the direction s . The eigenvalues will change trivially and we have $s\hat{\sigma} |\uparrow s\rangle = s |\uparrow s\rangle$ and $s\hat{\sigma} |\downarrow s\rangle = -s |\downarrow s\rangle$. Then we can summarize the eigenvalues and eigenstates of the density matrix in the following way:

$$\frac{\hat{I} + s\hat{\sigma}}{2} |\uparrow s\rangle = \frac{1+s}{2} |\uparrow s\rangle, \quad \frac{\hat{I} + s\hat{\sigma}}{2} |\downarrow s\rangle = \frac{1-s}{2} |\downarrow s\rangle.$$

5.4 Magnetic rotation for $|\uparrow\rangle \longrightarrow |\downarrow\rangle$. We must implement a π -rotation of the polarization vector and we can choose the rotation vector $(\pi, 0, 0)$ which means π -rotation around the x -axis. In magnetic field ω , the polarization vector s satisfies the classical equation of motion $\dot{s} = \omega \times s$ meaning that s will rotate around the direction ω of the field at angular velocity ω . Accordingly, we can choose the field to point along the x -axis: $\omega = (\omega, 0, 0)$. The rotation angle π is achieved if we switch on the field for a period $t = \pi/\omega$.

5.5 Interrelated qubit physical quantities.

$$\begin{aligned} \hat{P}_n + \hat{P}_{-n} &= \hat{I} \\ 2\hat{P}_n - \hat{\sigma}_n &= 2\hat{P}_{-n} + \hat{\sigma}_n = \hat{I} \end{aligned}$$

5.6 Mixing non-orthogonal polarizations. Since the qubit density matrix is a linear function of the polarization vector, mixing the density matrices means averaging their polarization vectors with the mixing weights. Therefore our mixture has the following polarization vector:

$$s = \frac{1}{3} \times (0, 0, 1) + \frac{2}{3} \times (1, 0, 0) = (1/3, 0, 2/3).$$

Problems of Chap. 6

6.1 Universality of Hadamard and phase operations. The unitary rotations $\hat{U}$ of the qubit space are, apart from an irrelevant phase, equivalent to the spatial rotations of the corresponding Bloch sphere. This time the three Euler angles ψ, θ, ϕ are the natural parameters. We can write the unitary rotations, corresponding to the spatial ones, into this form:

$$\hat{U}(\psi, \theta, \phi) = \exp\left(-\frac{i}{2}\psi\hat{\sigma}_z\right) \exp\left(-\frac{i}{2}\theta\hat{\sigma}_x\right) \exp\left(-\frac{i}{2}\phi\hat{\sigma}_z\right).$$

The middle factor, too, becomes rotation around the z -axis if we sandwich it between two Hadamard operations because $\hat{H}\hat{\sigma}_x\hat{H} = \hat{\sigma}_z$. We can thus express the r.h.s. in the desired form:

$$\hat{U}(\psi, \theta, \phi) = \hat{T}(\psi)\hat{H}\hat{T}(\theta)\hat{H}\hat{T}(\phi).$$

6.2 Statistical error of qubit determination. Out of N , we allocate N_x, N_y, N_z qubits to estimate s_x, s_y, s_z , respectively. We learned that the estimated value of s_x takes this form:

$$\frac{N_{\uparrow x} - N_{\downarrow x}}{N_{\uparrow x} + N_{\downarrow x}} = \frac{2N_{\uparrow x}}{N_x} - 1,$$

because on a large statistics $N_x = N_{\uparrow x} + N_{\downarrow x}$ the ratio $N_{\uparrow x}/N_x$ converges to the q -theoretical prediction $p_{\uparrow x} = \langle \uparrow x | \hat{\rho} | \uparrow x \rangle \equiv \frac{1}{2}(1 + s_x)$. The statistical error of the estimation takes the form $2\Delta N_{\uparrow x}/N_x$ and we are going to determine the mean fluctuation $\Delta N_{\uparrow x}$. The statistical distribution of the count $N_{\uparrow x}$ is binomial:

$$p(N_{\uparrow x}) = \binom{N_x}{N_{\uparrow x}} p_{\uparrow x}^{N_{\uparrow x}} p_{\downarrow x}^{N_{\downarrow x}},$$

hence the mean squared fluctuation of the count $N_{\uparrow x}$ takes the form $(\Delta N_{\uparrow x})^2 = N_x p_{\uparrow x} p_{\downarrow x} = N_x(1 - s_x^2)/4$. This yields the ultimate form of the estimation error:

$$\Delta s_x = \sqrt{\frac{1 - s_x^2}{N_x}},$$

and we could get similar results for Δs_y and Δs_z .

6.3 Fidelity of qubit determination. If the state $|n\rangle$ sent by Alice and the polarization $\hat{\sigma}_m$ chosen by Bob were fixed then the structure of the expected fidelity of Bob's guess would be this:

$$|\langle n|m\rangle|^2 p_{\uparrow m} + |\langle n|-m\rangle|^2 p_{\downarrow m}.$$

Here we have understood that Bob's optimum guess must always be the post-measurement state $|\pm m\rangle$ based on the measurement outcome $\hat{\sigma}_m = \pm 1$, respectively. Now we recall that $|\langle n|m\rangle|^2 = p_{\uparrow m} = \cos^2(\vartheta/2)$ where $\cos \vartheta = nm$, and

$|\langle \mathbf{n} | -\mathbf{m} \rangle|^2 = p_{\perp m} = \sin^2(\vartheta/2)$. Hence the above fidelity takes the simple form $\cos^4(\vartheta/2) + \sin^4(\vartheta/2)$ which we rewrite into the equivalent form $\frac{1}{2} + \frac{1}{2} \cos^2(\vartheta)$. The average of $\cos^2(\vartheta) = (\mathbf{n}\mathbf{m})^2$ over the random independent $\mathbf{n}$ and $\mathbf{m}$ yields $1/3$ therefore the expected fidelity of Bob's guess becomes $2/3$.

6.4 Post-measurement depolarization. Let $\hat{\sigma}_n$ denote the polarization chosen by Bob. The non-selective measurement induces the change $\hat{\rho} \rightarrow \hat{P}_n \hat{\rho} \hat{P}_n + \hat{P}_{-n} \hat{\rho} \hat{P}_{-n}$ of the state. Inserting the Pauli-representation of $\hat{\rho}$ and the projectors $\hat{P}_{\pm n}$ yields:

$$\begin{aligned} \frac{\hat{I} + s\hat{\sigma}}{2} &\rightarrow \frac{\hat{I} + \hat{\sigma}_n}{2} \frac{\hat{I} + s\hat{\sigma}}{2} \frac{\hat{I} + \hat{\sigma}_n}{2} + \frac{\hat{I} - \hat{\sigma}_n}{2} \frac{\hat{I} + s\hat{\sigma}}{2} \frac{\hat{I} - \hat{\sigma}_n}{2} = \\ &= \frac{\hat{I} + s\hat{\sigma}}{4} + \frac{\hat{I} + s\hat{\sigma}_n \hat{\sigma} \hat{\sigma}_n}{4} . \end{aligned}$$

Since Bob's choice is random regarding $\mathbf{n}$ we shall average $\mathbf{n}$ over the solid angle. Averaging the structure $\hat{\sigma}_n \hat{\sigma} \hat{\sigma}_n$ yields $-\hat{\sigma}/3$ hence the average influence of Bob's non-selective measurements can be summarized as:

$$\frac{\hat{I} + s\hat{\sigma}}{2} \rightarrow \frac{\hat{I} + s\hat{\sigma}/3}{2} .$$

6.5 Anti-linearity of polarization reflection. Let us calculate the influence of the anti-unitary transformation $\hat{T}$ on a pure state qubit:

$$\hat{T} |\mathbf{n}\rangle = \hat{T} \left(\cos \frac{\theta}{2} |\uparrow\rangle + e^{i\varphi} \sin \frac{\theta}{2} |\downarrow\rangle \right) = -\cos \frac{\theta}{2} |\downarrow\rangle + e^{-i\varphi} \sin \frac{\theta}{2} |\uparrow\rangle = e^{-i\varphi} |-\mathbf{n}\rangle .$$

6.6 General qubit effects. We can suppose that the weights w_n are non-vanishing. First, we have to impose the conditions $|\mathbf{a}_n| \leq 1$ since otherwise the matrices would be indefinite. Second, the request $\hat{\Pi}_n \geq 0$ implies the conditions $w_n > 0$. And third, the request $\sum_n \hat{\Pi}_n = \hat{I}$ implies the conditions $\sum_n w_n = 1$ and $\sum_n w_n \mathbf{a}_n = 0$.

Problems of Chap. 7

7.1 Schmidt orthogonalization theorem. Let r be the rank of $\hat{c}$ and let us consider the non-negative matrices $\hat{c}\hat{c}^\dagger$ and $\hat{c}^\dagger\hat{c}$ of rank r both. Their spectrum is non-negative and identical. Indeed, if $\hat{c}^\dagger\hat{c}|R\rangle = w|R\rangle$, i.e., w and $|R\rangle$ are an eigenvalue and a (normalized) eigenvector of $\hat{c}^\dagger\hat{c}$ then $|L\rangle = \hat{c}|R\rangle/\sqrt{w}$ will be a (normalized) eigenvector of $\hat{c}\hat{c}^\dagger$ with the same eigenvalue w . This can be seen by direct inspection. Now we determine the r positive eigenvalues w_λ for $\lambda = 1, 2, \dots, r$ and the corresponding orthonormal eigenstates $|\lambda; R\rangle$ of $\hat{c}^\dagger\hat{c}$. Then, by $|\lambda; L\rangle = \hat{c}|\lambda; R\rangle/\sqrt{w_\lambda}$, we define the r orthonormal eigenstates of $\hat{c}\hat{c}^\dagger$ which belong to the common positive eigenvalues w_λ , for $\lambda = 1, 2, \dots, r$. Now we can see that

$$\hat{c}|\lambda; R\rangle = \sqrt{w_\lambda}|\lambda; L\rangle ,$$

for all $\lambda = 1, 2, \dots, r$. We have thus proven that there exists the following Schmidt decomposition of the matrix $\hat{c}$:

$$\hat{c} = \sum_{\lambda=1}^r \sqrt{w_\lambda} |\lambda; L\rangle \langle \lambda; R| .$$

7.2 Swap operation. For convenience, we introduce $\hat{\sigma}^{(\pm)} = (\hat{\sigma}_x \pm i\hat{\sigma}_y)/2$ instead of $\hat{\sigma}_x, \hat{\sigma}_y$. Now we express the Pauli matrices in the up-down basis:

$$\hat{\sigma}^{(+)} = |\uparrow\rangle\langle\downarrow|, \quad \hat{\sigma}^{(-)} = |\downarrow\rangle\langle\uparrow|, \quad \hat{\sigma}_z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow| .$$

Substituting these expressions we obtain:

$$\begin{aligned} \frac{\hat{I} \otimes \hat{I} + \hat{\sigma} \otimes \hat{\sigma}}{2} &= \frac{\hat{I} \otimes \hat{I} + 2\hat{\sigma}^{(+)} \otimes \hat{\sigma}^{(-)} + 2\hat{\sigma}^{(-)} \otimes \hat{\sigma}^{(+)} + \hat{\sigma}_z \otimes \hat{\sigma}_z}{2} = \\ &= |\uparrow\uparrow\rangle\langle\uparrow\uparrow| + |\downarrow\downarrow\rangle\langle\downarrow\downarrow| + |\uparrow\downarrow\rangle\langle\downarrow\uparrow| + |\downarrow\uparrow\rangle\langle\uparrow\downarrow| , \end{aligned}$$

which is indeed the swap matrix $\hat{S}$.

7.3 Singlet density matrix. The singlet state $\hat{\rho}(\text{singlet})$ is invariant under rotations of the Bloch sphere. Therefore $\hat{\rho}(\text{singlet})$ must be of the form:

$$\hat{\rho}(\text{singlet}) = \frac{\hat{I} \otimes \hat{I} + \lambda \hat{\sigma} \otimes \hat{\sigma}}{4}$$

because there are no further rotational invariant mathematical structures. We could determine the parameter λ from the pure state condition $[\hat{\rho}(\text{singlet})]^2 = \hat{\rho}(\text{singlet})$, yielding $\lambda = -1$. However, we can spare these calculations if we recall the swap $\hat{S}$. It is Hermitian, rotation invariant and idempotent: $\hat{S}^2 = \hat{I} \otimes \hat{I}$. Hence we get the singlet state directly in the form:

$$\hat{\rho}(\text{singlet}) = \frac{\hat{I} \otimes \hat{I} - \hat{S}}{2} = \frac{\hat{I} \otimes \hat{I} - \hat{\sigma} \otimes \hat{\sigma}}{4} .$$

7.4 Local measurement of expectation values. Alice and Bob will determine $\langle \hat{A} \otimes \hat{B} \rangle$ and $\langle \hat{A}' \otimes \hat{B}' \rangle$ separately on two independent sub-ensembles and will finally add them since the expectation value is additive. Still we have to show that the expectation value of a tensor product, like $\langle \hat{A} \otimes \hat{B} \rangle$, can be determined in local measurements. We introduce the local spectral expansions $\hat{A} = \sum_{\lambda} A_{\lambda} \hat{P}_{\lambda}$ and $\hat{B} = \sum_{\mu} B_{\mu} \hat{Q}_{\mu}$. Alice and Bob perform local measurements of $\hat{A}$ and $\hat{B}$ in coincidence, yielding the measurement outcomes $A_1, B_1, A_2, B_2, \dots, A_r, B_r, \dots, A_N, B_N$ where A_r is always an eigenvalue A_{λ} and the case is similar for the B_r 's. Then Alice and Bob can calculate the q-expectation value asymptotically:

$$\langle \hat{A} \otimes \hat{B} \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_r A_r B_r .$$

To see that this is indeed the right expression of $\langle \hat{A} \otimes \hat{B} \rangle$ we have to rethink the nonlocal measurement of $\hat{A} \otimes \hat{B}$ itself. Its spectral expansion is

$$\hat{A} \otimes \hat{B} = \sum_{(\lambda, \mu)} (A_\lambda B_\mu) (\hat{P}_\lambda \otimes \hat{Q}_\mu) ,$$

and the corresponding q-measurement will obviously yield the same statistics of the outcomes $A_r B_r$ like in case of the local-measurements.

7.5 Local measurement of certain nonlocal quantities. If we measure $\hat{\sigma}_z \otimes \hat{\sigma}_z$ on a singlet state we always get -1 and the singlet state remains the post-measurement state. In the attempted local measurement, the entanglement is always destroyed and we get either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$ for the post-measurement state. Obviously the degenerate spectrum of $\hat{\sigma}_z \otimes \hat{\sigma}_z$ plays a role in the nonlocality. If, in the general case, we suppose $\hat{A} \otimes \hat{B}$ has a non-degenerate spectrum then the post-measurement states will be the same pure states in both the nonlocal measurement of $\hat{A} \otimes \hat{B}$ and the simultaneous local measurements of $\hat{A}$ and $\hat{B}$.

7.6 Nonlocal hidden parameters. Let the further hidden parameter ν take values 1, 2, 3, 4 marking whether Alice and Bob measures $\hat{A} \otimes \hat{B}$, $\hat{A}' \otimes \hat{B}$, $\hat{A} \otimes \hat{B}'$ or $\hat{A}' \otimes \hat{B}'$, respectively. Then, according to the hidden variable concept, the assignment of all four polarization values will uniquely depend on the composite hidden variable $r\nu$:

$$\hat{A} = A_{r\nu} = \pm 1 , \quad \hat{A}' = A'_{r\nu} = \pm 1 , \quad \hat{B} = B_{r\nu} = \pm 1 , \quad \hat{B}' = B'_{r\nu} = \pm 1 .$$

Contrary to the local assignment (7.36), the above assignments are called nonlocal since the hidden variable $r\nu$ is nonlocal: it depends on both Alice's and Bob's measurement setup. The statistical relationships, cf. (7.37), become modified:

$$\langle \hat{A} \otimes \hat{B} \rangle = \lim_{N_1 \rightarrow \infty} \frac{1}{N_1} \sum_{r \in \Omega_1} A_{r1} B_{r1} ; \quad N_1 = |\Omega_1| ,$$

$$\langle \hat{A}' \otimes \hat{B} \rangle = \lim_{N_2 \rightarrow \infty} \frac{1}{N_2} \sum_{r \in \Omega_2} A_{r2} B_{r2} ; \quad N_2 = |\Omega_2| ,$$

etc. for the other two cases $\nu = 3, 4$. The assignments are independent for the four different values of ν . There is no constraint combining the $A_{r\nu}$'s with different ν 's. Hence it has become straightforward to reproduce the above said four q-theoretical predictions including of course correlations $\langle \hat{C} \rangle$ that are higher than 2.

7.7 Does teleportation clone the qubit? The selective post-measurement state of the two qubits on Alice's side is one of the four maximally entangled Bell-states. Therefore the reduced state of the qubit that she had teleported is left in the totally mixed state independently of its original form as well as of the four outcomes of Alice's measurement. Note that the form (7.47) of the three-qubit pre-measurement state shows that Alice's measurement outcome is always random. The four outcomes have probability $1/4$ each.

Problems of Chap. 8

8.1 All q-operations are reductions of unitary dynamics. Given the trace-preserving q-operation $\mathcal{M}\hat{\rho} = \sum_n \hat{M}_n \hat{\rho} \hat{M}_n^\dagger$, we have to construct the unitary interaction matrix $\hat{U}$ acting on the composite state of the system and environment. Let us introduce the composite basis $|\lambda\rangle \otimes |n; E\rangle$ where $\lambda = 1, 2, \dots, d$ and $n = 1, 2, \dots, d_E$. Let us define the influence of $\hat{U}$ on a subset of the composite basis:

$$\hat{U}(|\lambda\rangle \otimes |1; E\rangle) = \sum_{n=1}^{d_E} \hat{M}_n |\lambda\rangle \otimes |n; E\rangle, \quad \lambda = 1, 2, \dots, d.$$

This definition is possible because the above map generates orthonormal vectors:

$$\sum_{m=1}^{d_E} \langle \mu | \hat{M}_m^\dagger \otimes \langle m; E | \sum_{n=1}^{d_E} \hat{M}_n |\lambda\rangle \otimes |n; E\rangle = \sum_{n=1}^{d_E} \langle \mu | \hat{M}_n^\dagger \hat{M}_n |\lambda\rangle = \delta_{\lambda\mu}.$$

The further matrix elements of $\hat{U}$, i.e. those not defined by our first equation above, can be chosen in such a way that $\hat{U}$ is unitary on the whole composite state. Using this definition of $\hat{U}$ in the equation (8.3) of reduced dynamics we can directly inspect that the resulting operation is $\mathcal{M}\hat{\rho} = \sum_n \hat{M}_n \hat{\rho} \hat{M}_n^\dagger$, as expected.

8.2 Non-projective effect as averaged projection. Let us substitute the proposed form of the effects $\hat{I}_n$ into the equation $p_n = \text{tr}(\hat{I}_n \hat{\rho})$ introduced for non-projective measurement in Sect. 4.4.2:

$$\text{tr}(\hat{I}_n \hat{\rho}) = \text{tr}(\text{tr}_E \hat{P}_n \hat{\rho}_E \hat{\rho}) = \text{tr}(\hat{P}_n \hat{\rho}_E \hat{\rho}).$$

In this formalism, i.e., without the $\otimes$'s, the matrices of different subsystems commute hence $\hat{\rho}_E \hat{\rho} = \hat{\rho} \hat{\rho}_E$. Thus we obtain the following result: $\text{tr}(\hat{I}_n \hat{\rho}) = \text{tr}(\hat{P}_n \hat{\rho} \hat{\rho}_E)$. We recognize the coincidence of the r.h.s. with the r.h.s. of (8.18). Since this coincidence is valid for all possible $\hat{\rho}$, it verifies the proposed form of $\hat{I}_n$.

8.3 Q-operation as supermatrix. We start from the Kraus representation $\mathcal{M}\hat{\rho} = \sum_n \hat{M}_n \hat{\rho} \hat{M}_n^\dagger$. We take the matrix elements of both sides and we also sandwich the $\hat{\rho}$ between the identities $\sum_{\lambda'} |\lambda'\rangle \langle \lambda'|$ and $\sum_{\mu'} |\mu'\rangle \langle \mu'|$ on the r.h.s.:

$$\langle \lambda | \mathcal{M}\hat{\rho} | \mu \rangle = \langle \lambda | \sum_n \hat{M}_n \sum_{\lambda'} |\lambda'\rangle \langle \lambda'| \hat{\rho} \sum_{\mu'} |\mu'\rangle \langle \mu'| \hat{M}_n^\dagger | \mu \rangle.$$

Comparing the r.h.s. with $\sum_{\lambda', \mu'} \mathcal{M}_{\lambda\mu\lambda'\mu'} \rho_{\lambda'\mu'}$, we read out the components of the supermatrix: $\mathcal{M}_{\lambda\mu\lambda'\mu'} = \sum_n \langle \lambda | \hat{M}_n | \lambda' \rangle \langle \mu' | \hat{M}_n^\dagger | \mu \rangle$.

8.4 Environmental decoherence, time-continuous depolarization. The equation takes the Lindblad form with $\hat{H} = 0$ and with three hermitian Lindblad matrices identified by the Cartesian components of $(\hat{\sigma}/2\sqrt{\tau})$. For convenience, we shall use the Einstein convention to sum over repeated indices, e.g.: $s\hat{\sigma} = s_a \hat{\sigma}_a$. We write

the r.h.s. of the master equation into the equivalent form $-(1/8\tau)[\hat{\sigma}_b, [\hat{\sigma}_b, \hat{\rho}]]$ and insert $\hat{\rho} = \frac{1}{2}(\hat{I} + s_a \hat{\sigma}_a)$ into it. The master equation reduces to:

$$\dot{s}_a \hat{\sigma}_a = -\frac{1}{8\tau}[\hat{\sigma}_b, [\hat{\sigma}_b, s_a \hat{\sigma}_a]] = -\frac{1}{\tau} s_a \hat{\sigma}_a ,$$

which means the simple equation $\dot{\mathbf{s}} = -\mathbf{s}/\tau$ for the polarization vector. Its solution is $\mathbf{s}(t) = e^{-t/\tau} \mathbf{s}(0)$. Therefore the τ may be called depolarization time, or decoherence time as well.

8.5 Kraus representation of depolarization. The map should be of the form $\mathcal{M}\hat{\rho} = (1 - 3\lambda)\hat{\rho} + \lambda\hat{\sigma}\hat{\rho}\hat{\sigma}$ with $0 \leq \lambda \leq 1/3$ since there exist no other isotropic Kraus structures for a qubit. The depolarization channel decreases the polarization vector $\mathbf{s}$ by a factor $1 - w$ and we have to find the parameter λ as function of w . Inserting $\hat{\rho} = \frac{1}{2}(\hat{I} + \mathbf{s}\hat{\sigma})$ we get

$$\mathcal{M}\frac{\hat{I} + \mathbf{s}\hat{\sigma}}{2} = \frac{\hat{I} + (1 - 4\lambda)\mathbf{s}\hat{\sigma}}{2} ,$$

which means that $\lambda = w/4$. Four Kraus matrices make the depolarization channel: $\sqrt{1 - 3w/4}\hat{I}$ and the three components of $\sqrt{w/4}\hat{\sigma}$.

Problems of Chap. 9

9.1 Positivity of relative entropy. We can write:

$$S(\rho' \parallel \rho) = \sum_{\mathbf{x}} \rho(\mathbf{x}) \log \frac{\rho(\mathbf{x})}{\rho'(\mathbf{x})} .$$

We invoke the inequality $\ln \lambda > 1 - \lambda^{-1}$ valid for $\lambda \neq 1$ and apply it to $\lambda = \rho/\rho'$. This yields:

$$S(\rho' \parallel \rho) > \frac{1}{\ln 2} \sum_{\mathbf{x}} \rho(\mathbf{x}) \left[1 - \frac{\rho'(\mathbf{x})}{\rho(\mathbf{x})} \right] = 0 ,$$

which always holds if $\rho' \neq \rho$.

9.2 Concavity of entropy. Suppose we have a long message $x_1^{(1)} x_2^{(1)} \dots x_n^{(1)}$ where $\rho_1(\mathbf{x})$ is the apriori distribution of one letter. Let S_1 stand for the single-letter entropy $S(\rho_1)$. The number of the typical messages is 2^{nS_1} so that their shortest code is nS_1 bits. Consider a second message from the same alphabet and suppose the single-letter distribution $\rho_2(\mathbf{x})$ is different from $\rho_1(\mathbf{x})$. Let us concatenate the two messages:

$$x_1^{(1)} x_2^{(1)} x_3^{(1)} \dots x_n^{(1)} x_1^{(2)} x_2^{(2)} x_3^{(2)} \dots x_m^{(2)} ,$$

where the two lengths n and m may be different. Obviously, the number of the typical ones among such composite messages is $2^{nS_1} \times 2^{mS_2}$ and their shortest code is $nS_1 + mS_2$ bits. Now imagine that we permute the $n+m$ letters randomly. On one

hand, the composite messages become usual $(n + m)$ -letter-long messages where the single letter distribution is always the same, i.e., the mixture $\rho = w_1\rho_1 + w_2\rho_2$ with weights $w_1 = n/(n + m)$ and $w_2 = m/(n + m)$. Therefore the number of the typical messages is $2^{(n+m)S(\rho)}$ and the shortest code is $(n + m)S(\rho)$ bits. On the other hand, we can inspect that the number of the typical messages $2^{(n+m)S(\rho)}$ is greater than $2^{nS_1} \times 2^{mS_2}$ because the number of inequivalent permutations has increased: the first n letters have become permutable with the last m letters. This means that $(n + m)S(\rho) > nS_1 + mS_2$ which is just the concavity of the entropy: $S(w_1\rho_1 + w_2\rho_2) > w_1S(\rho_1) + w_2S(\rho_2)$, for $\rho_1 \neq \rho_2$.

9.3 Subadditivity of entropy. We can make the choice $\rho'_{AB}(x, y) = \rho_A(x)\rho_B(y)$. To calculate $S(\rho'_{AB}||\rho_{AB}) = -S(\rho_{AB}) - \sum_{x,y} \rho_{AB}(x, y) \log \rho'_{AB}(x, y)$, we note that the second term is $-\sum_{x,y} \rho_{AB}(x, y) \log[\rho_A(x)\rho_B(y)] = S(\hat{\rho}_A) + S(\hat{\rho}_B)$. Hence the positivity of the relative entropy $S(\rho'_{AB}||\rho_{AB}) \geq 0$ proves subadditivity: $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$.

9.4 Coarse graining increases entropy. Let us identify our system by the k -partite composite system of the k bits $x_1, x_2, \dots, x_k$. Then the coarse grained system corresponds to the $(k - 1)$ -partite sub-system consisting of the first $k - 1$ bits $x_1, x_2, \dots, x_{k-1}$. The coarse grained state $\tilde{\rho}$ is just a reduced state w.r.t. ρ . Hence we see that coarse graining increases the entropy because reduction does it.

Problems of Chap. 10

10.1 Subadditivity of q-entropy. Let us calculate

$$S(\hat{\rho}_A \otimes \hat{\rho}_B || \hat{\rho}_{AB}) = -S(\hat{\rho}_{AB}) - \text{tr}[\hat{\rho}_{AB} \log(\hat{\rho}_A \otimes \hat{\rho}_B)]$$

and note that the second term is $S(\hat{\rho}_A) + S(\hat{\rho}_B)$. Hence the Klein inequality $S(\hat{\rho}'||\hat{\rho}) \geq 0$ proves the subadditivity: $S(\hat{\rho}_{AB}) \leq S(\hat{\rho}_A) + S(\hat{\rho}_B)$.

10.2 Concavity of q-entropy, Holevo entropy. We assume a certain environmental system E and a basis $\{|n; E\rangle\}$ for it. Let us construct a composite state:

$$\hat{\rho}_{big} = \sum_n w_n \hat{\rho}_n \otimes |n; E\rangle \langle n; E| .$$

Note that the reduced state of the system is invariably $\hat{\rho} = \sum_n w_n \hat{\rho}_n$ and the reduced state of the environment is $\hat{\rho}_E = \sum_n w_n |n; E\rangle \langle n; E|$. Subadditivity guarantees that $S(\hat{\rho}_{big}) \leq S(\hat{\rho}) + S(\hat{\rho}_E)$. Let us calculate and insert the entropies $S(\hat{\rho}_{big}) = S(w) + \sum_n w_n S(\hat{\rho}_n)$ and $S(\hat{\rho}_E) = S(w)$ which results in the desired inequality: $\sum_n w_n S(\hat{\rho}_n) \leq S(\hat{\rho})$.

10.3 Data compression of the non-orthogonal code. The density matrix of the corresponding 1-letter q-message reads:

$$\hat{\rho} = \frac{|\uparrow z\rangle \langle \uparrow z| + |\uparrow x\rangle \langle \uparrow x|}{2} = \frac{\hat{I} + \hat{\sigma}_n / \sqrt{2}}{2} ,$$

which is a partially polarized state along the skew direction $\mathbf{n} = (1, 0, 1)/\sqrt{2}$, cf. (6.29). The eigenvalues of this density matrix are the following:

$$p_+ = \frac{1 + 1/\sqrt{2}}{2}, \quad p_- = \frac{1 - 1/\sqrt{2}}{2},$$

hence its von Neumann entropy amounts to:

$$S(\hat{\rho}) = -p_+ \log p_+ - p_- \log p_- \approx 0.60.$$

According to the q-data compression theorem, we can compress one qubit of the q-message into 0.6 qubit on average, and this is the best maximum faithful compression.

10.4 Distinguishing two non-orthogonal qubits: various aspects. In fact, we measure the polarization component orthogonal to the polarization of the single-letter density matrix. The measurement outcomes ± 1 on both q-states $|\uparrow z\rangle, |\uparrow x\rangle$ will appear with probabilities p_+ and p_- (cf. Prob. 10.3), in alternating order of course:

$$\begin{aligned} p(y = +1|x = 0) &= p_+, & p(y = -1|x = 0) &= p_- \\ p(y = +1|x = 1) &= p_-, & p(y = -1|x = 1) &= p_+. \end{aligned}$$

Regarding the randomness of the input message the output message, too, becomes random: $H(Y) = 1$. Hence the information gain takes this form and value:

$$I_{\text{gain}} = H(Y) - H(Y|X) = 1 + p_+ \log p_+ + p_- \log p_- \approx 0.40.$$

10.5 Simple optimum q-code. The q-data compression theory says that a pure state q-code is not compressible faithfully (i.e.: allowing the same accessible information) if and only if the single-letter average state has the maximum von Neumann entropy. In our case, we must assure the following:

$$\frac{|R\rangle\langle R| + |G\rangle\langle G| + |B\rangle\langle B|}{3} = \frac{\hat{I}}{2},$$

which is possible if we chose three points on a main circle of the Bloch-sphere, at equal distances from each other.

Problems of Chap. 11

11.1 Creating the totally symmetric state.

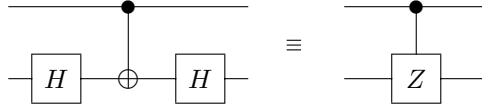
$$\begin{aligned} |S\rangle &\equiv \frac{1}{2^{n/2}} \sum_{x=1}^{2^n-1} |x_n x_{n-1} \dots x_1\rangle = \sum_{x_n=0}^1 \sum_{x_{n-1}=0}^1 \dots \sum_{x_1=0}^1 |x_n\rangle \otimes |x_{n-1}\rangle \otimes \dots \otimes |x_1\rangle \\ &= \hat{H}|0\rangle \otimes \hat{H}|0\rangle \otimes \dots \otimes \hat{H}|0\rangle \equiv \hat{H}^{\otimes n}|0\rangle^{\otimes n}. \end{aligned}$$

11.2 Constructing Z-gate from X-gate.

$$\hat{H}\hat{X}\hat{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \hat{Z}.$$

The inverse relationship follows from $\hat{H}^2 = \hat{I}$.

11.3 Constructing controlled Z-gate from cNOT-gate.

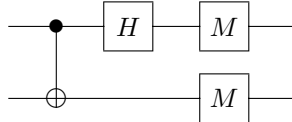


11.4 Q-circuit to produce Bell states. Let us calculate the successive actions of the H-gate and the cNOT-gate:

$$\begin{array}{llll} |00\rangle & \longrightarrow & |00\rangle + |10\rangle & \longrightarrow & |00\rangle + |11\rangle & \longrightarrow & |\Phi^{(+)}\rangle \\ |01\rangle & \longrightarrow & |01\rangle + |11\rangle & \longrightarrow & |01\rangle + |10\rangle & \longrightarrow & |\Psi^{(+)}\rangle \\ |10\rangle & \longrightarrow & |00\rangle - |10\rangle & \longrightarrow & |00\rangle - |11\rangle & \longrightarrow & |\Phi^{(-)}\rangle \\ |11\rangle & \longrightarrow & |01\rangle - |11\rangle & \longrightarrow & |01\rangle - |10\rangle & \longrightarrow & |\Psi^{(-)}\rangle \end{array}$$

The trivial factors $1/\sqrt{2}$ in front of the intermediate states have not been denoted.

11.5 Q-circuit to measure Bell states. The task is the inverse task of preparing the Bell states. Since both the H-gate and the cNOT-gate are the inverses of themselves, respectively, we can simply use them in the reversed order w.r.t. the circuit that prepared the Bell states (cf. Prob. 11.4):



The boxes $\boxed{M}$ stand for projective measurement of the computational basis.

References

1. M. A. Nielsen, I. L. Chuang: *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge 2000)
2. J. Preskill: *Quantum Computation and Information*, (Caltech 1998); <http://theory.caltech.edu/people/preskill/ph229/>
3. D. Bouwmeester, A. Ekert, A. Zeilinger (editors): *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*, (Springer, Berlin Heidelberg New York 2000)
4. L. D. Landau, E. M. Lifshic: *Course of Theoretical Physics: Mechanics*, (Pergamon, Oxford 1960)
5. M. C. Gutzwiller: *Chaos in Classical and Quantum Mechanics*, (Springer, Berlin Heidelberg New York 1991)
6. J. von Neumann: *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, Princeton 1955)
7. A. Peres: *Quantum Theory: Concepts and Methods*, (Kluwer, Dordrecht 1993)
8. K. Kraus: *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, (Springer, Berlin Heidelberg New York, 1983)
9. P. Busch, P. J. Lahti, P. Mittelstadt: *The Quantum Theory of Measurement*, (Springer, Berlin Heidelberg New York 1991)
10. E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, K. Kupsch, I. O. Stamatescu: *Decoherence and the Appearance of a Classical World in Quantum Theory*, 2nd edit. (Springer, Berlin Heidelberg New York 2003)
11. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden: *Rev. Mod. Phys.* **74**, 145 (2002)
12. A. Ekert, P. Hayden, H. Inamori: *Basic Concepts in Quantum Computation*, (Les Houches lectures 2000); Los Alamos e-print arXiv: quant-ph/0011013
13. L. Diósi: Weak measurements in quantum mechanics. In: *Encyclopedia of Mathematical Physics*, vol 4, ed by J.P. Francoise, G.L. Naber, S.T. Tso (Elsevier, Oxford 2006) pp 276-282
14. R. F. Werner: *Phys. Rev. A* **40**, 4277 (1989)
15. W. K. Wootters, W. K. Zurek: *Nature* **299**, 802 (1982)
16. I. D. Ivanovic: *Phys. Lett. A* **123**, 257 (1987)
17. S. Wiesner: *SIGACT News* **15**, 77 (1983)
18. C. H. Bennett: *Phys. Rev. Lett.* **68**, 3121 (1992)
19. C. H. Bennett, G. Brassard: Quantum cryptography: Public key distribution and coin tossing, In: *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, (IEEE Press, New York 1984)
20. A. Einstein, B. Podolsky, N. Rosen: *Phys. Rev.* **48**, 696 (1935)
21. W. F. Stinespring: *Proc. Am. Math. Soc.* **6**, 211 (1955)
22. A. Peres: *Phy. Rev. Lett.* **77**, 1413 (1996)

23. M. Horodecki, P. Horodecki, R. Horodecki: Phys. Lett. A **223**, 1 (1996)
24. J. S. Bell: Physics **1**, 195 (1964)
25. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt: Phys. Rev. Lett. **49**, 1804 (1969)
26. S. Popescu: Phys. Rev. Lett. **74**, 2619 (1995)
27. C. H. Bennett, S. J. Wiesner: Phys. Rev. Lett. **69**, 2881 (1992)
28. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters: Phys. Rev. Lett. **70**, 1895 (1993)
29. G. Lindblad: Commun. Math. Phys. **48**, 199 (1976)
30. V. Gorini, A. Kossakowski, E. C. G. Sudarshan: J. Math. Phys. **17**, 821 (1976)
31. C. E. Shannon: Bell System Tech. J. **27**, 379, 623 (1948)
32. B. Schumacher: Phys. Rev A **51**, 2738 (1995)
33. A. S. Holevo: Problems of Inf. Transm. **5**, 247 (1979)
34. R. P. Feynman: Int. J. Theor. Phys. **21**, 467 (1982)
35. D. Deutsch: Proc. Roy. Soc. **400**, 97 (1985)
36. D. Deutsch, R. Jozsa: Proc. Roy. Soc. London A **439**, 533 (1992)
37. L. K. Grover: A fast quantum mechanical algorithm for database search. In: *Proc. of the 28th annual STOC* (Assoc. for Comp. Machinery, New York 1996)
38. R. R. Tucci: *QC Paulinesia*, <http://www.ar-tiste.com/PaulinesiaVer1.pdf> 2004
39. P. W. Shor: Algorithms for quantum computation: discrete logarithm and factoring. In: *35th Annual Symposium on Foundations of Computer Science*, (IEEE Press, Los Alamitos 1994)

Index

- Alice, Bob 36, 44
- Alice, Bob, Eve 48
- Bell
 - basis, states 58
 - inequality 63
 - nonlocality 62
- bit 11
- channel capacity 83
- code 80
 - optimum 83
 - q- 88
 - superdense 64
- completely positive map 69
- contrary to classical 20, 22, 23, 29
- data compression 80, 89
- decoherence 23
- density matrix 19, 34
- entanglement 19, 74
 - as resource 91
 - dilution 94
 - distillation 57, 93
 - maximum 58
 - measure 55
 - two-qubit 57
- entropy
 - conditional 82
 - relative 79, 88
 - Shannon 55, 79
 - von Neumann 56, 87
- equation
 - master 25, 75
 - of motion 6, 20, 35
- expectation value 8, 23
- fidelity 36, 44
- function evaluation 100
- information, q-information 11, 84
 - accessible 91
 - mutual 82
 - theory 79, 87
- irreversible
 - master equation 25
 - operation 10, 29, 41
 - q-measurement 23
 - reduced dynamics 75
- Klein inequality 88
- Kraus form 69
- Lindblad form 75
- local
 - Hamilton 56
 - operation 56
 - physical quantity 57
- LOCC 74
- measurement 6, 21, 35
 - continuous 24
 - in pure state 26
 - indirect 71, 73
 - non-projective 9, 19, 23, 46, 73
 - non-selective 8, 23
 - projective 7, 22, 46
 - selective 8, 22
 - unsharp 9
- message 80, 88
 - typical 81
- mixing 5, 20, 42, 43
- nonlocality
 - Bell 62
 - Einstein 59
- operation 5, 20, 69

- depolarization 40
- irreversible 84, 95
- local 56
- logical 39
- one-qubit 39
- reflection 40
- selective 6, 20, 72
- Pauli 32
 - matrices 32
 - representation 32, 37
- physical quantities 7, 21, 35
 - compatible 25
- q-algorithm
 - Fourier 104
 - oracle problem 101
 - searching 103
- q-banknote 47
- q-channel 49, 75
- q-circuit 105
- q-computation
 - parallel 99
 - representation 31, 37
- q-correlation 28, 64
 - history 59
- q-cryptography 48, 84
- q-entropy 87
- q-gate 105
 - universal 40
- q-information
 - hidden 40
- q-key 48
- q-protocol 47
- q-state
 - cloning 45
 - determination 42, 43, 45
 - indistinguishability 45, 47
 - no-cloning 44, 47
 - non-orthogonal 45
 - preparation 42
 - purification 54
 - unknown 39, 43
- qubit 16, 31
 - unknown 36
- reduced dynamics 10, 29, 70
- rotational invariance 33
- Schmidt decomposition 53
- selection 5, 20
- state space 5, 19
 - discrete 12
- state, q-state
 - mixed 5, 19
 - pure 5, 19
 - separable 10, 28
- superposition 16, 19
- system
 - bipartite 53
 - collective 11, 29
 - composite 9, 27, 53
 - environmental 54
 - open 75
- teleportation 65

Lecture Notes in Physics

For information about earlier volumes
please contact your bookseller or Springer
LNP Online archive: springerlink.com

Vol.667: W. D. Heiss (Ed.), Quantum Dots: a Doorway to Nanoscale Physics

Vol.668: H. Ocampo, S. Paycha, A. Vargas (Eds.), Geometric and Topological Methods for Quantum Field Theory

Vol.669: G. Amelino-Camelia, J. Kowalski-Glikman (Eds.), Planck Scale Effects in Astrophysics and Cosmology

Vol.670: A. Dinklage, G. Marx, T. Klinger, L. Schweikhard (Eds.), Plasma Physics

Vol.671: J.-R. Chazottes, B. Fernandez (Eds.), Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems

Vol.672: R. Kh. Zeytounian, Topics in Hypersonic Flow Theory

Vol.673: C. Bona, C. Palenzuela-Luque, Elements of Numerical Relativity

Vol.674: A. G. Hunt, Percolation Theory for Flow in Porous Media

Vol.675: M. Kröger, Models for Polymeric and Anisotropic Liquids

Vol.676: I. Galanakis, P. H. Dederichs (Eds.), Half-metallic Alloys

Vol.677: A. Loiseau, P. Launois, P. Petit, S. Roche, J.-P. Salvetat (Eds.), Understanding Carbon Nanotubes

Vol.678: M. Donath, W. Nolting (Eds.), Local-Moment Ferromagnets

Vol.679: A. Das, B. K. Chakrabarti (Eds.), Quantum Annealing and Related Optimization Methods

Vol.680: G. Cuniberti, G. Fagas, K. Richter (Eds.), Introducing Molecular Electronics

Vol.681: A. Llor, Statistical Hydrodynamic Models for Developed Mixing Instability Flows

Vol.682: J. Souchay (Ed.), Dynamics of Extended Celestial Bodies and Rings

Vol.683: R. Dvorak, F. Freistetter, J. Kurths (Eds.), Chaos and Stability in Planetary Systems

Vol.684: J. Dolinšek, M. Vilfan, S. Žumer (Eds.), Novel NMR and EPR Techniques

Vol.685: C. Klein, O. Richter, Ernst Equation and Riemann Surfaces

Vol.686: A. D. Yaghjian, Relativistic Dynamics of a Charged Sphere

Vol.687: J. W. LaBelle, R. A. Treumann (Eds.), Geospace Electromagnetic Waves and Radiation

Vol.688: M. C. Miguel, J. M. Rubi (Eds.), Jamming, Yielding, and Irreversible Deformation in Condensed Matter

Vol.689: W. Pötz, J. Fabian, U. Hohenester (Eds.), Quantum Coherence

Vol.690: J. Asch, A. Joye (Eds.), Mathematical Physics of Quantum Mechanics

Vol.691: S. S. Abdullaev, Construction of Mappings for Hamiltonian Systems and Their Applications

Vol.692: J. Frauendiener, D. J. W. Giulini, V. Perlick (Eds.), Analytical and Numerical Approaches to Mathematical Relativity

Vol.693: D. Alloin, R. Johnson, P. Lira (Eds.), Physics of Active Galactic Nuclei at all Scales

Vol.694: H. Schwoerer, J. Magill, B. Beleites (Eds.), Lasers and Nuclei

Vol.695: J. Dereziński, H. Siedentop (Eds.), Large Coulomb Systems

Vol.696: K.-S. Choi, J. E. Kim, Quarks and Leptons From Orbifolded Superstring

Vol.697: E. Beaurepaire, H. Bulou, F. Scheurer, J.-P. Kappler (Eds.), Magnetism: A Synchrotron Radiation Approach

Vol.698: S. Bellucci (Ed.), Supersymmetric Mechanics – Vol. 1

Vol.699: J.-P. Rozelot (Ed.), Solar and Heliospheric Origins of Space Weather Phenomena

Vol.700: J. Al-Khalili, E. Roeckl (Eds.), The Euroschool Lectures on Physics with Exotic Beams, Vol. II

Vol.701: S. Bellucci, S. Ferrara, A. Marrani, Supersymmetric Mechanics – Vol. 2

Vol.702: J. Ehlers, C. Lämmerzahl, Special Relativity

Vol.703: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1

Vol.704: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2

Vol.705: P. Bhattacharyya, B.K. Chakrabarti (Eds.), Modelling Critical and Catastrophic Phenomena in Geoscience

Vol.706: M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, E.K.U. Gross (Eds.), Time-Dependent Density Functional Theory

Vol.707: A.V. Shechetilov, Calculus and Mechanics on Two-Point Homogenous Riemannian Spaces

Vol.708: F. Iachello, Lie Algebras and Applications

Vol.709: H.-J. Borchers and R.N. Sen, Mathematical Implications of Einstein-Weyl Causality

Vol.710: K. Hutter, A.A.F. van de Ven, A. Ursescu, Electromagnetic Field Matter Interactions in Thermoelastic Solids and Viscous Fluids

Vol.711: H. Linke, Controlled Nanoscale Motion in Biological and Artificial Systems

Vol.712: W. Pötz, J. Fabian, U. Hohenester (Eds.), Modern Aspects of Spin Physics

Vol.713: L. Diósi, A Short Course in Quantum Information Theory