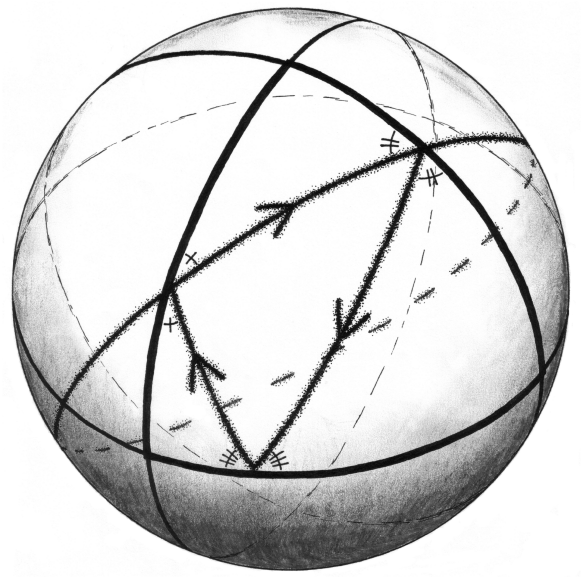
Exploring the Set of Quantum States

Åsa Ericsson



Doctoral thesis in Theoretical Physics
Fysikum
Stockholms Universitet, 2007

Thesis for the degree of Doctor of philosophy in Theoretical Physics

Department of Physics
Stockholm University
Sweden

ISBN 978-91-7155-475-8 (pp. i-xii, 1-82)

Printed by Universitetsservice US-AB, Stockholm 2007

Cover illustration by Sören Holst
A bouncing Bures-Uhlmann geodesic.

Abstract

Quantum mechanical properties of finite dimensional quantum systems are used within the field of quantum information. In this thesis the set of states (density matrices) for such systems is studied and described, largely in geometrical terms. The introductory part also acquaints the reader with relevant background about majorization, bistochastic matrices, mutually unbiased bases, Hadamard matrices and entanglement, with the aim to make the papers attached easier to read.

Paper I considers Peres' criterion for separability, for two qubit states. Paper II deals with the problem of how density matrices can be mixed from pure states, especially what probability distributions over pure states that are possible. In Paper III the set of bistochastic matrices—Birkhoff's polytope—and the subset of unistochastic matrices is studied, with a detailed description in dimensions 3 and 4. In Paper IV it is seen how the states of a complete set of mutually unbiased bases form a polytope in the set of density matrices, with certain combinatorial properties. A search for mutually unbiased bases in dimension 6 is presented in Paper VI, which includes a thorough discussion on 6 by 6 Hadamard matrices. Paper V presents a result about geodesics in the set of quantum states with respect to the curved Bures-Uhlmann geometry.

"I must apologize
for not having penetrated quantum mechanics
deeply enough."

Albert Einstein [32]

Preface

This is my doctoral thesis in Theoretical Physics. It deals with quantum states, also known as density matrices. I first came across density matrices when I started my project as a Masters student. Gently guided by my supervisor, my views were widened. I got to know about quantum information, and working on concrete problems I came to appreciate the algebra I never understood in my linear algebra course. (Why, oh why, didn't anyone tell us that matrices and eigenvalues actually mean something?) That there are some weird things going on in quantum mechanics I learned in my first physics course, but I wasn't aware that these mysterious features can be used in a marvelous way within quantum information. (Admittedly the weirdness is also 'in use' when quantum mechanics marvelously explains plenty of physics.) For a beginner quantum info is not only fascinating, it is also fairly accessible without first ploughing through thick books. During my Masters project, I got the taste for it. I wanted to continue with research. And I wanted to learn more of quantum information.

I am happy I got the chance to go on as a grad student—it's been a privilege. I have now had five more years at Fysikum, Stockholms universitet. The research I have done during this time is presented in this theses. Six published papers, along with an introductory text containing a lot of what I have learnt. So, am I now an expert? That's not how I feel. There is so much, much more I would like to know and master. At the same time I have realized, while striving to write the introductory text, that I have taken in quite a lot these years. Certainly more than I can put down in print in this thesis.

I am very much indebted to Professor Ingemar Bengtsson. He is an excellent supervisor, and without him I would not have been able to carry this thesis to term. He has generously shared of his knowledge and proposed problems to study, always interested in any of my progress and willingly given of his time in discussions. I have much appreciated his eagerness to tell me what's on his mind and we have shared the joy of illustrating physics with hand drawn pictures. With great patience he has guided me, constantly encouraged me and given abundant advice—especially during these last months when I have been struggling to finish my thesis, often drained of creativity.

I will miss Ingemar. But he is not the only one who will be missed, now that I am leaving Fysikum. I am grateful to many, who have contributed to making our workplace such an enjoyable one. So thank you,

all my friends in ‘kof’, ‘cops’, and former ‘fop’ groups. Many times, when lacking inspiration to work, I’ve travelled to Fysikum just to meet people whom I knew could cheer me up. We have had many lunches, coffee and dinner breaks, discussing almost anything in the world, and we’ve been watching movies—thanks, Maria!—as well as playing table tennis—thanks, Emil and Narit!—and much more. Occasionally we’ve been working. Especially I want to thank Sören, with whom I have shared a room during all my time at Fysikum. We have had many interesting discussions, often disagreeing furiously. But he is always helpful and ready to give advice (as he has been a student of Ingemar’s, too, he is the perfect support whenever I actually don’t agree with my supervisor). During my first years here I also benefitted from sharing a room with Johan, who could help a novice in quantum oriented questions.

Besides working with Ingemar, I have also had the opportunity to coauthor papers with Karol Życzkowski, Jan-Åke Larsson, Marek Kuś, Wojciech Tadej, and Wojciech Bruzda. Thanks to all of you! For providing linguistic advice, in parts of this thesis, I thank Subhash Chaturvedi.

There are also people outside Physics in my life. I want to thank my dear friend Sara, with whom most things can be chewed over. I am also thankful to friends in my church: to everyone in my “cell group” and to Maria for providing a shoulder to lean on in hard times. Finally I thank my mom and dad for their constant support.

Åsa Ericsson

Stockholm, August 2007

List of Accompanying Papers

The following papers are included in this thesis. I have tried to specify my contributions to the papers I have coauthored, although it is difficult to say who has done what of the research presented. Results emerge from the exchange of ideas and from discussions in front of the black board.

Paper I Å. Ericsson,
Separability and the stella octangula,
Phys. Lett. A **295** (2002) 256.

Ingemar Bengtsson proposed the question studied in this paper.

Paper II I. Bengtsson and Å. Ericsson,
How to mix a density matrix,
Phys. Rev. A **67** (2003) 012107.

The main idea is mine. I found the counterexamples and why Nielsen's construction do not always give different pure states.

Paper III I. Bengtsson, Å. Ericsson, M. Kuś, W. Tadej
and K. Życzkowski,
Birkhoff's polytope and unistochastic matrices,
$N = 3$ *and* $N = 4$,
Commun. Math. Phys. **259**, 307-324 (2005).

I participated in discussions, especially about the geometrical parts. The geometrical descriptions were primarily done in Stockholm by me and Ingemar Bengtsson, whereas all published results from computer calculations were done by our colleagues in Poland.

Paper IV I. Bengtsson and Å. Ericsson,
Mutually unbiased bases and the complementarity polytope,
Open Sys. & Information Dyn. (2005) **12**: 107-120.

I did the work presented in Section 4 and I wrote the manuscript.

Paper V Å. Ericsson,
Geodesics and the best measurement for distinguishing quantum states,
J. Phys. A: Math. Gen. **38** (2005) L725-L730.

I acknowledge Ingemar Bengtsson for showing me the works of Uhlmann on geometry of quantum state space, and the work of Fuchs and Caves, on optimal distinguishing measurements.

Paper VI I. Bengtsson, W. Bruzda, Å. Ericsson, J.-Å. Larsson, W. Tadej and K. Życzkowski,
Mutually unbiased bases and Hadamards of Order Six,
J. Math. Phys. **48**, 052106 (2007).

I did substantial work analyzing data from computer calculations to convert the essentials into understandable form. I participated in crucial discussions, especially about equivalences.

Contents

Chapter 1

Introduction

"One of the most fascinating aspects of recent work in fundamental quantum theory is the emergence of a new notion, the concept of quantum information, which is quite distinct from its classical counterpart. It provides a new perspective for all foundational and interpretational issues and highlights new essential differences between classical and quantum theory."

Richard Jozsa [56]

1.1 A (very) brief history of quantum mechanics

In the beginning of the twentieth century the concept of the 'quantum' was invented to model physics that did not fit the classical theory. It was the rise of a sweeping transformation that would renew physics in the years to come. In the mid nineteen-twenties the theory of quantum mechanics was formulated, essentially as we know it today. It was tremendously efficient for describing various properties of the constituents of our universe. But how should reality be understood if quantum mechanics is the way to model it? The theory seemed so weird!

The determinism of classical physics were lost in favor of quantum probabilities. In his 1927 paper on the uncertainty principle Heisenberg said [48]:

"Even in principle, we cannot know the present in all detail. For that reason everything observed is a selection from a plen-

itude of possibilities and a limitation on what is possible in the future."

This is interconnected to the complementarity of observables, the absence of definite values of complementary quantities. As Bohr wrote [20]:

> *"[T]he indivisibility of the quantum of action is itself, from the classical point of view, an irrational element which inevitably requires us to forgo a causal mode of description and which . . . forces us to adopt a new mode of description designated as complementary in the sense that any given application of classical concepts precludes the simultaneous use of other classical concepts which in a different connection are equally necessary for the elucidation of the phenomena."*

Or with the words of Pauli [66]:

> *"One can see it with p-eyes and one can see it with q-eyes, but if one opens both eyes then one goes astray."*

Not only was classical determinism superseded by quantum probabilities, there were also these strange "spooky actions at a distance", as clarified in the famous "EPR-paper" [33]. This phenomenon was considered by Schrödinger, who was the first to call states 'entangled' [72]:

> *"When two systems . . . enter into temporary . . . interaction . . . and when after a time . . . the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives [the quantum states] have become entangled."*

He comments upon the "EPR-paper":

> *"Attention has recently been called to the obvious but very disconcerting fact that even though we restrict the disentangling measurements to one system, the representative obtained for the other system is by no means independent of the particular choice of observations which we select . . . It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it."*

The strangeness of quantum mechanics was discussed and there were struggles to try to intertwine quantum mechanics with classical concepts. But to most physicists these abstrusenesses were, for a long time, regarded as subtleties no need to worry about. After all, quantum mechanics was successful and there were so many discoveries to make, through application of this new theory. With quantum mechanics the stability of everyday matter could be explained. Quantum mechanics was also fruitfully applied to key properties of matter on all scales, from elementary particles to stars.

But the oddities remained. Somehow, until the late twentieth century, they did not enter the applications manifestly.

1.2 Quantum information —exploiting the oddities

The advances of technology has lead to the possibility of manipulating systems of single or a few atoms or photons with such a high precision that the strangeness of quantum mechanics can be investigated experimentally. Perhaps most famous are the Bell-experiments which display the non-local character of entangled quantum states.

This capability of controlling single quantum systems has, together with novel theoretical insights, given rise to the new field of quantum information science. There has been a shift in perspective. Instead of seeing the impossibilities resulting from quantum uncertainties and the weirdness of entanglement one asks: How can the oddities of quantum mechanics be employed?

The outgrowth is the discovery of many fascinating phenomena, wherein quantum properties are utilized for information processing. Creative scientists have figured out how, for example, entanglement can be used for quantum teleportation and dense coding, how chryptography can benefit from the limited distinguishability of quantum states which is a consequence of complementarity, and how superpositions of quantum states enable quantum computing. The stumbling blocks in the efforts of understanding the foundations of quantum mechanics have become indispensable resources in quantum information science.

Several ideas have been successfully implemented experimentally. The basic quantum system in most applications is a qubit, that is, a two level quantum system. There are various physical realizations in use—photons, electrons, ions, molecules, quantum dots and superconducting circuits—

all having different advantages and different drawbacks.

To prepare one single qubit in an arbitrary state, to perform unitary transformations and to measure it can, these days, be done without too much trouble. But what is really new is the possibility to create interesting states also in the higher dimensional Hilbert space of several qubits. The interaction between qubits can be controlled so that superpositions of product states can be prepared. We will have a look at just one example of what can be done in the laboratory.

When ions are used as qubits the two relevant quantum levels are the electronic ground state and one of the exited states, whereas the rest of the energy levels are unoccupied. The ions can be trapped with electromagnetic fields, they can be controlled with laser pulses and with highly sensitive CCD cameras fluorescence light can be detected to measure the states of the ions. In our example the experimentalists used calcium ions. As qubit levels the ground state $S_{1/2}$ and the metastable exited state $D_{5/2}$ where employed. In one experiment they managed to have up to eight calcium ions in a row in the same ion trap [45]. The dimension of the Hilbert space of states for all eight qubits is thus $2^8 = 256$. With lasers and with the aid of electrostatic interactions between the ions all eight have been prepared in a so called W-state: $|\mathrm{W}\rangle = \frac{1}{\sqrt{N}}(|\mathrm{D}\cdots\mathrm{DDS}\rangle + |\mathrm{D}\cdots\mathrm{DSD}\rangle + \cdots + |\mathrm{SD}\cdots\mathrm{D}\rangle)$. This is a highly entangled state, where every ion is entangled with all the others. From measurement results it has been successfully established that the prepared state really is close to the W-state and that it is entangled.

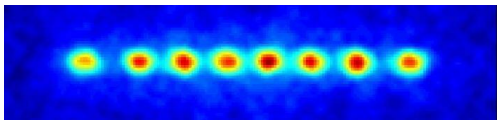


Figure 1-1: *Eight entangled calcium ions in a linear Paul trap.*

When such things can be done in the laboratory, there is a need for a deeper study of the properties of quantum states. Some technicalities addressed in the youth of quantum mechanics, which did not get much attention back then, have now become relevant not only for foundational questions but primarily for their implications for quantum information processing. A case in point is the discussion of mixtures of quantum states by Schrödinger in the mid thirties, when he generalized the situation of the EPR-paper. It is, for instance, of importance for the possible transformations that might be realized between entangled states. This work by Schrödinger is one thing that will be adressed in this thesis,

along with ideas and concepts that have been introduced into quantum mechanics only quite recently.

1.3 What will come—and what will not

What have been said here motivates investigations of the details in the formalism of quantum mechanics. One central part are the states of the theory. This is the issue where this thesis is hoped to give a modest contribution. On the forthcoming pages aspects of the set of quantum states will be explored.

Quantum systems used for quantum information processing are most often those having a finite number of levels, by which I mean that the relevant Hilbert spaces for the states are finite dimensional. We will restrict ourselves to the states of such systems. Some facts can be generalized to the infinite case, and also to the continuous case, but for others the cautions needed when handling infinite dimensions might prohibit straightforward generalizations.

In chapters 2 and 3 we will see that the set of quantum states make up a convex set, conveniently described as a set in a vector space. A central feature, which is also the theme of Paper II, is the multitude of ways any mixed state can be obtained as convex combinations of pure states. This is based upon Schrödingers work from the mid thirties. In the latter part of chapter 3 a curved geometry on the set of quantum states is introduced—it is the "geometry of distinguishability". Geodesics of this geometry appear in Paper V.

Chapter 4 explains the concepts of majorization and bistochastic matrices—mathematics which finds its way into physics in various places. The set of bistochastic matrices—Birkhoff's polytope—is studied in Paper III.

In chapter 5 we encounter complementary observables in finite dimensions, commonly called "mutually unbiased bases". The central question is: Can we find "complete sets" of mutually unbiased bases. Papers IV and VI are also devoted to these bases.

And in chapter 6 we will have a glance on the "magical entangled states". They are a resource for quantum information processing. A visualization of the set of entangled two-qubit states is found in Paper I.

In this thesis I will not try to say what quantum states are, in the ontological sense, or how they shall (or not even how they can) be interpreted; it is certainly an interesting issue, but far too intricate for this thesis. My ambition is to not presuppose any interpretation, and merely

conciliate with an instrumentalistic view; still the attentive reader might scent shortcomings.

I end this introduction by subscribing to the words of Weyl [87]:

> *"The development of quantum theory has only been made possible by the enormous* refinement of experimental technique. *which has given us an almost direct insight into atomic processes. If in the following little is said concerning the experimental facts, it should not be attributed to the mathematical haughtiness of the author; to report on these things lies outside his field. Allow me to express now, once and for all, my deep respect for the work of the experimenter and for his fight to wring significant facts from an inflexible Nature, who says so distinctly "No" and so indistinctly "Yes" to our theories."*

Chapter 2

Quantum States

The concept of a *state* is central in theories of physics. To describe what is meant by a state is in general not so easy, but one might say something like this:

Every physical system is in some particular state, and depending on which state that "is the case", there are different predictions about what will be observed. When the system is specified the state is supposed to include all that can be said about the system, within the framework we have chosen to work in. Framework here means theory but could also include system dependent assumptions, like for example, Hamiltonians.*

In quantum mechanics so-called pure states are given by complex vectors in a Hilbert space associated to the quantum mechanical system. These vectors can be combined to more general states, described by matrices, and the space of these matrices is the state space for the quantum mechanical system. This we will now explore.

2.1 Pure states and mixed states †

"We ... assume each state of a dynamical system at a particular time corresponds to a ket vector ..."[30]. This or something similar is what you learn in most textbooks on quantum mechanics. Such states, represented by vectors $|\psi\rangle$ in a Hilbert space are called *pure states*. We will only deal with states in finite dimensional Hilbert spaces $\mathcal{H}^N$ of (complex)

* This obviously just gives a hint about what is meant by a state. One intricate aspect I have avoided is what is meant by a system. Another remark is that also within the same theory, 'state' can be dealt with quite differently—just think of the Schrödinger and the Heisenberg pictures in quantum mechanics.

† A nice source of knowledge of these things is Preskill's Lecture Notes [68].

dimension N. Every vector $c|\psi\rangle$, for any $c \in \mathbb{C}$, represents the same state as $|\psi\rangle$, hence the one-to-one correspondence is between physical states and rays in Hilbert space. This set of rays is the complex projective space $\mathbb{C}\mathrm{P}^{N-1}$. It has $2N-2$ real dimensions. We will choose our state vectors to be normalized, $|\langle\psi|\psi\rangle| = 1$, always achievable because of the arbitrary number c.

So, how can we make predictions about measurement outcomes from a state vector $|\psi\rangle$? Every measurement that might be performed on a quantum system can be described by a *POVM*, a positive operator valued measure. This is a set of Hermitian operators E_i —called POVM-elements—acting on the Hilbert space $\mathcal{H}^N$. The index i labels the measurement outcomes. Every POVM-element must be non-negative, $E_i \geq 0$, and the full set must be complete, which means that $\sum_i E_i = \mathbb{1}$ (the identity operator).[‡] Upon measurement, the outcome labeled i will occur with probability

$$P_i = \langle\psi|E_i|\psi\rangle \ . \tag{2-1}$$

The non-negativity of the POVM-elements ensures that the probabilities are positive and completeness that they sum to one. Quantum mechanics is a probabilistic theory: the best predictions we can get is the probabilities for the possible outcomes of any measurement.

Although this is not the place to expand upon the theory of quantum measurements I will use a few lines to relate POVMs with the more commonly known projective measurements, also called von Neumann measurements. Such a measurement is what we get whenever the POVM-elements are orthogonal one-dimensional projectors, $E_i = |e_i\rangle\langle e_i|$. This is often said to be a measurement of an observable, represented by the Hermitian operator

$$O = \sum_{i=1}^{N} \lambda_i |e_i\rangle\langle e_i| \ . \tag{2-2}$$

Here the eigenvalues λ_i correspond to the values of the possible outcomes. Any Hermitian operator can be written in the form (2-2)—the spectral representation—with eigenvectors $|e_i\rangle$ forming an ON-basis, and any such operator corresponds to an observable that can be measured. If the operator has a degenerate spectrum some POVM-elements have higher rank. They are projectors onto the eigenspaces corresponding to the degenerate eigenvalues.

[‡] This is enough as far as the outcomes are concerned. If one is also interested in the state after the measurement, measurement operators M_i, with $E_i = M_i^\dagger M_i$, are required.

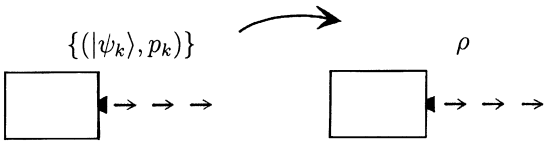


Figure 2-1: *If a source emits particles described by states $|\psi_k\rangle$ with probabilities p_k, it will be possible to determine the statistics of any measurement from the mixed state $\rho = \sum_k p_k \, |\psi_k\rangle\langle\psi_k|$.*

Now consider a situation where we are given states from a set $\{|\psi_k\rangle\}_{k=1}^{M}$. Suppose each state $|\psi_k\rangle$ is given with probability p_k, $\sum_k p_k = 1$; this might be the result of some preparation process in the lab. Then the probability for the outcome i is

$$P_i = \sum_{k=1}^{M} p_k \, \langle\psi_k|E_i|\psi_k\rangle = \mathrm{Tr}\left(E_i \sum_{k=1}^{M} p_k \, |\psi_k\rangle\langle\psi_k| \right) = \mathrm{Tr}(E_i \, \rho) \ , \qquad (2\text{-}3)$$

where, in the last step, we defined

$$\rho \equiv \sum_{k=1}^{M} p_k \, |\psi_k\rangle\langle\psi_k| \ . \qquad (2\text{-}4)$$

The operator ρ is called a density matrix. It represents a general state in quantum mechanics: a *mixed state*. The density matrix includes all there is to say about the probabilities for possible outcomes of measurements.[§]

From equation (2-4) it follows that density matrices ρ fulfill the following conditions:

$$\begin{array}{lll} (a) & \rho^\dagger = \rho & \text{(hermiticity)} \\ (b) & \rho \geq 0 & \text{(non-negative eigenvalues)} \\ (c) & \mathrm{Tr}\,\rho = 1 & \text{(normalization)} \end{array} \qquad (2\text{-}5)$$

Also the converse is true: any Hermitian, non-negative operator on $\mathcal{H}$, with unit trace, is a density matrix. Moreover, this is the most general quantum state one can have. It can always be written in the same form as in (2-4). But this does not necessarily mean that the state has been prepared by mixing pure states $\{|\psi_k\rangle\}_{k=1}^{M}$ with probabilities p_k, as will be made clear later.

[§] von Neumann's original term for what is nowadays usually called a density matrix, or density operator, was "statistical operator" [85]—in my opinion a term that conveys more of what it stands for. Nevertheless I will use the more conventional term.

If we have the state $|\psi_{k'}\rangle$ with probability $p_{k'} = 1$ for some $k = k'$, we recover the formulas for pure states. The density matrix is then the projector

$$\rho = |\psi'_k\rangle\langle\psi'_k| \ , \quad \text{with} \ \ \rho^2 = \rho \ . \tag{2-6}$$

All projectors fulfilling $\rho^2 = \rho$ and $\mathrm{Tr}\rho = 1$ correspond to pure states (since any unit trace projector has rank one).

Another way to look at density matrices emerges from considering systems composed of two or more parts. A bipartite system has two subsystems, A and B, and the pure states are rays in a tensor product Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $\{|\, i\rangle_A\}$ and $\{|\mu\rangle_B\}$ be orthonormal bases in $\mathcal{H}_A$ respectively $\mathcal{H}_B$. A general pure state can be expanded as

$$|\Psi\rangle = \sum_{i,\mu} c_{i\mu} |\, i\rangle_A \otimes |\mu\rangle_B \ , \ \text{where} \ \sum_{i,\mu} |c_{i\mu}|^2 = 1 \ . \tag{2-7}$$

If we only have access to subsystem A, all our POVM-elements will be of the form $E_i = E_{Ai} \otimes \mathbb{1}_B$; we can choose any POVM in $\mathcal{H}_A$ but we cannot act with anything in $\mathcal{H}_B$ other than the identity operator. Let's look at the probabilities for the outcomes.

$$P_i = \langle\Psi|E_{Ai} \otimes \mathbb{1}_B|\Psi\rangle = \sum_{j,\nu,i,\mu} c^*_{j\nu}\, c_{i\mu}\, {}_A\langle j|E_{Ai}|\, i\rangle_A\, {}_B\langle\nu|\mathbb{1}_B|\mu\rangle_B = \mathrm{Tr}\,(E_{Ai}\rho_A) \ . \tag{2-8}$$

In the last step we have introduced the *reduced density matrix* ρ_A. It is defined as the partial trace over subsystem B of the density matrix $\rho = |\Psi\rangle\langle\Psi|$:

$$\rho_A \equiv \mathrm{Tr}_B\, \rho \equiv \sum_\kappa {}_B\langle\kappa|\rho|\kappa\rangle_B \ , \tag{2-9}$$

which in this case is

$$\rho_A = \sum_{i,j} \Big(\sum_\mu c^*_{j\mu}\, c_{i\mu} \Big) |\, i\rangle_A\, {}_A\langle j| \ . \tag{2-10}$$

The conditions (2-5) are satisfied by ρ_A and this is really the density matrix for system A, when considered as a system on its own, independently of system B. ρ_A is the only operator for which $\mathrm{Tr}\,(E_{Ai}\rho_A)$ would give the correct probabilities for all POVM-elements $E_i = E_{Ai} \otimes \mathbb{1}_B$. Also when the combined system is in a mixed state ρ, the states of the subsystems are obtained by "tracing out" the other systems, according to (2-9). Conversely, every mixed state ρ can be obtained by partial trace of some pure state for a larger bipartite system; there exist "purifications" of ρ in

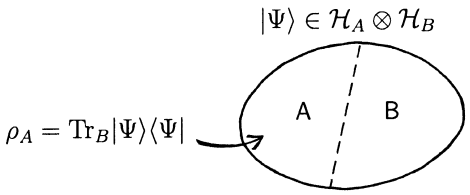


Figure 2-2: *The state of a subsystem is obtained from the state for the whole system by taking the partial trace over the other subsystem(s).*

the state space of the bipartite (perhaps imagined) system. We will come back to purifications in section 3.2, where it will be explained how they give rise to an interesting geometry.

Also entanglement will be considered later (in chapter 6), but let me note one thing about the connection between entanglement and mixed states. Whenever a combined system is in a pure state, the reduced state of a subsystem will be pure if and only if it is not entangled with the rest of the system.

A useful result for bipartite systems is the following. For every state $|\Psi\rangle \in \mathcal{H}_A^N \otimes \mathcal{H}_B^M$, it is always possible to find orthonormal bases $\{|\, i'\rangle_A\}$ and $\{|\mu_i'\rangle_B\}$, in $\mathcal{H}_A^N$ respectively $\mathcal{H}_B^M$, such that

$$|\Psi\rangle = \sum_{i=1}^{\min\{N,M\}} c_i |\, i'\rangle_A \otimes |\mu_i'\rangle_B \quad , \quad \text{with } c_i \text{ real} \geq 0 \; . \tag{2-11}$$

This is called the *Schmidt decomposition*. Note that here we only have a sum over one index, giving $\min\{N, M\}$ terms, while in equation (2-7) there are NM terms. Using the bases of the Schmidt decomposition, the formula for the reduced density matrix for system A simplifies to

$$\rho_A = \mathrm{Tr}_B\, \rho = \sum_i c_i^2 \, |\, i'\rangle_{A\ A}\langle\, i'| \; . \tag{2-12}$$

The reduced state for system B will be described by the same matrix, except for some zero eigenvalues in the case where the systems A and B have different dimensions.

2.2 Mixing density matrices

We have seen two different ways of arriving at a density matrix for describing the state of a quantum system. First it was through the idea of

an *ensemble*, a mixture of pure states. With an ensemble I mean a set of (different) pure states and associated probabilities: $\{(|\psi_k\rangle, p_k)\}_{k=1}^{M}$, $p_k \geq 0$, $\sum_i p_k = 1$.¶ The corresponding density matrix is given by (2-4). The second way was as a reduced density matrix for a subsystem, as in equation (2-9). But we can use the spectral decomposition to see that the state described by the reduced density matrix equally well could have arisen from an ensemble in the subsystem: any density matrix ρ can be seen as a mixture of its eigenvectors $|e_i\rangle$ with the probabilities given by the eigenvalues λ_i,

$$\rho = \sum_{i=1}^{N} \lambda_i |e_i\rangle\langle e_i| \ . \tag{2-13}$$

The ensemble $\{(|e_i\rangle, \lambda_i)\}$ is called the eigenensemble (or perhaps *an* eigenensemble, since it is not unique if the eigenvalues are degenerate). It follows that whatever ensemble we chose, with two or more non-orthogonal states, there is another ensemble—the eigenensemble—giving the same density matrix. Thus, there can be different ways to mix the same density matrix. But the state of the system is nevertheless fully determined by the density matrix, from which the probabilities of the outcomes of any measurement on the system can be calculated.

Since a quantum state, even if it is pure, does not predict with certainty the outcomes of most measurements, it is sometimes said that there are quantum probabilities intrinsic to the quantum systems, in contrast to classical probabilities which arise only because we don't know every detail. Using this language the probabilities p_i in equation (2-4) are classical, and the states $|\psi_i\rangle$ "contain" quantum probabilities. But if every mixed state can be considered as an ensemble in several different ways, it is hard to make sense of this division of the probabilities. There is, for example, no well justified way to say how much of the probability for some measurement outcome is classical and how much is quantum. It is thus reasonable to speak only of one sort of probabilities. The difference is, that within quantum theory the probabilities have a more fundamental role, compared to that in classical theories. Thus, I believe, to better understand and find satisfying interpretations of quantum states (if at all possible—it might be highly dependent upon personal taste), one needs

¶ Unfortunately the term "ensemble" is used frequently with slightly different meanings in quantum mechanics. In contrast to this mathematical definition it often refers to a set of systems prepared in the same way or, at least, described with the same state (density matrix). However, the two views are not contradictory, and they might "coincide".

to know what one means by probabilities. And here I will stop this discussion, because probabilities is surely not an easy subject—and it is not the target of this thesis.

Back to the possible ways to mix a quantum state: What mixings can we have? What states and with what probabilities? This is the subject being explored in Paper II [9]. Already Schrödinger wrote about it [73], in 1936 when he considered a generalization of the EPR-scenario [33]. He wrote an exciting paper showing how an experimenter can, with some non-zero probability, force a subsystem into any pure state of her choice, by only manipulating the other part of the system.[‖] This possibility hinges on quantum entanglement, which will be considered in chapter 6. Schrödinger makes the comment: *"The statement is hardly more than a corollary to a theorem about 'mixtures' for which I claim no priority but the permission of deducing it ... for it is certainly not* well *known."* Unfortunately this theorem about mixtures did not become well known from then on either. That is why the theorem is often referred to as the HJW-theorem, after Hughston, Jozsa and Wootters [52]. Their paper is similar to Schrödinger's, except that it was written in 1993, when the interest in quantum entanglement and quantum information was on the rise.

Here follows the theorem, as it is stated in Paper II. Remember that $|e_i\rangle$ stands for an eigenvector of ρ, while $|\psi_i\rangle$ can be some other state vector.

Schrödinger's Mixture Theorem:
A density matrix ρ having the diagonal form

$$\rho = \sum_{i=1}^{N} \lambda_i \, |e_i\rangle\langle e_i| \tag{2-14}$$

can be written in the form

$$\rho = \sum_{i=1}^{M} p_i \, |\psi_i\rangle\langle\psi_i| \ , \qquad p_i > 0 \ , \quad \sum_{i=1}^{M} p_i = 1 \tag{2-15}$$

[‖] This sounds really baffling, but already the formulation of the sentence presupposes an interpretation of quantum states as some kind of property of a system. Within other interpretations this phenomenon might seem less mysterious, although still out of reach of classical physics.

if and only if there exists a unitary $M \times M$ matrix U such that

$$|\psi_i\rangle = \frac{1}{\sqrt{p_i}} \sum_{j=1}^{N} U_{ij} \sqrt{\lambda_j}\, |e_j\rangle \ , \quad i = 1, \ldots, M \ , \quad M \geq N \ . \tag{2-16}$$

One should not mistake U to be an operator on the Hilbert space. Instead U acts on the list of eigenvectors $|e_j\rangle$, in such a way that the state $|\psi_i\rangle$ is a superposition of the vectors $|e_j\rangle$ with coefficients $c_j^{(i)}$ computed from the i:th row of U ($c_j^{(i)} = \sqrt{\lambda_j / p_i}\ U_{ij}$). For more comments useful in understanding Schrödinger's mixture theorem consult Paper II. (Note, however, that the meanings of M and N are reversed in Paper II—an inconvenience turning up because here I have chosen N to denote the dimension of Hilbert space.)

The probabilities with which the states $|\psi_i\rangle$ occur, in (2-15), are found from (2-16). If we take the scalar product of $\sqrt{p_i}|\psi_i\rangle$ with itself we get

$$p_i = \sum_{j=1}^{M} B_{ij}\, \lambda_j \ , \quad \text{where } B_{ij} = |U_{ij}|^2 \ . \tag{2-17}$$

By construction B is a bistochastic matrix. This, and some of its consequences, were noted by Uhlmann [78]. These matrices will be discussed in chapter 4 and are also the subject of Paper III [11]. Here it suffices to say that equation (2-17) gives a means of characterizing different probability distributions p_i, consistent with a given density matrix. This was studied by Nielsen, who also gave a procedure for finding state vectors corresponding to a probability distribution p_i [64]. However, the procedure sometimes results in a set of state vectors, such that several vectors actually give the same state. One could, for example, find a pure state

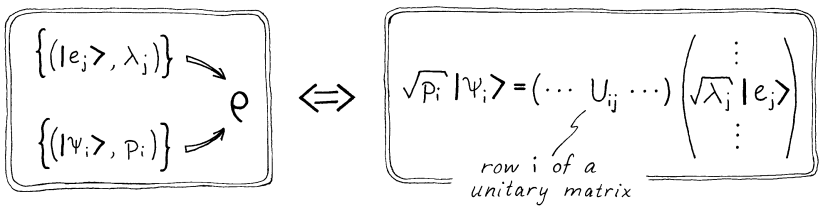


Figure 2-3: *Illustration of Schrödinger's mixture theorem.*

expressed as a mixture of itself with different probabilities.** To avoid this, we need to impose the reasonable requirement that states in an ensemble should be different. In Paper II [9] the difficulties are pointed out, and a partial characterization of the possible probability distributions is given.

Equation (2-17) gives the probabilities in terms of the eigenvalues of ρ and information—encoded in the bistochastic matrix B—about the states $|\psi_i\rangle$ of the ensemble. Let's now examine the slightly different question: With what probability can an arbitrary state $|\psi_1\rangle$ be included in an ensemble giving the density matrix ρ? The first requirement for a non-zero probability p_1 is that $|\psi_1\rangle$ lies in the span of ρ, since it has to be a superposition of the eigenstates of ρ, in accordance with the mixture theorem. To find an expression for p_1 corresponding to a state $|\psi_1\rangle$ we will once again use (2-16), but this time we take the scalar product of $\sqrt{p_1}\,|\psi_1\rangle$ with $1/\sqrt{\lambda_k}\,|e_k\rangle$. We get the following:

$$\sqrt{p_1}\,\frac{\langle e_k|\psi_1\rangle}{\sqrt{\lambda_k}} = U_{1k} \quad \Rightarrow \quad p_1\,\frac{\langle\psi_1|e_k\rangle\langle e_k|\psi_1\rangle}{\lambda_k} = |U_{1k}|^2 \tag{2-18}$$

$$\Rightarrow \quad p_1\,\langle\psi_1|\left(\sum_{k=1}^{N}\frac{|e_k\rangle\langle e_k|}{\lambda_k}\right)|\psi_1\rangle = \sum_{k=1}^{N}|U_{1k}|^2 \tag{2-19}$$

The operator within parentheses on the left hand side is nothing but the inverse of ρ. Thus

$$p_1 = \frac{1}{\langle\psi_1|\rho^{-1}|\psi_1\rangle}\sum_{k=1}^{N}B_{1k} \le \frac{1}{\langle\psi_1|\rho^{-1}|\psi_1\rangle} \le \langle\psi_1|\rho|\psi_1\rangle\ . \tag{2-20}$$

The first inequality is an equality whenever the number of states M is equal to the dimension N or when the last $M-N$ elements of the first row of B_{1k} are zero. This is when $|\psi_1\rangle$ is linearly independent of the other states included in the ensemble. In this case, these other states do not span the range of ρ and they cannot be used to replace $|\psi_1\rangle$ in the ensemble. The second inequality is an equality if and only if $|\psi_1\rangle$ is an eigenvector of ρ (as can be proved by remembering that ρ is positive with unit trace).

From what is said here about mixtures, it is not easy to get an intuition of how combinations of different pure states can give the same density

** This does not cause any problem in applications to entanglement transformation considered by Nielsen [64].

matrix. This is where the geometry comes in. In what follows we will several times return to this problem. As a first step to gain more insights into the mixed states, we will study convexity—one of the basic properties of this set.

2.3 Convexity

For any two density matrices ρ_1 and ρ_2, we can construct mixtures of them. Take the first with probability p and the second with probability $1-p$. This yields a new density matrix ρ,

$$\rho = p\,\rho_1 + (1-p)\rho_2 \ , \quad 0 < p < 1 \ . \tag{2-21}$$

What we have done here is to take a *convex combination* of two density matrices. Since this is again a density matrix, the density matrices form a convex set; we denote this set $\mathcal{S}$. When we talk of mixtures of quantum states, what is meant is convex combinations. For example, the mixture in equation (2-15) is a convex combination of several pure states.

The *extreme elements* of a convex set are those elements, which cannot be written as a convex combination of any other. Among the density matrices, this is exactly the pure quantum states. (One way to see this is to consider the trace of ρ^2: $\mathrm{Tr}\,\rho^2 = \mathrm{Tr}\,\rho = 1$ for pure states, but for every mixture one gets $\mathrm{Tr}\,\rho^2 < 1$.)

The *convex set* is the *convex hull*, that is, the set of all convex combinations, of the extreme elements.* The extreme elements are part of the boundary of the set, but they do not have to make up the whole boundary. A condition for a density matrix to lie at the boundary $\partial\mathcal{S}$ of the set $\mathcal{S}$ can easily be stated (although not easily checked, unless the dimensions is low). Since the eigenvalues of density matrices are non-negative, the

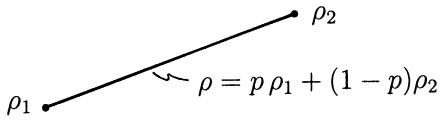


Figure 2-4: *ρ is a convex combination of ρ_1 and ρ_2.*

* This is true for bounded convex sets, which will be our concern here. Unbounded convex sets may contain elements that are not convex combinations of extreme elements.

boundary of $\mathcal{S}$ consists of matrices which have at least one eigenvalue equal to zero and hence have a vanishing determinant,

$$\det \rho_{\partial\mathcal{S}} = 0 \ . \tag{2-22}$$

The natural midpoint of the set of quantum states is the *maximally mixed state*

$$\rho_0 \equiv \frac{1}{N}\mathbb{1} \ . \tag{2-23}$$

It can be obtained as the convex combination of all pure states with equal weight.[†] ρ_0 is also called the matrix of ignorance (and denoted ρ_* in some places in the papers). It is the state such that no outcome of a (non-degenerate) von Neumann measurement is more probable than any other.

A convex set that has only a finite number of extreme elements, is called a *polytope*. We will encounter polytopes later on. In Euclidean space a polytope can be equivalently defined as a bounded intersection of a finite set of half-spaces.[‡] (A half-space is a hyperplane together with every point on one side of the plane). Or differently phrased, and perhaps easier to visualize: the polytope is what is left if one takes away the complements to the defining half-spaces. As an example, a cube in three dimensional space is the intersection of six half-spaces.

The smallest number of extreme elements, whose convex hull is a d-dimensional polytope, is $d+1$. Such a polytope is called a *simplex*, or more precisely a d-simplex. Equivalently a d-simplex is the intersection of $d+1$ half-spaces, which is the smallest number needed to give a d-dimensional polytope. The 2-simplex is a triangle and the 3-simplex is a tetrahedron. Within a simplex every element is a convex combination of the extreme elements in a unique way. This is true only for simplicies.

Simplices turn up whenever one has probability distributions over a finite number of outcomes. Any point in a *probability simplex* corresponds to a probability distribution. The extreme elements correspond to the cases where one outcome has probability one; see figure 2-5. In fact, the set of states of a classical system is a probability simplex. The extreme elements are "pure classical states" and every classical state, that

[†] A general sum over all pure states requires an integration over the unitary group (usually using the Haar measure), but here it is enough to acknowledge that every pure state is part of equally many orthonormal bases; the convex combination of every basis gives ρ_0.

[‡] More generally, every convex set is an intersection of half-spaces, according to the Hahn-Banach separation theorem [4].

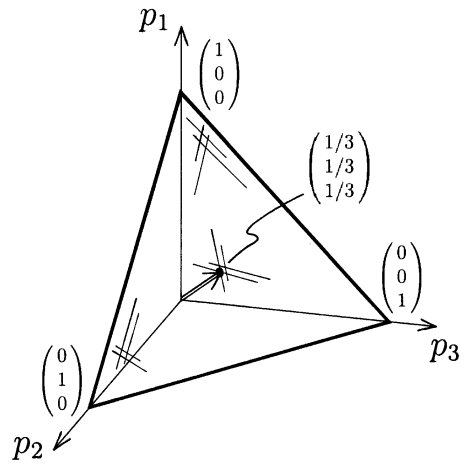


Figure 2-5: *A probability simplex: all probability vectors $\vec{p}$ are convex combinations of the extreme elements.*

is, a point in the simplex, is a unique mixture of the pure states. Therefore, the property of quantum states, that any mixture can be written as a convex combination in several different ways (as described in the last section), is in stark contrast to the classical case.

That any element in a d-simplex is a convex combination of $d+1$ extreme elements is true by definition. However, this is true in any convex set, according to Carathéodory's theorem: Not more than $d+1$ extreme elements are needed in a convex combination to give any element in a convex set of dimension d. The set of quantum states is special, in that so many pure states never are needed to give any convex combination. N pure states are always enough to give any mixture in this set of dimension $d = N^2 - 1$; just choose the eigenvectors. What is a little more surprising is that if we start with any pure state, we can mix it with other pure states, such that a mixture of them all will give any full rank density matrix we might want—and it will always be enough with at most $N-1$ added pure states. This is actually a consequence of the theorem about mixtures. Equation (2-16) shows that any pure state can be part of an ensemble yielding any given full rank density matrix.

We end this section with some additional terminology about convex sets. At the boundary of a convex set there are faces. A *face* is the convex hull of a subset of the extreme points, such that no point in the face can be written as a convex combination including points not in the face. This definition includes the extreme points and the whole set as

faces, although most of the time when one uses this term it is the other "proper" faces one is interested in. For the three dimensional cube, the faces are (besides the whole cube and its corners) the twelve *edges* and the six sides. The faces of dimension one less than the full convex set are called *facets*. These things and some more facts on convex sets can be found in ref. [12].

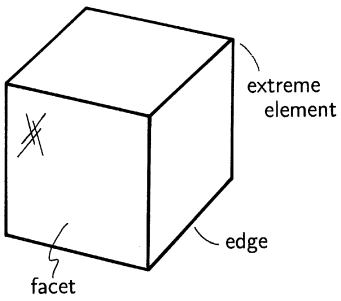


Figure 2-6: *A cube is a polytope with 8 extreme elements, 12 edges and 6 facets.*

Chapter 3

Geometry of the Set of Quantum States

We have seen that the set of quantum states is a convex set. If we introduce a distance between states, we equip the set with a geometrical structure. The set will have a definite "shape" and we can describe both the set and single states in geometrical terms. Two different geometries will be described here.

Furthermore density matrices can be regarded as vectors in a vector space. If the Hilbert space has dimension N, the set of density matrices, has dimension $N^2 - 1$. As soon as $N > 2$ this is so high that it is difficult to see what the full picture is. Nevertheless, it can for many purposes be convenient to view the set of quantum states as a convex set in a vector space.

First we study the Hilbert-Schmidt geometry. It is the simplest case: the set of quantum states is embedded in a flat Euclidean space. The picture we obtain is advantageous, for instance, in understanding the different ways a density matrix can be mixed from pure states. For the qubit this picture is the well known Bloch ball.

Another important geometry is the Bures-Uhlmann geometry. Here the geometry is curved, thus harder to visualize. But now the distances have a physical meaning—they correspond to how well it is possible to distinguish, through measurements, between pairs of quantum states.

3.1 Hilbert-Schmidt geometry

The set $\mathcal{S}$ of all quantum states for an N-level system is the set of non-negative Hermitian operators, with unit trace (equation (2-5)). For any pair of Hermitian operators A and B, define the *Hilbert-Schmidt distance* $D_{\mathcal{HS}}$ to be given by

$$D^2_{\mathcal{HS}}(A,B) \equiv \frac{1}{2}\mathrm{Tr}(A-B)^2 \ , \tag{3-1}$$

The Hermitian matrices then forms a Euclidean space, in which the set of quantum states is a convex subset. We will describe some properties of this set.

We noted earlier (in section 2.3) that $\rho_0 = \frac{1}{N}\mathbb{1}$ is a natural midpoint of the set $\mathcal{S}$. Here we will see more of what this means. The distance between ρ_0 and an arbitrary state ρ depends only on the eigenvalues λ_i of ρ:

$$D^2_{\mathcal{HS}}(\rho_0,\rho) = \frac{1}{2}\sum_{i=1}^{N}\left(\frac{1}{N}-\lambda_i\right)^2 = \frac{1}{2}\left(\sum_{i=1}^{N}\lambda_i^2 - \frac{1}{N}\right) \ . \tag{3-2}$$

The more even (or equal) the eigenvalues are, the closer to the midpoint ρ_0 lies ρ. This statement can be made exact in terms of majorization, as we will see in chapter 4.

The maximal distance from ρ_0 is obtained for states with one eigenvalue equal to one and the rest equal to zero, that is, for the pure states. Therefore all pure states lie on a hypersphere—the outsphere—with radius

$$R_{\text{outsphere}} \equiv D_{\mathcal{HS}}(\rho_0,\rho_{\text{pure}}) = \sqrt{\frac{N-1}{2N}} \ . \tag{3-3}$$

The set $\mathcal{S}$ is the convex hull of the pure states on this outsphere.

The minimal distance to a boundary state is obtained when only one eigenvalue is zero and all the rest are equal. These density matrices at the boundary that are closest to ρ_0 lie then on a hypersphere—the insphere—with radius

$$R_{\text{insphere}} \equiv D_{\mathcal{HS}}(\rho_0,\rho_{\partial\mathcal{S}\text{closest}}) = \sqrt{\frac{1}{2N(N-1)}} \ . \tag{3-4}$$

Every matrix inside this sphere is a density matrix. And the boundary $\partial\mathcal{S}$, consisting of matrices with vanishing determinant (equation (2-22)), lies between the two hyperspheres centered around ρ_0.

The two spheres coincide for dimension $N = 2$, having the same radius $D_{\mathcal{HS}} = 1/2$. This is the Bloch sphere of pure states for a qubit, which enclose all the mixed states. For higher dimensions the outer sphere will increase with N, to a radius $1/\sqrt{2}$ when the dimension goes to infinity, while the inner sphere shrinks, and finally consists only of ρ_0. In terms of the dimension N, the radius of the outsphere is $N-1$ times the radius of the insphere; $R_{\text{outsphere}} = (N-1)R_{\text{insphere}}$. The set of pure states ($\mathbb{C}P^{N-1}$), situated on the outsphere, is $2(N-1)$ dimensional. Thus, whenever $N > 2$, they form only a submanifold of the boundary, which is $N^2 - 2$ dimensional. Most of the outsphere does not correspond to quantum states. And most of the boundary does not correspond to pure states.

For every point in the boundary $\partial\mathcal{S}$ which lies on the outsphere, the point in the boundary in the opposite direction from ρ_0 lies on the insphere. That is, opposite to a pure state is always a "closest state", and vice versa. This is so since ρ_0 is a convex combination of the states $\rho_{\text{pure}} = \text{diag}(1, 0 \dots 0)$ and $\rho_{\partial\mathcal{S}\text{closest}} = \frac{1}{N-1}\text{diag}(0, 1 \dots 1)$. The lines between states such as these are the longest lines through the set of density matrices via ρ_0; the length is $\sqrt{N}/\sqrt{2(N-1)}$. The shortest lines are between states like $\rho_1 = \frac{1}{N}\text{diag}(1 \dots 1, 2, 0)$ and $\rho_2 = \frac{1}{N}\text{diag}(1 \dots 1, 0, 2)$; the length is $1/N$. This can be recognized by inspecting where the longest and shortest lines go through the tetrahedron in figure 3-1. Some more considerations gives that for an arbitrary boundary state ρ at a distance $D_{\mathcal{HS}}$ the distance to the opposite boundary state is $\frac{1}{N\lambda_{\max}-1}D_{\mathcal{HS}}$, where $\lambda_{\max}$ is the largest eigenvalue of ρ. The distances are equal when $\lambda_{\max} = 2/N$.

At every point in the boundary $\partial\mathcal{S}$ there are some directions in which $\partial\mathcal{S}$ is curved and contains a circle [47]. For every pair of one non-zero eigenvalue and one zero eigenvalue a state $\rho_{\partial\mathcal{S}}$ lies on a curve that can be written

$$\rho(\alpha) = \begin{pmatrix} \lambda_1\cos^2\alpha & & & \lambda_1\cos\alpha\ \sin\alpha \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_{N-1} \\ \lambda_1\cos\alpha\ \sin\alpha & & & \lambda_1\sin^2\alpha \end{pmatrix}, \quad 0 \le \alpha < \pi\ . \tag{3-5}$$

This is a circle in the boundary (since $\det\rho(\alpha) = 0$). Moreover, inserting phase factors $e^{\pm i\phi}$ in the off-diagonal elements it becomes a full sphere in the boundary.

There is also at every point in $\partial\mathcal{S}$, except at the pure states, at least

one direction in which the boundary is a straight line [47]. A non-pure boundary state $\rho_{\partial\mathcal{S}}$ have at least two positive eigenvalues, say λ_1 and λ_2. This state lies on a line since, for example,

$$\rho(p) = p \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + (1-p) \begin{pmatrix} 0 & & & \\ & \frac{\lambda_2}{1-\lambda_1} & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} , \qquad (3\text{-}6)$$

gives such a line, with $\rho(\lambda_1) = \rho_{\partial\mathcal{S}}$. A state of rank r can be seen as a convex combination of r projectors $|e_i\rangle\langle e_i|$, with $|e_i\rangle$ eigenvectors of the state. Thus, the boundary is straight in r orthogonal directions.

A special set of pure quantum states, used almost all the time, is an orthogonal basis. How can we characterize it geometrically? The distance between any pair of orthogonal states, for example $|1\rangle = (1, 0 \dots 0)^{\mathrm{T}}$ and $|2\rangle = (0, 1, 0 \dots 0)^{\mathrm{T}}$, is one (as can be found from (3-1)). Three orthogonal states will form an equilateral triangle and four a tetrahedron. The density matrices of the N basis states, sitting at the outsphere at unit distance from each other, form a regular simplex. This simplex spans (is contained within) an $N-1$ dimensional plane in the N^2-1 dimensional set of quantum states. Within the cross-section of this plane, all states sit in the simplex. They are the states diagonal in the basis considered, possible to express as mixtures of the basis states. These states are in some sense

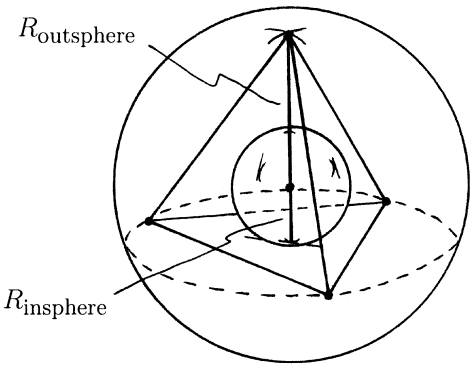


Figure 3-1: *An orthonormal basis forms a simplex in a subspace of Hermitian unit trace matrices; for a 4-level system this is a tetrahedron. The circumscribed and the inscribed spheres of such a simplex coincides with the outpshere and the insphere of the set of quantum states.*

classical. For measurements of any observable with the basis vectors as eigenvectors, the states correspond to probabilities in a probability simplex.

The sphere inscribed in such a simplex coincides (within the cross-section) with the insphere, as should be clear if one realize that it touches states opposite to the pure basis states. Every density matrix, projected to the subspace in question, lies in the simplex. In the center of the simplex is the matrix of ignorance ρ_0. Together with any two of the basis states it gives a triangle (except when $N = 2$), with two sides equal to $\sqrt{(N-1)/2N}$ (the radius of the sphere with pure states; equation (3-3)) and the third side equal to 1. The angle subtended at ρ_0 is decreasing from π for $N = 2$—this gives no triangle, only a straight line—to $\pi/2$ when the dimension goes to infinity.

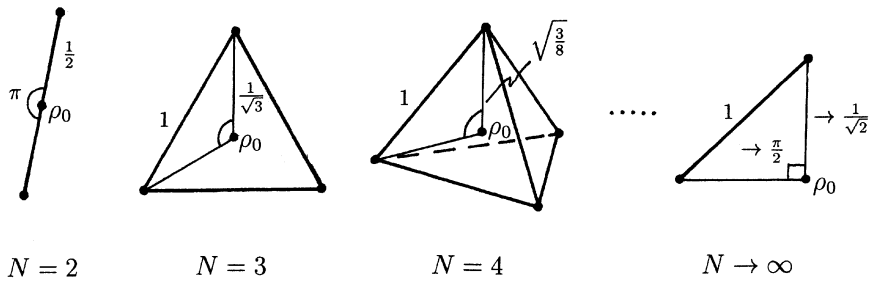


Figure 3-2: *The opening angle at the matrix of ignorance ρ_0 of lines to two orthogonal pure states decreases from π to $\frac{\pi}{2}$ when the dimension increases.*

We can also say something more about the set of states orthogonal to some basis states. If we choose all but two basis states, they will be orthogonal to all superpositions, and all mixtures thereof, of the two remaining basis states—this is like a Bloch ball with its pure states at unit distance from the first $N-2$ basis states. If we choose one basis state, it will be orthogonal to all superpositions, and all mixtures thereof, of the other basis states—this is like the set of states of an $N-1$ dimensional system. More generally in any dimension N: to n basis states, there is a set of orthogonal states, equivalent to the set of states of an $N-n$ dimensional system. The pure orthogonal states all lie at unit distance from the n first basis states. Conversely the first n basis states generate themselves a set equivalent to the set of an n dimensional system. The centers of these two orthogonal sets lie "opposite" each other, with the maximally mixed state in between them; this is illustrated in figure 3-3.

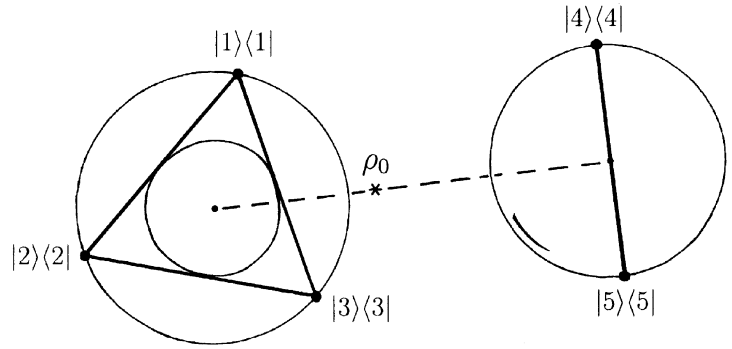


Figure 3-3: *A sketch of the set of quantum states in dimension 5. Two basis states generate a Bloch ball at the boundary, which is orthogonal to three basis states generating a set equivalent to the set of states for a 3 dimensional system.*

Without explicitly saying it, a vector space has been introduced above. We have a distance, $D_{\mathcal{HS}}(A, B)$ (equation (3-1)), and an origin, ρ_0 (equation (2-23)). Thereby we get, using the polarization formula, the *scalar product*

$$(A, B) = \frac{1}{4}\left[D^2(A + B, \rho_0) - D^2(A - B, \rho_0)\right] = \frac{1}{2}\Big[\mathrm{Tr}\, AB - \frac{1}{N}\Big] \ . \quad (3\text{-}7)$$

This expression might seem a bit strange, if you have seen the commonly used scalar product $\mathrm{Tr}AB$ of Hermitian matrices A and B. The difference is that we have ρ_0 as our origin, instead of the zero matrix. When studying quantum states we are only concerned with unit trace matrices. In the vector space with $\mathrm{Tr}AB$ as the scalar product, these lie in a hyperplane that does not contain the origin, and hence is not a subspace. But with (3-7) as the scalar product for Hermitian matrices, the quantum states will be situated in a subspace. Our studies take place within this subspace of $N^2 - 1$ dimensions. An equivalent way to put it is to use the simpler form $\mathrm{Tr}AB$, but at the same time represent any density matrix ρ with the traceless matrix $\sigma = \rho - \rho_0$ instead.

This vector space with the convex set of density matrices, as described in this section, is the scene for what is done in Papers I, II, IV and partly in VI. In Paper III the set of bistochastic matrices is seen as a convex set in a similar vector space.

We will look at a couple of examples of sets of density matrices, but first we get back to the issue concerning mixing density matrices. With the Hilbert-Schmidt geometry, a mass distribution with its center of mass

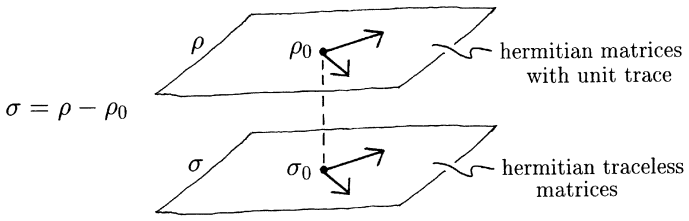


Figure 3-4: *With the scalar product* TrAB *traceless matrices* σ *form a subspace. This is not so for unit trace matrices* ρ*. These form a subspace if the scalar product* $\frac{1}{2}(\mathrm{Tr}AB - \frac{1}{N})$ *is used.*

gives a nice analogue of an ensemble giving a mixed state. An ensemble $\{(|\psi_k\rangle, p_k)\}_{k=1}^{M}$ gives the mixed state,

$$\rho = \sum_{k=1}^{M} p_k \, |\psi_k\rangle\langle\psi_k| = \sum_{k=1}^{M} p_k \, \rho_k \ . \tag{3-8}$$

The density matrices are the position vectors of the states, and from this formula we see that in this vector space, every ensemble $\{|\psi_k\rangle, p_k\}$ can be thought of as a mass distribution, with the corresponding density matrix at the center of mass. Masses p_k should be placed at appropriate positions on that part of the outsphere which consists of pure states $|\psi_k\rangle\langle\psi_k|$, so that it gives the right center of mass, that is, ρ . Figure 3-6, in the next subsection, show examples of this in the qubit case,

3.1.1 The Bloch ball

An overwhelming majority of all quantum information experiments deal with two-level systems—*qubits*—as its basic quantum systems. And almost all algorithms suggested for quantum computing, and other quantum information processes, are based on qubits. To get interesting results, several qubits are needed, but at the end one typically measures single qubits. Thus, it is worthwhile to have a closer look at the set of possible states for a qubit. However, this is not the only reason why this set, known as the Bloch ball, is quite familiar to most physicists. It is exceptional in several regards. It is three dimensional, thus easy for us to picture, and moreover just a round ball, and the whole boundary consists of pure states. Furthermore, if one considers spin$\frac{1}{2}$-particles the three dimensions can be thought of as our three space dimensions. Or rather,

the points on the Bloch sphere (that is, the boundary of the Bloch ball) corresponds to directions in space. For example, consider silver atoms in a Stern-Gerlach experiment [40, 70]. The internal angular momentum of an atom in a pure state can be regarded as being in the direction corresponding to the point representing the state on the Bloch sphere. Thus if the Stern-Gerlach magnet is oriented along that direction the outcome is certain to be "up" (and this is not the case for any other direction).

The pure states of a qubit lie on a sphere with radius 1/2, and conversely, every point on that sphere corresponds to a pure state ($\mathbb{C}\mathrm{P}^1$ is a sphere). Since there are only two eigenvalues, all boundary states are pure. Inside the sphere are the mixed states. Any orthogonal basis will correspond to two antipodal points, at unit distance from each other.*

A nice expression for a general 2×2 density matrix is given by

$$\rho = \frac{1}{2}\mathbb{1} + \vec{n} \cdot \vec{\sigma} \,, \tag{3-9}$$

where $\vec{n}$ is a vector in the *Bloch ball*, giving the position of the matrix, and $\vec{\sigma}$ is a vector with the Pauli matrices as its element. The eigenvalues of ρ are $\lambda = \frac{1}{2} \pm |\vec{n}|$. We see that $|\vec{n}| \leq 1/2$ for non-negative eigenvalues, with equality for pure states.† If we consider a spin $\frac{1}{2}$ system, a density matrix given by a vector $\vec{n}$ on the sphere, is simply the state of spin up in the direction of $\vec{n}$.

Note that equation (3-9) is an expansion of the matrix ρ in an ON-basis in the space of Hermitian matrices. $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ constitutes an ON-basis and $(\frac{1}{2}, n_x, n_y, n_z)$ are the expansion coefficients.

Special to two dimensions is that there exists a "universal NOT" operation—an operation that take any pure state to an orthogonal state. It has been proven that this is only possible for qubit states [84]. The orthogonal state is the antipodal point on the Bloch sphere, and hence the universal NOT implements an inversion of the Bloch ball. This is given by an antiunitary operator Θ. In any basis, to apply the NOT operation on a state vector $(a, b)^{\mathrm{T}}$, take the complex conjugate and then multiply with an antisymmetric matrix with unit determinant; the result is the wanted state vector $(b^*, -a^*)^{\mathrm{T}}$ (perhaps with some phase factor).

---

* It is interesting that Gleason's theorem, which ascertains that states should be represented by density matrices, given some reasonable requirements on non-contextuality of outcome probabilities, does not hold in two dimensions [42]. This is because there is only one pure state orthogonal to any given pure state.

† Often a factor of one-half is included before $\vec{n}$ in equation (3-9), so that the maximal length of the "Bloch vector" is one.

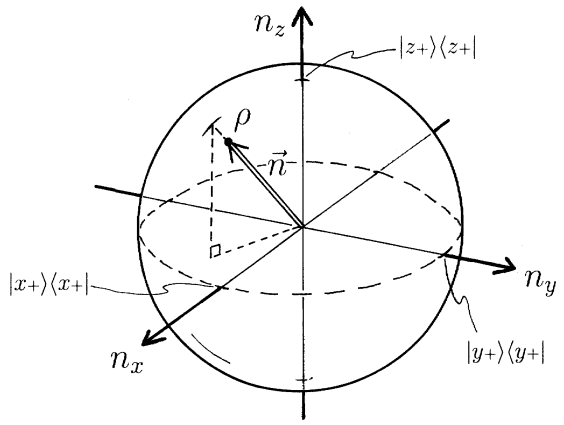


Figure 3-5: *Every qubit state corresponds to a point in the Bloch ball, conveniently represented by a Bloch vector $\vec{n}$.*

There is also a noteworthy property of mixtures, special to qubits. Starting with an arbitrary pure state, there is always another pure state, such that a mixture of them will give any (non-pure) density matrix we might ask for. This is evident with the sphere in mind: a chord drawn from the first given pure state, to the wanted density matrix, will end at a second pure state. This is not true when $N > 2$, since in higher dimensions most of the boundary consists of non-pure states.

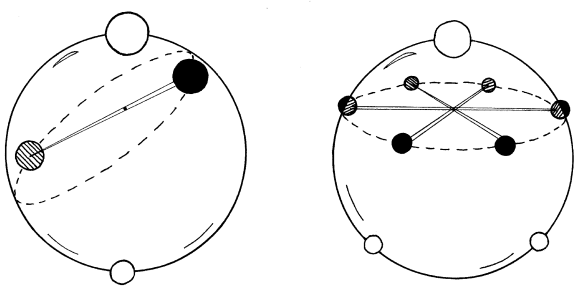


Figure 3-6: *Four ensembles giving the same mixed state; one "black" and one "white" ensemble are shown on each Bloch sphere.*

3.1.2 The states of a qutrit

A three-level quantum system is sometimes called a qutrit. I will briefly note some properties of the set of qutrit states. This might be regarded as an example which summarizes properties discussed above.

The set of all qutrit states is 8-dimensional. The pure states forms a 4-dimensional subset on a 7-dimensional hypersphere with radius $1/\sqrt{3}$. The rest of the boundary lies somewhere between this outsphere and an insphere with radius $1/(2\sqrt{3})$. On this insphere the boundary matrices have the eigenvalues $\lambda = 0,\ 1/2,\ 1/2$.

The three points corresponding to an ON-basis in Hilbert space $\mathcal{H}^3$ are vertices of a triangle with unit side lengths. This triangle is inscribed in the outsphere and has an inscribed sphere with the same radius as the insphere. In such a subspace of an ON-basis all states are diagonal in that basis and lie within the triangle.

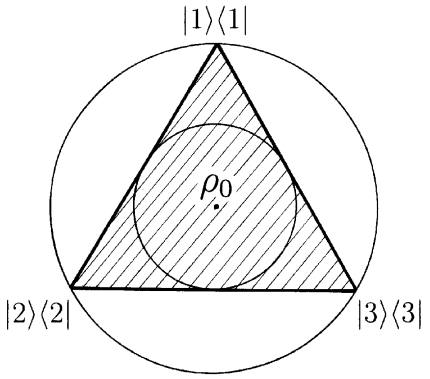


Figure 3-7: *A two dimensional cross section of the set of qutrit states, spanned by an ON-basis.*

The first state of a basis is orthogonal to any superpositions of the two other basis states. These two states generate themselves a Bloch ball of states. Thus, to any pure qutrit state there is a whole set, equivalent to a Bloch ball, orthogonal to it. Figure 3-8 show a possible three dimensional cross-section through the set, which includes first a basis, and also one great circle of the Bloch ball generated by two of the basis states.

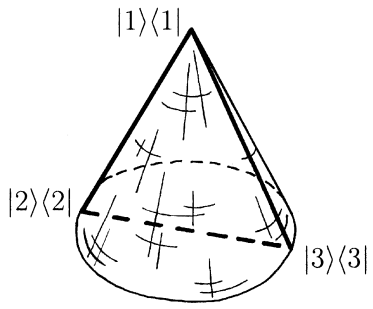


Figure 3-8: *Two basis states generate a Bloch ball orthogonal to the third basis state. A disc of the Bloch ball together with the orthogonal state gives a cone of states.*

3.2 Bures-Uhlmann geometry

As opposed to the nice, flat Hilbert-Schmidt geometry, the curved, unwieldy Bures-Uhlmann geometry has an operational meaning.* It is closely related to distinguishability of quantum states, as will be touched upon later. But this was not clear when it was introduced. Instead it begun with a generalization of "transition probabilities" for state vectors, $|\langle\psi|\phi\rangle|^2$, to mixed states. This was done in terms of "purifications", and this is where we will begin this survey. What will be explained here is based upon work by Uhlmann [79, 80, 81].†

Consider a mixed state ρ—an operator on the Hilbert space $\mathcal{H}$. We want to "purify" ρ. That is, we want to find a pure state $|\Psi\rangle$ for a composite system, such that the reduced state for one of the subsystems is ρ. This is always possible with some state vector $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}$. Let $|\Psi\rangle$ be written in the Schmidt decomposition form (equation(2-11)),

$$|\Psi\rangle = \sum_i c_i |\, i\rangle_A \otimes |\mu_i\rangle_B \,. \tag{3-10}$$

From the reduced state

$$\rho = \mathrm{Tr}_B |\Psi\rangle\langle\Psi| = \sum_i c_i^2 \,|\, i\rangle_{A\ A}\langle\, i| \,, \tag{3-11}$$

* At least the infinitesimal distances.

† My understanding of the Bures-Uhlmann geometry has also benefitted from lecture notes by Uhlmann [82], and ref. [12].

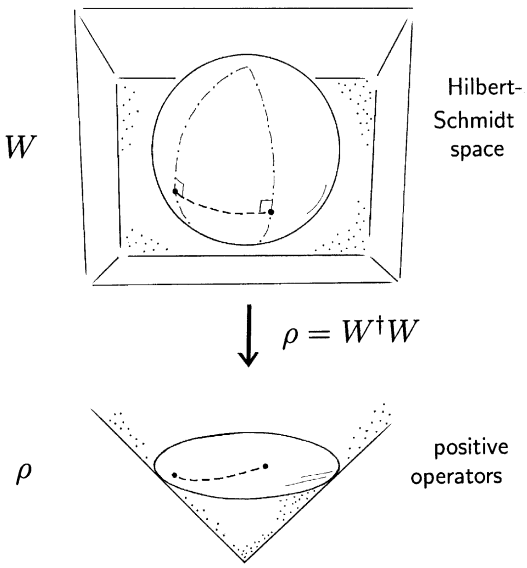


Figure 3-9: *Purifications (fibres) on the unit sphere in Hilbert-Schmidt space are projected to mixed states (unit trace positive matrices).*

we see that $|\Psi\rangle$ is a "purification" of ρ whenever the basis states $|\,i\rangle_A$ are the eigenvectors of ρ, and the Schmidt coefficients c_i are the square roots of the eigenvalues. The basis $\{|\mu_i\rangle_B\}$ could be replaced by any basis, $\{U|\mu_i\rangle_B\}$, for some unitary operator U, thus giving a set of purifications as large as the set of all unitaries on $\mathcal{H}$.

Now, let's make use of a small trick. Exchange all "ket"-vectors $|\,j\rangle_B$ with "bra"-vectors ${}_B\langle\, j|$. This turns every pure state $|\Psi\rangle = \sum_{i,j} c_{ij}|\,i\rangle_A|\,j\rangle_B$ into an operator $W = \sum_{i,j} c_{ij}|\,i\rangle_A\,{}_B\langle\, j|$ on $\mathcal{H}$. We call W a purification of the reduced state ρ, now given by

$$\rho = WW^\dagger\ . \tag{3-12}$$

In this formulation we again see that whenever W is a purification, so is WU, for any unitary U. What we have here is a kind of fibre bundle construction (except that not all fibres are isomorphic). Equation (3-12) can be understood as a projection from a bundle space of operators W—the Hilbert-Schmidt space—to a base manifold of positive operators ρ. Right multiplication with the unitary group gives the fibres WU. Normalization is not needed for this to work, but we are only interested in normalized quantum states, thus we assume $\mathrm{Tr}\, WW^\dagger = 1$.

In the bundle space we define distances D by

$$D^2(W_1, W_2) \equiv \mathrm{Tr}(W_1 - W_2)(W_1^\dagger - W_2^\dagger) \; ; \tag{3-13}$$

this coincides, for Hermitian operators, with the Hilbert-Schmidt distance $D_{\mathcal{HS}}$ (equation (3-1)) we had between density matrices, except for a factor of one half.

Distances between operators ρ_1 and ρ_2 in the base manifold are defined as the length of the shortest path between the corresponding fibres of purifications in the bundle space. And this is the *Bures distance* D_B:[‡]

$$D_B^2(\rho_1, \rho_2) \equiv \min_{U_1, U_2} D^2(W_1 U_1, W_2 U_2) = \tag{3-14}$$

$$\mathrm{Tr}\,\rho_1 + \mathrm{Tr}\,\rho_2 - \max_U \mathrm{Tr}(W_1 U^\dagger W_2^\dagger + W_2 U W_1^\dagger) = 2 - 2 \max_U \mathrm{Tr}(W_1^\dagger W_2 U) \; .$$

It is enough to find the extremum over only one unitary, $U = U_2 U_1^\dagger$. The extremum is attained if U is chosen so that the eigenvalues of $W_1^\dagger W_2 U$ are real and positive. To get an expression of the distance in terms of the density matrices ρ_1 and ρ_2, use the polar decomposition $W_1 = \sqrt{\rho_1}\, V$ (some unitary V) and look at the square

$$(W_1^\dagger W_2 U)^2 = W_1^\dagger W_2 U\, U^\dagger W_2^\dagger W_1 = V^\dagger \sqrt{\rho_1}\, \rho_2 \sqrt{\rho_1}\, V \; . \tag{3-15}$$

If we take the square root of this, and then take the trace we get

$$\max_U \; \mathrm{Tr}\,(W_1^\dagger W_2 U) = \mathrm{Tr} \sqrt{\sqrt{\rho_1}\, \rho_2 \sqrt{\rho_1}} \; . \tag{3-16}$$

The square of this trace is the generalized "transition probability" defined by Uhlmann [79], perhaps more generally known as the *fidelty* F, as it was named by Jozsa [55]. In terms of purifying state vectors, it can be stated as follows:

$$F(\rho_1, \rho_2) \equiv \max_{\textit{purifications}} |\langle \Psi_1 | \Psi_2 \rangle|^2 = \left(\mathrm{Tr} \sqrt{\sqrt{\rho_1}\, \rho_2 \sqrt{\rho_1}} \right)^2 \; . \tag{3-17}$$

The Bures distance is thus given by

$$D_B^2(\rho_1, \rho_2) = 2 \left(1 - \mathrm{Tr} \sqrt{\sqrt{\rho_1}\, \rho_2 \sqrt{\rho_1}} \right) = 2 \left(1 - \sqrt{F(\rho_1, \rho_2)} \right) \; . \tag{3-18}$$

[‡] This distance has been defined by Bures in a more general setting [24]. The corresponding metric is often called Bures metric, but the infinitesimal distances are due to Uhlmann, and it was Uhlmann who first used it for quantum states.

This is an explicit formula for any pair of quantum states. But it is not trivial to actually compute this distance—to find square roots of operators one needs to diagonalize them.

The distance (3-18) between density matrices corresponds to the length of a straight line between a special pair of purifications in the Hilbert-Schmidt space of operators. However, we want to study only normalized states, $\mathrm{Tr}\, WW^\dagger = 1$, which means a restriction to the unit sphere in the Hilbert-Schmidt space. Therefore it makes sense to use the arc length between the operators instead. The minimal distance D_B then corresponds to a minimal arc length d_B, called the *Bures angle*. With the help of figure 3-10 we readily find that the cosine of the angle is simply the square root of the fidelity:

$$\cos d_B(\rho_1, \rho_2) = \sqrt{F(\rho_1, \rho_2)}\ . \tag{3-19}$$

Let's see what this is in those cases where the square root operators are easily computed. First, let one of the states be the maximally mixed state ρ_0:

$$\cos d_B(\rho, \rho_0) = \frac{1}{\sqrt{N}}\,\mathrm{Tr}\sqrt{\rho} = \frac{1}{\sqrt{N}}\sum_i \sqrt{\lambda_i}\ . \tag{3-20}$$

Secondly, let one of the states be a pure state $\rho_{\mathrm{pure}} = |\psi\rangle\langle\psi|$:

$$\cos d_B(\rho, \rho_{\mathrm{pure}}) = \sqrt{\langle\psi|\rho|\psi\rangle}\ . \tag{3-21}$$

From anyone of these two formulas we find the distance from the center ρ_0 to the pure states to be given by $\cos d_B(\rho_0, \rho_{\mathrm{pure}}) = 1/\sqrt{N}$. Thus, the distance d_B is $\pi/4$ for qubits and increases with the dimension N to $\pi/2$ when N goes to infinity. For boundary states closest to the center (3-20)

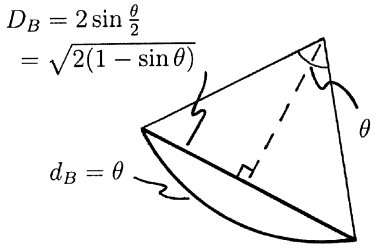


Figure 3-10: *The relation between Bures D_B distance and Bures angle d_B.*

gives $\cos d_B(\rho_0, \rho_{\partial\mathcal{S}\text{closest}}) = \sqrt{1-1/N}$. This distance is $\pi/4$ for qubits and goes to zero when the dimension goes to infinity.

Next we will look at the restrictions to only pure states and to only diagonal states. We give these cases sections by themselves.

3.2.1 Pure states and the Fubini-Study geometry

For a pair of pure states we obtain

$$\cos d_B\Big(|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|\Big) = |\langle\psi_1|\psi_2\rangle| \ . \tag{3-22}$$

The distance $\arccos|\langle\psi_1|\psi_2\rangle|$ is called the *Fubini-Study distance*, and to emphasize that we here only consider pure states we denote it $d_{\mathcal{FS}}$. It can be defined in terms of vectors in the Hilbert space $\mathcal{H}$ as described below. The construction is equivalent to the one of Bures distance, if we confine ourselves to pure states. The Fubini-Study distance [38, 75] predates Bures distance, which is a generalization thereof. First define

$$D_{\mathcal{FS}}\Big(|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|\Big) \equiv \min_{\phi} \| \, |\psi_1\rangle - e^{i\phi}|\psi_2\rangle \, \| \ ,^{\S} \tag{3-23}$$

where the norm is given by $\| \, |\psi\rangle \, \| = \sqrt{\langle\psi|\psi\rangle}$. This is a distance in the set of all "vectors modulo phases", that is, rank one operators $|\psi\rangle\langle\psi|$. The phase ϕ is varied until the angle in Hilbert space between the two vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ is as small as possible. Evaluated for two normalized vectors it becomes

$$D_{\mathcal{FS}}\Big(|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|\Big) = \sqrt{2-2|\langle\psi_1|\psi_2\rangle|} \ . \tag{3-24}$$

This distance is the length of a curve including non-normalized vectors. If we wish to measure distances within the normalized vectors, we want the corresponding smallest angle, which is $\arccos|\langle\psi_1|\psi_2\rangle|$ (see again figure 3-10). Once more we arrive at the expression (3-22). This makes it evident that the path with the length given by (3-22) lies entirely within the set of pure states. In differential geometry one would say that, with respect to the Bures-Uhlmann metric, the set of pure states is a totally geodesic submanifold.

One more remark about this. Consider the Hilbert-Schmidt distance (3-1) (for operators on the non-extended Hilbert space):

$$D^2_{\mathcal{HS}} = \tfrac{1}{2}\mathrm{Tr}\Big(|\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2|\Big)^2 = 1 - |\langle\psi_1|\psi_2\rangle|^2 = 1 - \cos^2 d_{\mathcal{FS}} \ . \tag{3-25}$$

[§] Sometimes it is this distance that is called the Fubini-Study distance.

If we let $|\psi_1\rangle$ and $|\psi_2\rangle$ be close, and expand to second order in $d_{\mathcal{FS}}$ we obtain

$$D^2_{\mathcal{HS}} \approx 1 - (1 - \tfrac{1}{2} d^2_{\mathcal{FS}})^2 \approx d^2_{\mathcal{FS}} \tag{3-26}$$

—for infinitesimal distances $D_{\mathcal{HS}}$ coincides with $d_{\mathcal{FS}}$. Thus, if we use Hilbert-Schmidt geometry to measure the length along curves within the pure states we get the same as the Fubini-Study distances, which are the distances for pure states according to the Bures-Uhlmann geometry. But there is a significant difference: with the Bures-Uhlmann geometry there are no shorter paths between pure states, whereas the geodesics according to the Hilbert-Schmidt geometry take a short cut through the mixed states.

3.2.2 Commuting states in hyperoctants

Commuting states, diagonal in the same basis, lie according to the Hilbert-Schmidt geometry in a regular simplex (see section 3.1). What will this set of states be like according to the Bures-Uhlmann geometry? For two commuting states $\rho_1 = \mathrm{diag}(p_1, \ldots, p_N)$ and $\rho_2 = \mathrm{diag}(q_1, \ldots, q_N)$, the distance is given by (from (3-19))

$$\cos d_B(\rho_1, \rho_2) = \sum_{i=1}^{N} \sqrt{q_i}\sqrt{p_i} \ . \tag{3-27}$$

Define vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^N$, by $x_i = \sqrt{p_i}$ and $y_i = \sqrt{q_i}$. Then $\vec{x}$ and $\vec{y}$ are positive vectors with unit length, and $\cos d_B(\rho_1, \rho_2) = \vec{y} \cdot \vec{x}$. Thus, the distance d_B is the angle between the vectors $\vec{x}$ and $\vec{y}$, which is the same as the geodesic distance along the unit ($N-1$ dimensional) sphere. We see that for these states the Bures-Uhlmann geometry is locally that of a sphere. The simplex of commuting states is deformed into a hyperoctant of a sphere [51]. It is the hyperoctant since the diagonal elements p_i are mapped to positive components x_i; see figure 3-11.

The distance between any two basis states, at "vertices" of the hyperoctant, is $\pi/2$. This is the maximal possible distance.

3.2.3 Bures-Uhlmann geodesics

We will now have a look at the geodesics; these are used in paper V [36]. To any geodesic between a pair of density matrices there is a corresponding preimage in the Hilbert-Schmidt space, having the same length. This, too, is a geodesic, and since we are on the unit sphere it has to be (a

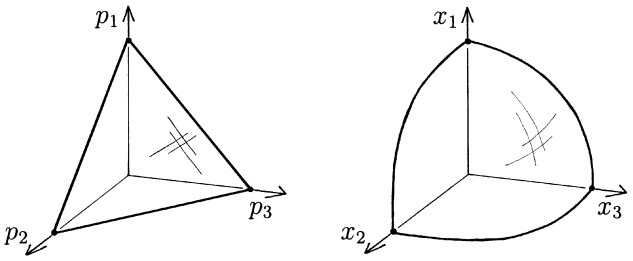


Figure 3-11: *The set of commuting 3-level quantum states: a flat simplex in Hilbert-Schmidt geometry turns into a round hyperoctant in the Bures-Uhlmann geometry.*

part of) a great circle. Moreover we have to require that this curve is everywhere perpendicular to the fibres, since it should be the shortest path. Imagining the fibres to be vertical, the curve must be horizontal. If we choose a representative W' (to determine the vertical height of the preimage) for one point ρ' on the geodesic, with ρ' invertible, these considerations will lead to a unique preimage, whose projection onto the density matrices gives the geodesic.

Let's turn this into formulas. We want an expression for the geodesic $\rho(\tau)$ connecting ρ_1 and ρ_2 . To get this we start by constructing a general formula for a geodesic $\rho(\tau)$, and then we require $\rho(\tau)$ to equal ρ_1 and ρ_2 for two values of τ.

- A geodesic on the unit sphere in the Hilbert-Schmidt space—a great circle—is given by

$$W(\tau) = W(0)\cos\tau + \dot{W}(0)\sin\tau \ , \tag{3-28}$$

where the operators $W(0)$ and $\dot{W}(0)$ fulfill the following conditions:

$$\begin{aligned} &\text{(i)} && \operatorname{Tr} W(0)W^\dagger(0) = 1 \ , \\ &\text{(ii)} && \operatorname{Tr} \dot{W}(0)\dot{W}^\dagger(0) = 1 \ , \\ &\text{(iii)} && \operatorname{Tr}\left[W(0)\dot{W}^\dagger(0) + W^\dagger(0)\dot{W}(0)\right] = 0 \ . \end{aligned} \tag{3-29}$$

The last condition is that $W(0)$ and $\dot{W}(0)$ should be orthogonal. It is obtained using the scalar product corresponding to the distance in (3-13) (and the zero matrix as origin).

The parameter τ is nothing but the distance from $W(0)$.

- The condition for "horizontality" reads

$$\dot{W}^\dagger(0)W(0) = W^\dagger(0)\dot{W}(0) \ . \tag{3-30}$$

  Every vertical vector, parallell to the fibre $W(0)U$, is the derivative of some curve $W(0)U(\sigma) = W(0)e^{iH\sigma}$ along the fibre. Here H is Hermitian. Thus $\dot{W}(0)$ has to be orthogonal to $iW(0)H$, for any H—this gives (3-30).

  Together with (iii) it follows that $\mathrm{Tr}\, W^\dagger(0)\dot{W}(0) = 0$.

Note that if the conditions (3-29) and (3-30) are satisfied for any point on the geodesic, they are satisfied for every point.

- The geodesic in the set of density matrices is

$$\rho(\tau) = W(\tau)W^\dagger(\tau) = W(0)W(0)^\dagger \cos^2\tau + \dot{W}(0)\dot{W}(0)^\dagger \sin^2\tau \ . \tag{3-31}$$

- Start the geodesic from ρ_1 (assumed to be invertible) and let it go through ρ_2. The distance between them is the Bures angle d_B. This means that we set

$$\begin{cases} W_1 = W(0) \ , \\ W_2 = W(d_B) \ , \quad \text{where} \quad \cos d_B = \mathrm{Tr}\sqrt{\sqrt{\rho_1}\,\rho_2\sqrt{\rho_1}} \ . \end{cases} \tag{3-32}$$

- The geodesic (3-28) becomes

$$W(\tau) = W_1\cos\tau + (W_2 - W_1\cos d_B)\frac{\sin\tau}{\sin d_B} \ , \tag{3-33}$$

  if one solves for $W(0)$ and $\dot{W}(0)$.

This gives an expression for the geodesic $\rho(\tau)$ in terms of the purifications W_1 and W_2. But these are not independent. If W_1 is the chosen representative for the fiber of purifications that maps to ρ_1 then W_2 will be determined from the horizontality condition (3-30).

Instead of finding out directly what W_2 should be, we will see how the geodesic can be expressed in terms of ρ_1 and a matrix M which can be thought of as determining the direction from ρ_A. M is such that

$$W_2 = MW_1 \ . \tag{3-34}$$

We obtain, from (3-33),

$$\rho(\tau) = \left[\mathbb{1}\cos\tau + (M - \mathbb{1}\cos d_B)\tfrac{\sin\tau}{\sin d_B}\right]\rho_1\left[\mathbb{1}\cos\tau + (M - \mathbb{1}\cos d_B)\tfrac{\sin\tau}{\sin d_B}\right]. \tag{3-35}$$

This will be our final formula for the geodesic, except that we need to determine d_B in terms of ρ_1 and M. We will find M in terms of ρ_1 and ρ_2, and from this the distance d_B in terms of ρ_1 and M.

From the "horizontality condition" (3-30) it follows that $W_1^\dagger W_2 = W_2^\dagger W_1$, for all W_1 and W_2 on the geodesic in the Hilbert Schmidt space. Thus $W_1^\dagger W_2$ is an Hermitian operator, having real eigenvalues. Therefore, since $W_1^\dagger W_1 = \rho_1$ is a positive operator, also $W_1^\dagger W_2$ is positive if W_2 is sufficiently close to W_1. If the points are continously separated we have

$$W_1^\dagger W_2 > 0 \tag{3-36}$$

until one of the operators will have one zero eigenvalue. This is when either ρ_1 or ρ_2 is at the boundary of the set of density matrices.

Since $W_1^\dagger W_2 = W_1^\dagger M W_1$ is positive, so is M. And from $W_2 = M W_1$ we get:

$$\begin{aligned} \rho_2 = M\rho_1 M \quad &\Leftrightarrow \quad \sqrt{\rho_1}\,\rho_2\sqrt{\rho_1} = \sqrt{\rho_1}\,M\rho_1 M\sqrt{\rho_1} \\ &\Leftrightarrow \quad \sqrt{\rho_1}\,\rho_2\sqrt{\rho_1} = (\sqrt{\rho_1}\,M\sqrt{\rho_1})^2 \end{aligned} \tag{3-37}$$

Together with the positivity requirement this yields

$$M = \rho_1^{-1/2}\sqrt{\rho_1^{1/2}\rho_2\,\rho_1^{1/2}}\,\rho_1^{-1/2}\ . \tag{3-38}$$

Note that such an operator has a name: it is the geometeric mean of the operators ρ_1^{-1} and ρ_2 [3]. The distance d_B in (3-35) is given by

$$\cos d_B = \mathrm{Tr}\, M\rho_1\ . \tag{3-39}$$

With this, the formula (3-35) for $\rho(\tau)$ is complete. If we want we can also rewrite it in terms of the density matrices ρ_1 and ρ_2 using (3-38). (This has been done also by Barnum [6], as I found out after the completion of Paper V.)

In Paper V [36] it is explained how any geodesic in the Hilbert-Schmidt space—a great circle—when projected to the set of density matrices will reach the boundary N times (counting degeneracies), where N is the dimension of the Hilbert space [81]. This means that the geodesic will be reflected at the boundary and 'bounce' back into the set. After N

bounces it will close itself. For commuting matrices we saw in the last section that the geometry is a round hyperoctant. The cover illustration shows what it looks like when $N = 3$. Normalized diagonal operators W in the Hilbert-Schmidt space form a two-sphere, on which a great circle in general will pass through six hyperoctants. Through the projection $W \to \rho = WW^\dagger$ all these octants will be mapped to the same positive octant and the great circle will give a triangle, which will be covered twice. (This positive octant is the selfsame as in figure 3-11).

3.2.4 Distinguishability of quantum states

A pair, or a larger set, of quantum states that have orthogonal support ($\mathrm{Tr}\,\rho_k\rho_l = 0$) are "one shot distinguishable": there is a measurement such that any of the states will with certainty result in a unique outcome and no two state will give the same outcome. This possibility to distinguish such states is reflected in the Bures-Uhlmann geometry. The Bures distance between two states is maximal—equal $\pi/2$—if and only if the states are one shot distinguishable. In figure 3-11 it is evident that the state at the top of the simplex is, in the curved Bures-Ulhmann case, at the same maximal distance from anyone of its orthogonal states lying on the 'line' between the other two vertices of the simplex. This is not so for the flat Hilbert-Schmidt case.¶

This suggests that the Bures-Uhlmann geometry has a physical meaning the Hilbert-Schmidt geometry does not have. In fact, for diagonal states, the situation is the same as for classical probability distributions and the Bures-Uhlmann metric coincides with the Fisher-Rao metric [37, 69]. The Fisher-Rao metric corresponds to a "statistical distance", which measures how well probability distributions can be distinguished. Wooters gives a readable account on statistical distance [91]. He also defines a statistical distance for pure quantum states, a definition later extended to mixed states by Fuchs and Caves [39] (see also Paper V). They showed that this distance is equivalent to Bures distance, thereby giving an operational definition for the Bures-Uhlmann geometry.

The definition of statistical distance for quantum states recognizes that the possibility to distinguish between two states depends on what measurement one performs. Fuchs and Caves found what the best measurement is for distinguishing between two states. This best measurement is actually the observable given by the operator M in equation (3-38). In

¶ For two pure orthogonal states the distance is maximal, $D_{\mathcal{HS}} = 1$, but mixed orthogonal states in dimension N can be as close as the distance $D_{\mathcal{HS}} = \sqrt{2/N}$.

Paper V [36] I give a proof of how this measurement is fully determined by the geodesic going through the two states. In fact, it is the 'bouncing' of the geodesic (mentioned in the last section)—itself a consequence of the projection of geodesics in the Hilbert-Schmidt space of purifications—that determines the best measurement.

We end these sketchy notes on distinguishability with an illustration of Bures-Uhlmann geometry for qubits.

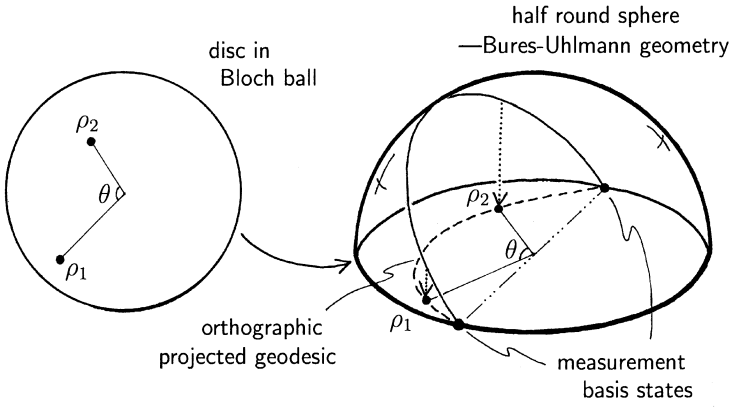


Figure 3-12: *A flat disc in the Bloch ball (with Hilbert-Schmidt geometry) will be half a sphere according to the Bures-Uhlmann geometry [51]. The figure show how to find what a Bures-Uhlmann geodesic through two density matrices ρ_1 and ρ_2 looks like in the Bloch ball. The geodesic meets the boundary in antipodal points, corresponding to the eigenvectors of the best measurement for distinguishing the two states.*[‖]

[‖] Note that the characterization of the best measurement is a bit more involved in higher dimensions; see Paper V [36].

Chapter 4

Majorization and Bistochastic Matrices

Bistochastic matrices map probability vectors to new probability vectors. We have seen an example of this: As a consequence of Schrödinger's mixture theorem a bistochastic matrix turned up which relates any possible probability distribution for mixing pure states to the eigenvalues of the density matrix: $p_i = \sum_j B_{ij}\,\lambda_j$ (equation (2-17)). This is just one context among many where bistochastic matrices—also called doubly stochastic—show up. Paper III [11] is devoted to this kind of matrices. Here we will see what majorization is—it is closely related to bistochastic matrices—and we will discuss Birkhoff's polytope, which is the set of all bistochastic matrices.

An $N \times N$ matrix B is said to be *bistochastic* if its matrix elements obey

$$\begin{aligned} &(i) && B_{ij} \geq 0 \ , \\ &(ii) && \textstyle\sum_i B_{ij} = 1 \ , \\ &(iii) && \textstyle\sum_j B_{ij} = 1 \ . \end{aligned} \tag{4-1}$$

A matrix fulfilling conditions (i) and (ii)—positivity-preserving and trace-preserving—is called stochastic. These are necessary and sufficient conditions for a matrix to map any probability vector $\vec{q}$, to another probability vector $\vec{p}$:

$$p_i = \sum_j B_{ij}\, q_j \ , \qquad q_j, p_i \geq 0 \ , \quad \textstyle\sum_j q_j = \sum_i p_i = 1 \ . \tag{4-2}$$

If also condition (*iii*) holds, the matrix is unital, which means that the uniform distribution, $p_i = \frac{1}{N}$, is a fixed point of the map. A stochastic map is a kind of contraction of a probability simplex. Bistochastic maps contracts the simplex towards the uniform distribution.

The bistochastic matrix that comes out from Schrödinger's theorem is obtained from a unitary matrix U, by taking the absolute value squared of the matrix elements:

$$B_{ij} = |U_{ij}|^2 \ . \tag{4-3}$$

A matrix B is said to be *unistochastic* if there exists such a unitary U. Unistochastic matrices turn up in different contexts within quantum theory; examples are mentioned in Paper III. To decide whether a given bistochastic matrix is unistochastic or not is in general a hard problem, solved completely only for $N = 2$ and $N = 3$. In Paper III [11] we report some new results regarding this question for the case $N = 4$.

Working with bistochastic matrices is often cumbersome. Some of the difficulties that arise have to do with the fact that the study of bistochastic matrices is not really a part of linear algebra. Already the definition of a bistochastic matrix presupposes a fixed basis—bistochastic matrices is a part of "matrix analysis".

4.1 Majorization

Majorization is a way to compare probability distributions. Let $\vec{x}$ and $\vec{y}$ be two probability vectors, that is, their elements are non-negative and sum to one: $x_i, y_i \geq 0$ and $\sum_i x_i = \sum_i y_i = 1$. Make sure that both vectors have the same number N of components, by adding zero elements if necessary. We will not care about the given order of the probabilities, but arrange the elements in decreasing order, denoted $\vec{x}^{\downarrow}$. The *majorization* relation is defined as follows (see for example [1, 2, 16]): $\vec{x}$ is majorized by $\vec{y}$—written $\vec{x} \prec \vec{y}$—if

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \quad , \qquad k = 1, \ldots, N \ . \tag{4-4}$$

The largest probability in $\vec{y}$ should be larger than the largest of $\vec{x}$. The sum of the two largest probabilities in $\vec{y}$ should be larger than the sum of the two largest of $\vec{x}$. The sum of the three largest probabilities ... and so on. When the sum goes up to $k = N$, there should be an equality; the sum of all the elements of $\vec{y}$ should equal the sum of the elements of $\vec{x}$.

Since we consider normalized probability vectors, this last requirement is satisfied automatically.

Majorization provides a partial preordering on the set of probability distributions: partial because any two vectors are not in general related by majorization, and 'pre' because $\vec{x} \prec \vec{y}$ and $\vec{y} \prec \vec{x}$ does not imply $\vec{x} = \vec{y}$ (it implies $\vec{x}^{\downarrow} = \vec{y}^{\downarrow}$, that is, the vector $\vec{y}$ is obtained by permuting the components of $\vec{x}$).

For any vector $\vec{x}$ we have

$$\left(\tfrac{1}{N}, \ldots, \tfrac{1}{N}\right) \prec (x_1, \ldots, x_N) \prec (1, 0, \ldots, 0) \tag{4-5}$$

—or in words: the uniform distribution $(\frac{1}{N}, \ldots, \frac{1}{N})$ is majorized by everything and the pure distribution $(1, 0, \ldots, 0)$ majorizes everything. When $\vec{x} \prec \vec{y}$ the probabilities in $\vec{x}$ are "more even" than in $\vec{y}$. There is "more certainty" in $\vec{y}$ than in $\vec{x}$.

Instead of the defining inequalities (4-4), majorization can be characterized in terms of bistochastic matrices. This relation between majorization and bistochastic matrices is given in the following theorem [46].

Theorem (Hardy, Littlewood, Polya):

$$\vec{x} \prec \vec{y} \quad \Leftrightarrow \quad \vec{x} = B\,\vec{y}\,, \quad \textit{for a bistochastic matrix } B\,. \tag{4-6}$$

Thus, bistochastic matrices take probability vectors to majorized—"more even"—vectors.

The simplest bistochastic matrices, except permutations, act trivially only on two components. These are the so called T-transforms, for example

$$T = \begin{pmatrix} t & 1-t & 0 \\ 1-t & t & 0 \\ 0 & 0 & 1 \end{pmatrix}\,, \quad 0 \le t \le 1. \tag{4-7}$$

The effect of this matrix is to make the first two components of a probability vector more equal. Any bistochastic matrix B can be obtained as a product of a sequence of T-transforms, which in each step makes a pair of probabilities more equal. Figure 4-1 indicates how T-transforms act in the simplex of 3-dimensional probability vectors.

As already mentioned, majorization is only a partial order on the set of probability vectors. Thus, we might ask: What probability distributions can be compared? First note that the set of vectors $\vec{x}$ that are majorized by a given vector $\vec{p}$ is convex. This follows from the fact that bistochastic

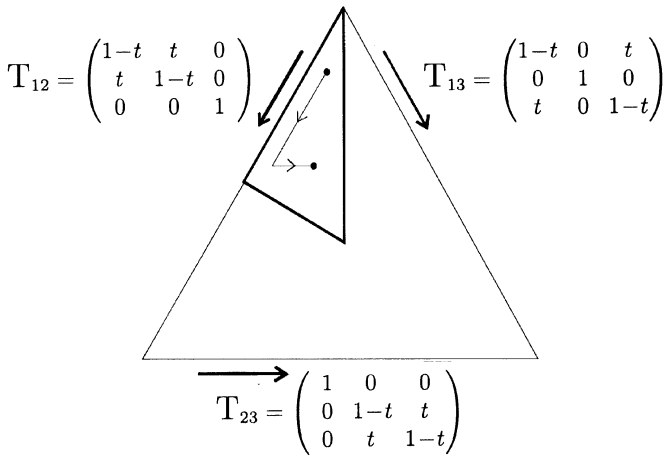


Figure 4-1: *The upper left area of the simplex contains probability vectors with the components ordered decreasingly. The three T-transforms move these probabilities in the directions indicated, and if the parameter t is no larger than $1/2$ the components are still decreasingly ordered. Every T-transform moves a probability closer (or not at all) to the uniform distribution $1/N(1,\ldots,1)$.*

matrices form a convex set.

$$\vec{x}_1 \prec \vec{p} \text{ and } \vec{x}_2 \prec \vec{p} \quad \Rightarrow \quad \vec{x}_1 = B_1\,\vec{p} \text{ and } \vec{x}_2 = B_2\,\vec{p} \quad \Rightarrow$$
$$p\,\vec{x}_1 + (1-p)\vec{x}_2 = (p\,B_1 + (1-p)B_2)\,\vec{p} \quad \Rightarrow \quad p\,\vec{x}_1 + (1-p)\vec{x}_2 \prec \vec{p}\,. \tag{4-8}$$

Furthermore, any majorized vector $\vec{x}$ can be written as a convex combination of the vector $\vec{p}$ and all vectors with the components of $\vec{p}$ permuted. Thus all majorized vectors form a polytope with $N!$ vertices, except in those cases where some components of $\vec{p}$ are equal whence the permutations yields fewer vertices. Figure 4-2 shows an example of what it looks like. The set of vectors that majorizes $\vec{p}$ is also depicted. What is left over is a set of vectors which are incomparable to $\vec{p}$. (From the figure it is also easy to convince oneself that any bistochastic matrix can be obtained as a product of T-transforms, since a sequence of T-transforms can move $\vec{p}$ to any majorized probability distribution.)

Some readers might have come to think of entropy when we mentioned that the probabilities in $\vec{x}$ are "more even" than in $\vec{y}$ if $\vec{x} \prec \vec{y}$. And yes, there is a relation between majorization and entropy. The entropy function is an example of a more general concept, that we now describe.

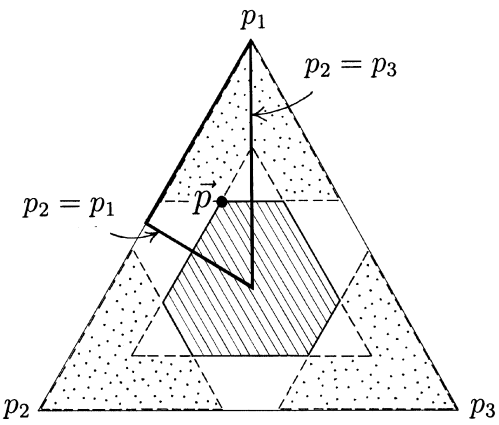


Figure 4-2: *The striped convex set of probability distributions are majorized by $\vec{p}$, and the dotted set majorizes $\vec{p}$. The blank areas are those probability distributions that are incomparable with $\vec{p}$ in terms of majorization.*

A real valued function ϕ on $\mathbb{R}^N$ is called *Schur-convex* if

$$\vec{x} \prec \vec{y} \quad \Rightarrow \quad \phi(\vec{x}) \leq \phi(\vec{y}) \ . \tag{4-9}$$

The term "convex" is used because in the expression $\vec{x} = B\,\vec{y}$, B can be understood as implementing a kind of averaging of $\vec{y}$, which yields $\vec{x}$. A Schur-convex function decreases when it's argument is averaged in this sense. Moreover, for any real valued function $f(x)$ on $[0,1]$ which is convex in the usual sense,* the function $\phi(\vec{x}) = \sum_i f(x_i)$ is Schur-convex. And conversely, if $\phi(\vec{x}) = \sum_i f(x_i) \leq \sum_i f(y_i) = \phi(\vec{y})$ for every convex function $f(x)$, then $\vec{x} \prec \vec{y}$ [16, 62]. There are also Schur-convex functions not stemming from an ordinary convex function. A simple example is the function $\phi(\vec{x}) = -\Pi_i\, x_i$. If $-f$ is Schur-convex f is said to be Schur-concave.

Since $f(x) = x^2$ is convex, the function

$$\phi(\vec{x}) = \sum_i x_i^2 \tag{4-10}$$

is an example of a Schur-convex function. In section 3.1 about the Hilbert-Schmidt geometry we found that the distance between the center ρ_0 and

* f is convex if $f(tx_a + (1-t)x_b) \leq tf(x_a) + (1-t)f(x_b)$, for all $t \in [0,1]$ (and any x_a, x_b).

any other density matrix ρ, depends only on the eigenvalues λ_i of ρ (equation (3-2)). In terms of the function ϕ this distance is $(\phi(\vec{\lambda}) - 1/N)/2$. It follows that density matrices with eigenvalues that are majorized by the eigenvalues $\vec{\lambda}$ of ρ lie closer to the midpoint, and those with eigenvalues majorizing $\vec{\lambda}$ lie further away.

An example of a Schur-concave function is the Shannon entropy,

$$H(\vec{x}) = -\sum_i x_i \log x_i \ , \tag{4-11}$$

which is used as a measure of information. The entropy can be said to measure the uncertainty of a random variable with probability distribution $\vec{x}$, or, differently phrased, it measures how much information that is gained on average when one learns the value of the random variable. Since the entropy is Schur-concave, we have $H(\vec{x}) \geq H(\vec{y})$ whenever $\vec{x} \prec \vec{y}$—a majorized probability distribution, being more "even", has a larger entropy and corresponds to higher uncertainty. Note however, that the converse does not hold. The Shannon entropy orders all probability distributions whereas majorization is only a partial order. For example, which one of the two vectors $\vec{x} = \frac{1}{9}(4,4,1)$ and $\vec{y} = \frac{1}{9}(5,2,2)$ is "most even" and which one is "most mixed"? The Shannon entropies are unequal—$H(\vec{x}) = 1,39$ and $H(\vec{y}) = 1,44$ (where the logarithm with base 2 has been used)—but these vectors are not comparable in terms of majorization.

According to the second law of thermodynamics the entropy of a closed system increases during any process. But even if a process is allowed by this criterion, it might not be possible for it to happen spontaneously. Consider, for example, three gas chambers of the same size with doors in between them, possible to open so that any pair of the

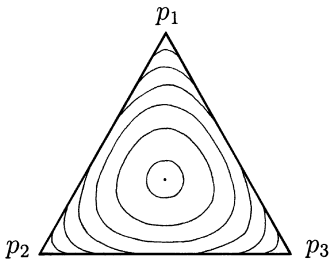
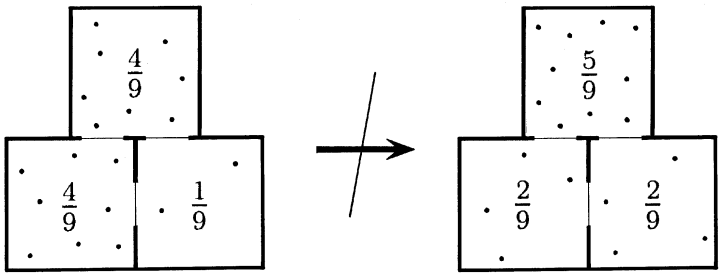


Figure 4-3: *Curves of constant Shannon entropy.*

chambers can be connected for some time. Assume the chambers contain fractions of the gas corresponding to the components of the vector $\vec{x}$ above. Can this state evolve into a state with fractions given by $\vec{y}$? Such a process would increase the entropy and are thus allowed by the second law. But it is not possible to increase the fraction 4/9 of the gas to 5/9 in the first chamber. Every possible process is stochastic, since the amount of gas does not change. Moreover the processes have to be bistochastic, since the homogeneous distribution (with one third of the gas in each chamber) will never change. Hence we see that the probability vector describing the final state should be majorized by the probability vector describing the initial state. In our example $\vec{x} \not\succ \vec{y}$ and there is no process going from the distribution $\vec{x}$ to $\vec{y}$. This issue is discussed by Mead in ref. [62], where this majorization criterion, which is stronger than requiring increasing entropy, is shown to be valid for a wide range of processes. For quantum systems, the same is true for the eigenvalues of the initial and final density matrices. There are also similar statements about systems coupled to a heat bath.



$$\tfrac{1}{9}(4,4,1) \not\succ \tfrac{1}{9}(5,2,2)$$

Figure 4-4: *It is impossible to open the doors between the gas chambers so that the distribution* $\vec{x} = \frac{1}{9}(4,4,1)$ *evolves to* $\vec{y} = \frac{1}{9}(5,2,2)$. *Such a process would increase entropy, but it would violate the majorization criterion.*

Now we will, once again, return to the issue about mixing density matrices. This is the main reason why majorization is studied in this thesis. Compare equation (2-17),

$$p_i = \sum_{j=1}^{M} B_{ij}\,\lambda_j \ , \quad \text{where } B_{ij} = |U_{ij}|^2 \ ,$$

for the probabilities p_i of some ensemble, with the characterization theo-

rem, equation (4-6),

$$\vec{x} \prec \vec{y} \quad \Leftrightarrow \quad \vec{x} = B\,\vec{y}\ , \quad \text{for a bistochastic matrix } B\ .$$

We see that if an ensemble $\{(|\psi_i\rangle, p_i)\}$ gives a density matrix with eigenvalues λ_i, then the vector $\vec{p}$ is majorized by the vector $\vec{\lambda}$, $\vec{p} \prec \vec{\lambda}$. Hence, the mixing of non-orthogonal states is always with a "more even" probability distribution than that for orthogonal states (the eigenvectors). In particular no probability p_i can be larger than the largest eigenvalue.

However the converse—that there exists an ensemble $\{(|\psi_i\rangle, p_i)\}$, giving ρ, for every probability vector $\vec{p}$ majorized by the eigenvalues $\vec{\lambda}$—is not a consequence. It is not true that such pure states $|\psi_i\rangle$, corresponding to the probabilities p_i, always are possible to find. Why this is a tricky problem is described in Paper II [9]. However, for most cases this seems to be possible. For the special case where all states should come with the same probability it is not difficult to find an ensemble. One example is to use the unitary $U_{kl} = \frac{1}{\sqrt{M}} e^{\frac{2\pi i}{M} kl}$ in equation (2-16), yielding the states

$$|\psi_k\rangle = \sum_{j=1}^{N} e^{\frac{2\pi i}{M} kl} \sqrt{\lambda_l}\, |e_l\rangle\ , \quad i = 1, \ldots, M\ , \quad M \geq N\ . \tag{4-12}$$

Examples of ensembles giving a certain qubit state are shown in figure 3-6.

4.2 Birkhoff's polytope

The set of bistochastic matrices is convex. This is easily seen by adding two matrices B_1 and B_2 with the weights p and $1-p$. As the matrix elements are non-negative and sum to one over each row and each column the convex combination is again a bistochastic matrix:

$$p\,B_1 + (1-p)B_2 = B\ , \quad 0 \leq p \leq 1\ . \tag{4-13}$$

The extreme points are those matrices with only one non-zero matrix element in each row and each column. These are the permutation matrices. Thus, the set of bistochastic matrices is the convex hull of the $N!$ permutation matrices. It is called *Birkhoff's polytope* [17]. The dimension of the set is $(N-1)^2$; the last row and last column of every matrix are constrained by the requirements (ii) and (iii) of (4-1), thus there are $(N-1)^2$ parameters to specify. With Hilbert-Scmidt distances,

$$D^2(B_1, B_2) \equiv \mathrm{Tr}(B_1 - B_2)^2\ , \tag{4-14}$$

the set becomes a regular polytope in Euclidean space. From Carathéodory's theorem (see section 2.3) it follows that it is enough to take a convex combination of not more than $N^2 - 2N + 2$ of the permutation matrices, to get any given bistochastic matrix.

The point in the middle of the polytope is called the *van der Waerden matrix*, $B_\star$ (denoted J_N in Paper III) It is the matrix with all entries equal, that is,

$$B_{\star ij} \equiv \frac{1}{N} \ . \tag{4-15}$$

In Paper III [11], Birkhoff's polytope is described and studied for $N = 3$ and $N = 4$. Birkhoff's polytope for $N = 3$ has a simple description. The three permutation matrices with positive determinant forms a regular triangle, as do the ones with negative determinant. These two triangles sit in two totally orthogonal planes. All convex combinations of the triangles forms the four dimensional polytope. Already for $N = 4$ it becomes much more complicated. It is not easy to view nine dimensional objects, but there are things that can be said to give a better understanding and a hunch about the set, built on three-dimensional intuition. In Paper III some new interesting results are reported.

One problem is to characterize the subset of the polytope that contains unistochastic matrices (equation (4-3)). When $N = 2$ every bistochastic matrix is unistochastic, but in all other dimensions this is not true anymore. The extreme points are still unistochastic, whereas there are non-unistochastic matrices in the polytope, so the subset of unistochastic matrices is not convex. To further characterize this subset, becomes quite involved—due to the diffuculty to decide whether a given matrix is unistochastic or not—already for $N = 4$, where we can tell only a part of the full story. The van der Waerden matrix $B_\star$ is unistochastic in all dimensions. When $N = 3$ every matrix close to $B_\star$ is unistochastic—there exists a ball around $B_\star$ containing only unistochastic matrices. We have found that this is not true when $N = 4$. Infinitesimally close to $B_\star$, there exist bistochastic matrices which are not unistochastic. For higher dimensions it follows from results on so called defects for Hadamard matrices that there is a ball of unistochastic around $B_\star$ whenever N is a prime, and also when $N = 6$ [77]. The result on defects in the six dimensional case appear also in Paper VI [8].

Another problem is to establish to what extent a unitary U is determined by the bistochastic matrix B, where $B_{ij} = |U_{ij}^2|$. For a general bistochastic matrix B the ambiguity is discrete. For special cases there exists continous sets of unitaries U corresponding to the same bistochastic

matrix B.

The unitary matrices giving rise to the van der Waerden matrix $B_\star$, is a research subject of its own, studied already in the nineteenth century [76] but still attracting a lot of attention. Those unitaries are the Hadamard matrices, sometimes called complex Hadamard matrices, to be more specific. In computer science real Hadamard matrices are used, for example in error correcting codes. In quantum information theory also complex Hadamards are interesting. This leads us to the next topic of this thesis, because they can, for instance, be used to represent mutually unbiased bases (as we will see in section 5.3).

Chapter 5

Mutually Unbiased Bases

One of the first things I learnt in quantum mechanics, which is very difficult to digest, is this: If you know exactly where a particle is, you can know nothing about its velocity, and if you know the velocity exactly it is equally probable to find the particle at any place in the universe. This is because the observables for position and momentum do not commute; $[q, p] = i\hbar\mathbb{1}$. In a similar way, for a spin $\frac{1}{2}$ particle: If the spin is known along the z-axis, nothing is known about the spin along the x-axis, or along the y-axis. These three spin operators do not commute; $[s_i, s_j] = i\epsilon_{ijk}s_k$.* Moreover, these observables are *maximally non-commutative*, since every eigenstate in one of the three bases has equal overlap with the eigenstates of the two other bases. The three bases are said to be mutually unbiased—or mutually conjugate, or complementary. Such bases are the concern of Papers IV and VI [10, 8].

More generally two orthonormal bases $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ in the Hilbert space $\mathcal{H}^N$ are said to be *unbiased* if

$$|\langle a_i|b_j\rangle|^2 = \frac{1}{N} \ , \quad \text{for all } i, j \ . \tag{5-1}$$

A whole set of bases is *mutually unbiased* if every pair of bases in the set is unbiased. Also a set of observables whose eigenbases are mutually unbiased, is said to be mutually unbiased. Such sets of several bases were first considered by Wiesner —under the name "conjugate"— in his seminal paper *Conjugate coding*. The paper *"treats a class of codes made possible by restrictions on measurement related to the uncertainty principal"*, and it was the starting point of quantum cryptography. Though

* In the finite dimensional case the commutator can never be proportional to the identity, since $\mathrm{Tr}[A, B] = 0 \neq \mathrm{Tr}\,\mathbb{1}$, whenever the trace is well defined ($N < \infty$).

he wrote it in the late sixties (according to ref. [**?**]), it was not published until 1983 [88]. Wiesner explains: *"Physically, if a system is in a state described by* $a_i,\ i = 1, ..., N$*, then it must have an equal probability of being found in any of the states* $b_i,\ i = 1, ..., N$ *and vice versa, if it is in a state* b_i *it must have an equal probability to be found in any* a_i*."*

Mutually unbiased bases—MUBs for short—have attracted a lot of interest the last years. They are interesting for studies in foundations of quantum mechanics (discussed, for example, in [22]) and they are used in quantum cryptography.† The famous BB84-protocol for secure quantum cryptography makes use of two unbiased bases for qubits [13]. The possibility to use several mutually unbiased bases has also been considered, and it is found that the security is increased [23, 7, 27]. However, it is noted that it comes with the cost of a lower rate for obtaining the key [27]. MUBs are also central to the entertaining "The Mean King's Problem" [83, 34, 74].

However, the main reason for the avalanche of papers about MUBs is probably related to the possibility to use MUBs for tomographic state determination, because in this context there is an obvious question about MUBs, not answered, despite a lot of effort. This is the question about the existence of "complete sets of MUBs", explained below.

Another reason for studying MUBs is their connection to discrete phase space and the possibility to describe quantum states with Wigner functions. I will say more about the quantum phase space later in this chapter. In some respects, the Wigner function is like a classical probability distribution. This is interesting for understanding the foundations of quantum mechanics, and its classical limit.

5.1 Complete sets of MUBs ?

Consider a source of quantum systems in a state ρ. The problem of state determination is to find what measurements should be performed to collect data from which the state can be estimated. We can start by measuring one observable—on a large enough (finite) number of systems—to obtain a probability distribution over N outcomes, where N is the dimension of the Hilbert space. This gives $N-1$ independent numbers. To determine the state ρ we need N^2-1 parameters. So we choose a new observable to measure, which gives us $N-1$ new numbers. It is reasonable

† It is the key distribution, that is, building up a secret cryptographic key common to sender and receiver, which can be made more secure using quantum mechanical properties.

to expect that it should be advantageous if the two chosen observables are unbiased, because then the outcomes of the two measurements are uncorrelated and we would gain as much information as possible. That this is so, has been shown by Wootters and Fields [93]. Their result is that, if the measurements to be done should be chosen ahead of the experiment and not adapted by and by, and if all measurements should be von Neumann measurements, then the statistical error will be minimized if the measurements are mutually unbiased. The number of observables/bases needed to determine the state is $N + 1$: $(N + 1)$ bases times $(N - 1)$ independent probabilities gives the $N^2 - 1$ parameters needed.

In two dimensional Hilbert space we know that there are three MUBs: the eigenbases of the spin operators along the x, y and z axes. This is just what we need to do an optimal state determination. But what about higher dimensional Hilbert spaces? Can we find MUBs there too? And how many?

Yes, mutually unbiased bases can be found in any dimensions. But it is not at all trivial to answer the question of how large a set of MUBs can be. An upper bound is known: the number of MUBs cannot exceed $N+1$ (see Paper IV)—the number needed for optimal state determination. We call a set of $N + 1$ MUBs a *complete set of MUBs*.

For prime dimensions, $N = p$, complete sets of MUBs were given by Ivanović in the early eighties [53]. He was considering the state determination problem and realized that the measurements needed could be found from *"orthogonal decompositions of the set of Hermitian matrices into commutative subsets"*. In Paper IV [10] this fact is described in simpler geometrical terms: MUBs lie in orthogonal planes in the set of density matrices, and if the set is complete these planes span the set of density matrices. In the late eighties, a generalization valid for prime power dimensions, $N = p^n$, was given by Wootters and Fields [93], when they also showed the optimality of using MUBs for state determination.

In the last several years these same MUBs (or unitarily equivalent ones) have appeared in several papers, written in new ways, with somewhat different approaches, see for example [5, 57, 31]. I think what formulation to use is mainly a matter of taste, although these re-phrasings can help to give some more insight into the problem. (That the bases in ref. [5] and in ref. [57] are equivalent with Wootters' and Fields' bases (ref. [93]) was shown by Godsil and Roy [43]. They pointed out that these bases are equivalent to a construction of "orthogonal frames" by Calderbank et. al. [26]. In ref. [31] Durt shows himself that his MUBs in odd dimensions, and in dimensions 2 and 4, are equivalent to earlier construc-

tions, whereas this is not clear for higher even dimensions. And according to ref. [26] there do exist inequivalent sets of MUBs in dimensions that are odd and composite powers of two.)

These constructions of MUBs rely on the existence of finite number fields $\mathbb{F}_N$, of order $N = p^n$. For odd prime powers a complete set of MUBs is given by the standard basis $|v_k\rangle_l = \delta_{kl}$, together with the N bases [93]

$$|v_k^{(r)}\rangle_l = \frac{1}{\sqrt{N}}\, e^{\frac{2\pi i}{p}\mathrm{tr}(rl^2+kl)} \quad , \qquad r,k,l \in \mathbb{F}_N \ . \tag{5-2}$$

The label r stands for the basis, k for the vector in that basis, and l for the component of the vector. The field trace used in the exponent is defined as

$$\mathrm{tr}\, x \equiv x + x^p + x^{p^2} + \cdots + x^{p^{n-1}} \quad , \quad x \in \mathbb{F}_{p^n} \ . \tag{5-3}$$

This is always an element in the prime field $\mathbb{F}_p$, and it is just a number, modulo p. If N is a prime, 'tr' can be ignored, because then $\mathrm{tr}\, x = x$. We will not use these explicit expressions, nor will we go deeper into finite fields.

For complete sets of MUBs in even dimensions $N = 2^n$, the same formula does not work.[‡] There is something special with even dimensions,[§] which is somehow related to the fact that the unit root $e^{2\pi i/p} = e^{\pi i} = -1$, is real. The formula (5-2) will only yield real vectors, and that is not enough. For a solution of this I refer to Wootters and Fields [93] or references [58, 57, 31].

The construction by Bandyopadhyay et. al. [5] yields the same MUBs, but it reveals more of the structure of these bases. They use generalized Pauli-matrices X and Z, defined by their action on the standard basis $\{|k\rangle\}_k^N$:

$$\begin{aligned} X\,|k\rangle &= |k+1\rangle \\ Z\,|k\rangle &= w^k\,|k\rangle \quad , \quad w = e^{\frac{2\pi i}{N}k} \end{aligned} \tag{5-4}$$

For prime dimensions the eigenvectors of the following matrices form a complete set of $N+1$ mutually unbiased bases:

$$X\,,\ XZ\,,\ XZ^2\,,\ \ldots\,,\ XZ^{N-1} \ \text{ and } \ Z\ . \tag{5-5}$$

[‡] Note, however, that Durt managed to write a single expression for MUBs in the odd and even cases [31].

[§] Or equivalently, there is something special with odd dimensions, if you prefer that view!

A basic feature of these operators is that any one of them permutes the eigenvectors of any other. For prime power dimensions MUBs can be constructed in a similar way, using tensor products of the operators above.

In non-prime power dimensional Hilbert spaces, no complete set of MUBs has been found. Number fields can not be used here, since all finite fields have prime power order. But a lower bound can be found, based on the ideas used for construction of MUBs in prime power dimensions. Factorize the dimension in prime powers: $N = p_1^{n_1} \cdots p_r^{n_r}$ (all p_i different). One can then find at least $p_m^{n_m} + 1$ MUBs, where $p_m^{n_m}$ is the minimum of all the prime powers in the factorization of N (see for example [57]).

Since the eighties, not much progress has been achieved in the search for larger sets of MUBs. Not until a couple of years ago, when a new approach was found, by Wocjan and Beth [90]. They use so called 'nets' from design theory, together with Hadamard matrices, to find mutually unbiased bases in square dimensions. Those nets can be found from mutually orthogonal Latin squares. This is interesting discrete mathematics, but what it is and how it works will not be described here. What is worth noting, is that for certain dimensions their construction gives more mutually unbiased bases, than the construction using finite fields. The lowest dimension where it is known to give more MUBs, is for $N = 26^2$, where it gives 6 MUBs instead of 5 (the number here depends on how many mutually orthogonal Latin squares that have been found.) It seems unlikely that these sets of MUBs will be used for applications, but the result is important since it shows that there can be alternative ways to find MUBs. (Another example of new sets of MUBs are given in Paper VI, but those sets are not larger than what have been found earlier.)

Let us summarize what is known for low dimensional cases. Using the construction with finite fields we get complete sets of MUBs in the dimensions $N = 2, 3, 4, 5, 7, 8$ and 9; these N are all primes or powers of primes. The lowest dimension for which the problem is not solved is $N = 6$. Despite a lot of work it is not known if there exists a complete set of MUBs, already for such a low dimension. A complete set requires 7 MUBs, while the finite field method gives only $2 + 1 = 3$. In Paper VI we search for MUBs in 6 dimensions, and find some new sets of three MUBs; more about this in section 5.3. Since then, numerical work have been performed which indicate that there can be no more than three MUBs [25]; see again section 5.3. The lack of definite answers, also for low dimensions, gives some indication on how difficult the problem is. The role of prime factorization shows the importance of number theory, when trying to understand finite dimensional quantum systems. The

problem about MUBs is one example where discrete mathematics enters in quantum information theory.

5.2 Discrete phase space and finite affine planes

From the density matrix ρ, we can calculate the probabilities for different outcomes of any measurement, especially for measuring the observables of a complete set of mutually unbiased bases—if they exist and we know them. Conversely, we can determine what density matrix ρ that represents the quantum state we have, using the probability distributions obtained by measuring $N+1$ mutually unbiased observables. Consequently, the quantum state can equally well be described by the density matrix ρ or by the set of probabilities for the MUB states. One way to "store" these probabilities is in a Wigner function, defined on some sort of phase space. In this way, a quantum state can equally well be described by a density matrix or by a Wigner function.

The function Wigner [89] introduced on phase space for describing quantum states is a quasi-probability distribution, in several regards similar to a probability distribution although it may attain negative values.[¶] Wootters has developed a way to define a similar Wigner function on a discrete phase space for finite dimensional systems, provided a complete set of MUBs exist [92]. More recently he has developed the idea together with Gibbons and Hoffman [41]; see also [94] for a nice presentation. The phase space they use is a finite affine plane. We will describe what affine planes are, since they also turn up in Paper IV [10].

To get an affine plane, consider a set of *points* $\{a_\alpha\}$, together with a set of *lines* $\{l_\omega\}$. The lines are subsets of the set of points. If two such lines have no points in common they are said to be parallel. These two sets constitute an *affine plane* if the following axioms are satisfied [15]:

(A1) If a_α and a_β are distinct points, there is a unique line l_ω such that $a_\alpha, a_\beta \in l_\omega$. – *Two points define a line.*

(A2) If $a_\alpha \notin l_\omega$, there is a unique line l_σ such that $a_\alpha \in l_\sigma$ and $l_\sigma \cap l_\omega = \emptyset$. – *Through any point not in a given line, there is a unique parallel line.*

(A3) There are at least two points on each line, and there are at least two lines. – *This excludes trivial cases.*

[¶] On the other hand, there are restrictions on the Wigner function not needed for probability distributions. For example, the Wigner function can not be very peaked.

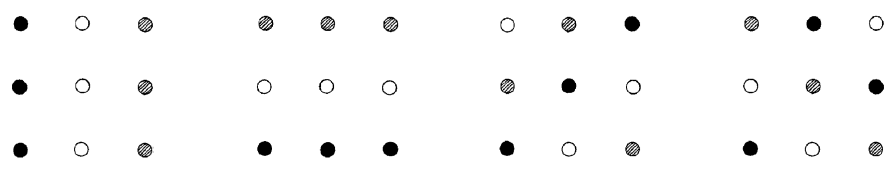


Figure 5-1: *Four copies of an affine plane of order 3, with the four pencils of three parallel lines marked.*

If the number of points is finite, then every line contains the same number of points, say N. This is the order of a finite affine plane. The total number of points is N^2, and the number of lines is $N^2 + N$. A pencil of parallel lines is a maximal set of non-intersecting lines. There are N lines in each pencil and $N + 1$ pencils in the affine plane.

It is known that finite affine planes exist of order $N = p^n$ (prime powers). For these orders it is possible to get an affine plane coordinatized by the elements in the field $\mathbb{F}_N$.[‖] These orders are the same as those dimensions where complete sets of MUBs have been found. But about affine planes more is known. Some orders are known not to be possible; these start with $N = 6, 14, 18, \ldots$ (because of a theorem [21]) and $N = 10$ (from computer calculations [28]). For other orders, starting with $N = 12, 15, 20, \ldots$, the question is still open.

Wootters and co-workers use an affine plane, of order N, coordinatized by a finite field, as a phase space of a finite dimensional quantum system. They associate lines in the plane with pure quantum states, and lay down some other requirements. This results in a phase space where every pencil of N parallel lines corresponds to a Hilbert space basis, and the $N + 1$ pencils give a complete set of MUBs. Their construction, taken together with the striking similarity of what is known about the existence of MUBs and of affine planes, makes it easy to guess that there might be some close connection between the two problems, and that they perhaps are equivalent. This may, however, be a rash conclusion. The way in which the structure of affine planes turns up in Paper IV might be an indication that the two problems are not that closely related after all. Certainly nothing definite can yet be said.

In Paper IV [10] we investigate how the MUBs sit in the space of density matrices. This means that we move to a space of higher dimension—to the mixed states instead of pure states. However, it does not necessarily

[‖] In some of these dimensions there exists also other, so-called non-Desarguesian, finite affine planes.

make things more complicated—the MUB states are ordered in a nice way in this vector space. There we define the *complementarity polytope* as the convex hull of all the states in a complete set of MUBs.** This polytope can be described, as a geometrical body, using only scalar products of the vectors for the extreme points; there is no need for explicit expressions of MUB states. Therefore, such a polytope can also be defined, regardless of the existence of any MUBs. The question about the existence of complete sets of MUBs can now be reformulated: there is a complete set of MUBs if and only if our polytope can be fitted inside the set of density matrices. Because, if this is possible, all the extreme points correspond to density matrices, and these density matrices will form a complete set of MUBs. We do not have an answer to when this is possible, but it is a new way of attacking the problem, that might bear fruit in the future.

What about affine planes in this picture? Surprisingly we find the following connection. It turns out that the problem of inscribing a regular simplex, in a certain way, in our polytope, is equivalent to the problem of finding $N-1$ mutually orthogonal Latin squares of size $N \times N$ (N is

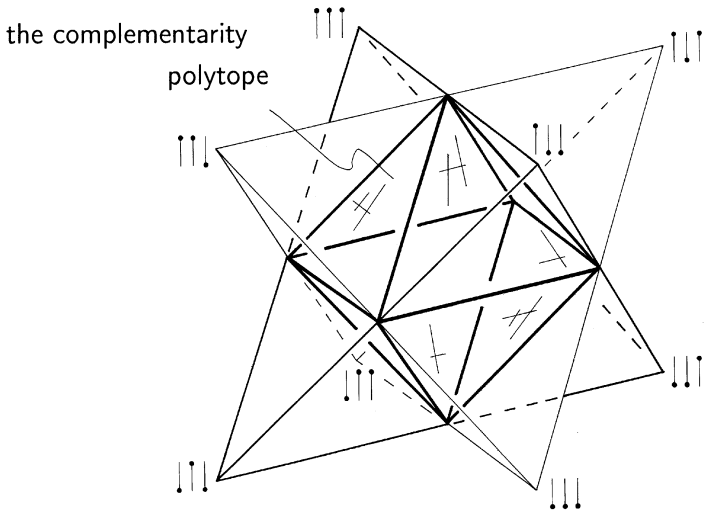


Figure 5-2: *The complementarity polytope for qubits is an octahedron, inscribed in two simplices; see Paper IV for notations, although in the paper the simplices are instead inscribed in the polytope.*

** For $N = 3$, if you take the convex combinations of two Birkhoff's polytopes sitting in orthogonal subspaces you will get the complementarity polytope. It is a nice connection, but it most probably does not mean anything.

1	2	3
3	1	2
2	3	1

1	2	3
2	3	1
3	1	2

Figure 5-3: *A pair of orthogonal Latin squares. In a Latin square the numbers 1 to N are arranged in such a way that each number occurs once in every row and in every column. Two Latin squares are orthogonal if the array of ordered pairs of numbers, obtained "superposing" the two Latin squares, contains all possible N^2 combinations.*

the dimension of Hilbert space; figure 5-3 show a pair of orthogonal Latin squares). Furthermore, it is known that the problem of finding these mutually orthogonal Latin squares is equivalent to the problem of finding an affine plane of order N. Thus, the simplex can be inscribed in our polytope if and only if there exists a finite affine plane.

This means that we have found the structure of an affine plane within our polytope, for any N where an affine plane exists. How this is related, if at all, to the possibility of arragning the polytope so that it fits in the set of density matrices, is not clear. That is, the relation between the affine plane and a complete set of MUBs remains to be understood. Hopefully, further investigations will unveil the connection, and perhaps yield a better understanding of finite dimensional quantum systems.

5.3 Hadamard matrices

Another route to mutually unbiased bases is to study Hadamard matrices. If we begin with the standard basis, $|v_k\rangle_l = \delta_{kl}$, every vector unbiased with it must have elements of the form $\frac{1}{\sqrt{N}}e^{i\phi}$, $\phi \in \mathbb{R}$. To get an unbiased basis, we need N such vectors, orthogonal to each other. This is equivalent to finding a complex Hadamard matrix, if we let the vectors be the columns of the matrix. An Hadamard matrix H is a unitary matrix with all its elements having the same modulus:

$$|H_{kl}| = \frac{1}{\sqrt{N}} \quad , \qquad k, l = 1, \ldots, N \ . \tag{5-6}$$

In some contexts Hadamard matrices are assumed to be real, but here we always refer to the more general case where the matrix elements might be complex. The MUB-states in equation (5-2) give us examples of

Hadamard matrices: for every $r \in \mathbb{F}_N$ there is an Hadamard matrix H with elements

$$H_{kl} = \frac{1}{\sqrt{N}}\, e^{\frac{2\pi i}{p}\mathrm{tr}(rl^2+kl)} \quad , \qquad k,l \in \mathbb{F}_N \ . \tag{5-7}$$

Every basis mutually unbiased with the standard basis corresponds in this way to an Hadamard matrix. Hence, finding Hadamard matrices is a way of searching for MUBs. This is the approach in Paper VI [8], where we seek for MUBs in six dimensions. All known Hadamard matrices are presented and equivalences between them are discussed. Here part of this work will be related; I refer to the paper for the full story. These Hadamard matrices are the possible candidates to be included in sets of MUBs. To get a set of bases, all of them mutually unbiased, one has to make sure that any two chosen Hadamard matrices H_1 and H_2 will correspond to a pair of unbiased bases. This requirement is that the scalar product of any column in H_1 with any column in H_2 has the modulus $1/\sqrt{N}$, that is, the product

$$H_1^\dagger H_2 = H_3 \tag{5-8}$$

has to be an Hadamard matrix, too.

Using the construction for complete sets of MUBs in prime power dimensions, at least three MUBs can be found in any dimension. To get such a triplet of MUBs in six dimensions, take three MUBs in two dimensions and three MUBs (out of the four possible) in three dimensions, and combine them pairwise. Take the tensor products of pairs of the two and three dimensional vectors to get a set of three MUBs in six dimensions. This yields the standard basis $\mathbb{1}$, the basis given by the Fourier matrix $\mathbf{F}$, with elements

$$F_{kl} = \frac{1}{\sqrt{N}}\, q^{kl} \quad , \qquad q = e^{2\pi i/N} \ , \tag{5-9}$$

and an enphased Fourier matrix, that is, the Fourier matrix multiplied from the left with a diagonal unitary matrix, $D\mathbf{F}$.

Instead of the last basis, $D\mathbf{F}$, there are several other possibilities for a third MUB. Grassl found all those, using an algebraic computer program [44].[††] Furthermore he established that none of these triplets—with the standard basis $\mathbb{1}$, the Fourier basis $\mathbf{F}$ and a third unbiased basis—can

[††] The relevant computation had actually been done earlier, in a search for biunimodular sequencies [18, 19]. The connection to MUBs is pointed out in Paper VI [8].

be extended to a larger set of MUBs. In Paper VI we have identified the Hadamard matrices representing the bases of these triplets. It is seen that the third basis in some cases are of another type of Hadamard matrices, inequivalent to the Fourier matrix (equivalence of Hadamard matrices is defined in Paper VI, section 2 [8]).

Unitarily equivalence for sets of MUBs is considered in section 3 of Paper VI , where the conditions for equivalence of pairs of unbiased bases are given. Let us here explicitly extend this to triplets of MUBs, and thereafter list the equivalences of all known triplets; this is only partly accomplished in Paper VI.

Acting with a unitary U on every basis vector in a MUB set gives an equivalent set, since this merely corresponds to a change of coordinates. Thus, if we always choose one of the bases to be the standard basis, represented by the identity $\mathbb{1}$, any triplet of MUBs might have the potential of generating two other equivalent triplets. Via multiplication with a unitary, any of the three bases can be chosen to be represented by the identity:

$$\{\mathbb{1}, H_1, H_2\} \approx \{H_1^\dagger, \mathbb{1}, H_1^\dagger H_2\} \approx \{H_2^\dagger, H_2^\dagger H_1, \mathbb{1}\} \ . \qquad (5\text{-}10)$$

('$\approx$' denotes the unitarily equivalence of MUB-sets.) However, when a basis is represented by an Hadamard matrix H we can always multiply from the right with any permutation P and any diagonal unitary D without changing the basis: $H \simeq HPD$ ('$\simeq$' denotes equivalence of Hadamard matrices). Employing this possibility it may turn out that some of the three triplets above correspond to the same set of MUBs.

What happens then, if we represent a basis with H_1PD instead of H_1, before applying the unitary transformation that brings this basis to the standard basis? Since we start out with the same bases as before, it is reasonable that unitary transformations give the same sets of MUBs as above. What we get is

$$\{\mathbb{1}, H_1PD, H_2\} \approx \{D^\dagger P^\dagger H_1^\dagger, \mathbb{1}, D^\dagger P^\dagger H_1^\dagger H_2\} \ . \qquad (5\text{-}11)$$

It doesn't look the same as in equation (5-10), however, if we act from the left with the unitary PD on all bases and then apply $D^\dagger P^\dagger$ from the right on the standard basis, this brings us back to the second set in (5-10). Admittedly the second triplet in (5-11) does not represent the same bases, but it differs only with the special unitary transformation PD, which entails that the bases at least remain in the same Hadamard equivalence classes.

Now we turn to the six dimensional MUB-triplets found by Grassl. Which of the triplets are equivalent? And are there new triplets to be found exploiting equivalences? We denote the Hadamard matrices as in Paper VI, section 2, where they are explicitly written out. Whenever a matrix is enphased it will be indicated by a tilde atop, thus, $D\mathbf{F} = \tilde{\mathbf{F}}$. In the list every triplet will be underlined the first time it appears. We start with the MUB-triplets including the Fourier basis, and exploit equation (5-10). Some equivalences have been ascertained through straightforward matrix multiplication, others figured out from the previous ones and remembering which matrices are circulant. The comments in the list are rather sketchy—the list should be understood as a supplement to Paper VI [8].

- $\underline{\{\mathbb{1}, \mathbf{F}(0,0), \tilde{\mathbf{F}}_A(0,0)\}} \approx \underline{\{\mathbf{F}(0,0), \mathbb{1}, \tilde{\mathbf{F}}_B(0,0)\}} \approx \{\mathbf{F}(0,0), \tilde{\mathbf{F}}_B(0,0), \mathbb{1}\}$

  $\tilde{\mathbf{F}}_A(0,0)$ and $\tilde{\mathbf{F}}_B(0,0)$ stands for two Fourier matrices differently enphased with 12th roots of unity. The two sets generated give the same MUB-triplet.

- $\underline{\{\mathbb{1}, \mathbf{F}(0,0), \tilde{\mathbf{F}}^{\mathrm{T}}(\frac{1}{6},0)\}} \approx \{\mathbf{F}(0,0), \mathbb{1}, \tilde{\mathbf{F}}^{\mathrm{T}}(\frac{1}{6},0)\} \approx \underline{\{\mathbf{F}(\frac{1}{6},0), \tilde{\mathbf{F}}(\frac{1}{6},0), \mathbb{1}\}}$

  $\tilde{\mathbf{F}}^{\mathrm{T}}(\frac{1}{6},0)$ stands for two possible matrices enphased with 12th roots of unity. Letting the Fourier matrix transform to the identity they transform into themselves. The last equivalence gives a new MUB-triplet. It is the first without the Fourier matrix. Moreover, the matrix $\mathbf{F}(\frac{1}{6},0)$ has not been in use before.

- $\underline{\{\mathbb{1}, \mathbf{F}(0,0), \tilde{\mathbf{C}}\}} \approx \underline{\{\mathbf{F}(0,0), \mathbb{1}, \tilde{\mathbf{F}}(\mathbf{0},\mathbf{0})\}} \approx \{\mathbf{F}(0,0), \tilde{\mathbf{C}}, \mathbb{1}\}$

  $\tilde{\mathbf{C}}$ stands for six possible circulant matrices enphased with 12th roots of unity. They generate six triplets which include Fourier matrices enphased with both 12th roots of unity and the number d (equation (10) in Paper VI).

This is all MUB-triplets including the Fourier basis. In conclusion we found that pairs of these triplets are unitarily equivalent with each other, but two of the triplets are also unitarily equivalent with a triplet including the standard basis $\mathbb{1}$ and two Hadamard matrices not equivalent (as Hadamard matrices) to the Fourier matrix, or any of the other bases found by Grassl.

We continue the list with the other MUB-triplets from Paper VI.

- $\underline{\{\mathbb{1}, \mathbf{F}(\frac{1}{6},\frac{1}{12}), \tilde{\mathbf{D}}(\frac{1}{8})\}} \approx \underline{\{\mathbf{F}^{\mathrm{T}}(\frac{1}{6},\frac{1}{12}), \mathbb{1}, \tilde{\mathbf{F}}(c_1,0)\}} \approx \underline{\{\mathbf{D}(-\frac{1}{8}), \tilde{\mathbf{F}}(c_1,0), \mathbb{1}\}}$

Computer calculations, using some rational roots of unity as phase factors, resulted in the first triplet. It generates two new triplets, with a "MUB-friendly" number c_1 (equation (66) in Paper VI).

- $\underline{\{\mathbb{1}, \mathbf{D}(0), \tilde{\mathbf{F}}(\frac{9}{24}+c_2, 0)\}} \approx \{\mathbf{D}(0), \mathbb{1}, \tilde{\mathbf{F}}(\frac{9}{24}+c_2, 0)\}$
  $\approx \underline{\{\mathbf{F}^{\mathrm{T}}(\frac{9}{24}+c_2, 0), \tilde{\mathbf{F}}^{\mathrm{T}}(\frac{9}{24}+c_2, 0), \mathbb{1}\}}$
  The first triplet here, found by computer search, includes a new "MUB-friendly" number c_2 (equation (67) in Paper VI). It generates yet a new MUB-triplet, not reported anywhere before.

These are the known sets of three MUBs in six dimensions. Our work shows that there exist MUB-triplets inequivalent to the triplets including the Fourier basis. Unfortunately we have not been able to either prove or disprove if the two new inequivalent triplets can be extended to larger sets of MUBs. Anyhow it is interesting to see that there are MUBs that do not—at least not in any apparent way—stem from a construction similar to the one used in the prime power case. Besides, it is known in four dimensions that there are pairs of unbiased bases impossible to extend to three or more MUBs, despite the fact that a complete set of five MUBs exists. Thus, it might be the case also in six dimensions that it is crucial what set to begin with if it should be possible to extend it.

All this said we have to admit there is later work indicating that no larger sets of MUBs than triplets exist in dimension six. In section 4 of Paper VI we introduce the following distance between two orthonormal bases $B_1 = \{|e_k\rangle\}$ and $B_2 = \{|f_l\rangle\}$ in Hilbert space:

$$D_c^2(B_1, B_2) = 1 - \frac{1}{N-1} \sum_{k=1}^{N} \sum_{k=1}^{N} \left(|\langle e_k | f_l \rangle|^2 - \frac{1}{N} \right)^2 . \qquad (5\text{-}12)$$

This distance is a generalization of the Fubini-Study distance: the Fubini-Study distance is a distance between one dimensional subspaces—every pure state is a ray in Hilbert space—and the distance we have here is a distance between $N-1$ dimensional subspaces—every basis defines a plane in Hilbert-Schmidt space [29]. The distance D_c can be thought of as a measure of "how much unbiased" two bases are, since it attains its maximal value, $D_c(B_1, B_2) = 1$, if and only if the two bases B_1 and B_2 are unbiased. For seven bases to be mutually unbiased all pairwise distances should be one and the function

$$f = \sum_{i,j=1}^{7} D_c^2(B_i, B_j) = \sum_{i,j=1}^{7} \left[1 - \frac{1}{5} \sum_{k,l=1}^{6} \left(|\langle e_k^{(i)} | e_l^{(j)} \rangle|^2 - \frac{1}{N} \right)^2 \right] . \qquad (5\text{-}13)$$

would attain its maximal value, $f_{\max} = 7$. Thus, finding the maximum of this function would answer the question whether a complete set of MUBs exists or not.

Of course, one can equivalently minimize the function

$$\tilde{f} = \sum_{1 \le i < j \le 7} \sum_{k,l=1}^{6} \left(|\langle e_k^{(i)} | e_l^{(j)} \rangle|^2 - \frac{1}{N} \right)^2 . \tag{5-14}$$

If a minimum of zero is attained a complete set of MUBs exists. Butterley and Hall have recently pursued such a minimization numerically [25]. The lowest value they reach is 1.584. Since their algorithm yields the minimum zero in some other dimensions, this gives quite strong evidence that no complete set exist. They also did a similar minimization for a set of four bases. The best result is 0.0512—a persuasive indication that not even four MUBs exist. However, one should be warned against taking this as a proof of non-existence of four or more MUBs in six dimensions. There can be other explanations of the numerical results. For example, it might be that in the non prime power cases there are many more local minima of the expression (5-14), thence the minimization will almost always get stuck in a local minimum and will never reach a global minimum of zero. It would be good to get more statistics. One check could be to consider only three bases in six dimensions. Since such MUB-triplets exist it would be good to see if the algorithm will find the minimum zero.

How does the minimum values of the function $\tilde{f}$ achieved numerically compare with the best sets of four respectively seven approximately mutually unbiased bases that occur in Paper VI [8]? We can obtain a set of four bases with all pairs except one unbiased. The "best unbiasedness" for the last pair corresponds to the distance $D_c^2 = 0.95$ (section VII of Paper VI), which implies $\tilde{f} = 0.25$. A set of seven "almost unbiased" bases can be obtained with bases unbiased with the Fourier basis (section VI of Paper VI; pick one basis from the square and four from one of the David's stars in figure 3). This set gives $\tilde{f} = 4.65$. Both of these values are much worse than what has been found numerically. However, it should hardly come as a surprise, since our sets contain many pairs of unbiased bases, and even several triplets of unbiased bases,[‡‡] whereas the numerics most likely correspond to sets of bases where no pair is unbiased.

[‡‡] 5 unbiased pairs out of 6 and 2 unbiased triplets out of 4, in the case of a four-basis set. 11 unbiased pairs out of 21 and 5 unbiased triplets out of 35, in the case of a seven-basis set.

Chapter 6

Entanglement

Perhaps the most fascinating in the theory of quantum mechanics, is the possibility of *entanglement*.* Correlations between subsystems in an entangled state can be stronger than in any classical state. Such 'quantum correlations' have been demonstrated in Bell-experiments.

Naturally, entanglement is often central in discussions on quantum foundations. Entanglement is also central in quantum information theory where it is seen as an important resource. Entangled states can be used for dense coding, for quantum teleportation and in protocols for quantum key distribution.

The most well known example of an entangled state is the singlet state for a pair of qubits,

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B\Big) \ . \tag{6-1}$$

The reduced density matrices of the singlet state are given by

$$\rho_A = \rho_B = \frac{1}{2}\Big(|0\rangle\langle 0| + |1\rangle\langle 1|\Big) = \frac{1}{2}\mathbb{1} \ . \tag{6-2}$$

This is the maximally mixed state and, in particular, it is not a pure state. Even though the combined system is described by a pure state $|\Psi\rangle$, there are no state vectors describing the states of the two subsystems separately. The subsystems are are said to be entangled. Moreover, these subsystems are maximally entangled, since the reduced states equals the maximally

* I can't refrain giving a comment to all Swedish speaking readers: The foremost translation of the word 'entanglement' is *'snärjelse'*, which is certainly to be preferred to 'sammanflätning', 'ihopflätning', 'ihoptrassling', 'intrassling', 'tilltrassling' or 'kvanttrassel', all of which have been in use occasionally.

mixed state. The correlations between the two spin$\frac{1}{2}$-particles in a singlet state is such that whenever the result of a measurement on one of the spins is 'up' (respectively 'down') in some direction $\vec{n}$, a measurement on the other spin along the same direction will always yield the opposite result 'down' (respectively 'up').

6.1 Magical entangled states

The peculiarity of entanglement makes itself apparent if we analyse what happens if a measurement is performed upon a subsystem of a bipartite system. This was first done by Einstein, Podolsky and Rosen, in their famous "EPR-paper" [33] and shortly thereafter by Schrödinger [72, 73]. The quotes by Schrödinger in the introduction, section 1.1, is from his first paper on entanglement. His second is the paper where he explained how density matrices can be seen as mixtures of different ensembles (see section 2.2)—this has an impact on how the effects of the measurements on a subsystem can be understood. Schrödinger shows how an experimenter can, without direct interference with a system, *"produce a non-vanishing probability of driving the system into any state he chooses"*.

Let's see what this refers to. Consider a state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, written in the Schmidt decomposition form (equation (2-11)):

$$|\Psi\rangle = \sum_{i=1}^{N} c_i \, |e_i\rangle_A \otimes |\mu_i\rangle_B \, . \tag{6-3}$$

$\{(|e_i\rangle_A, |c_i|^2)\}$ is the eigenensemble of ρ_A, the reduced density matrix for system A. Using Schrödinger's theorem about mixtures, we know that there exists some other ensemble $\{(|\psi_i\rangle_A, p_i)\}_{i=1}^N$, also giving the density matrix ρ_A. One of the states, say $|\psi_1\rangle_A$, can be chosen freely (within the span of ρ_A). From equation (2-16) it follows that there is a unitary matrix U such that

$$|e_i\rangle_A = \frac{1}{c_i} \sum_{j=1}^{N} U_{ij}^{-1} \sqrt{p_j} \, |\psi_j\rangle_A \, . \tag{6-4}$$

When inserted in equation (6-3), we get

$$|\Psi\rangle = \sum_{j=1}^{N} \sqrt{p_j} \, |\psi_j\rangle_A \otimes \left(\sum_{i=1}^{N} U_{ij}^{-1} |\mu_i\rangle_B \right) = \sum_{j=1}^{N} \sqrt{p_j} \, |\psi_j\rangle_A \otimes |\mu'_j\rangle_B \, , \tag{6-5}$$

with a new orthonormal basis $\{|\mu'_j\rangle_B\}$ in $\mathcal{H}_B$, defined by the last equality. Measuring in this new basis in subsystem B, we have probability p_1 for the

outcome $|\mu_1\rangle_B$.[†] The corresponding state for system A is $|\psi_1\rangle_A$—system A has been *"steered"* into the state $|\psi_1\rangle_A$.[‡] The experimenter can, just by handling system B, choose any state for system A, with a non-vanishing probability.

One can easily find this conclusion *"disconcerting"*, as Schrödinger did. How disturbing it is depends probably on how you interpret quantum states. Anyhow, this is a nonclassical feature and I think it is safe to say that whatever it is that is 'truly quantum' in this, will influence what interpretations that might be available. But this is completely irrelevant when it comes to the question on how these 'magical' entangled states might be used. The usefulness of entanglement is not dependent upon foundational matters. We will relate some results concerning this usefulness, but first we ought to define which states are entangled.

6.2 Entangled or separable?

A bipartite quantum state is said to be *entangled* if it is not separable. And it is *separable* if it is possible to write it as a statistical mixture of tensor products of states for the two subsystems:

$$\rho = \sum_i p_i \, \rho_i^A \otimes \rho_i^B \ . \tag{6-6}$$

It follows that the set of separable states is a convex subset of the set of all quantum states: convex combinations of separable states give new statistical mixtures of product states. Although this is a definition easy to state, it is in general a hard problem to determine whether a given density matrix ρ is entangled or not.

For pure states in the composite Hilbert space, the state is entangled if and only if the reduced density matrices are mixtures. Looking back at equations (2-11) and (2-12), we see that a pure state is separable if and only if there is only one term in the Schmidt decomposition.

For a general state, a necessary condition for separability is given by Peres [67]: if ρ is separable, then the partial transposed matrix ρ^{T_B} is a density matrix (that is, it has non-negative eigenvalues; see ref. [67] or

[†] The POVM-elements of this measurement are $E_j = \mathbb{1}_A \otimes |\mu'_j\rangle_{B\,B}\langle\mu'_j|$, $j = 1, \ldots, N$.

[‡] The careful reader might notice that we now talk about the state after the measurement, and this is not given by the POVM. The post-measurement state is proportional to $UE_1|\Psi\rangle$, where we assume that the unitary U is of the form $\mathbb{1}_A \otimes U_B$ since we do not interact with system A.

Paper I [35] for definition of ρ^{T_B}). For 2×2 and 2×3 systems this condition is also sufficient [50]. But for higher dimensional systems it is not enough to guarantee separability. States with positive partial transpose, $\rho^{\mathrm{T}_B} \geq 0$, are often referred to as *PPT-states*. Those PPT-states that are not separable are always *bound entangled*, or undistillable, that is, states from which it is impossible to obtain entanglement in the form of singlet states only by means of local operations and classical communication. The set of PPT-states is yet another convex set.

In the case of 2×2 systems, what is this 'partial transpose'? Which states are entangled and which are separable? Some understanding can be gained by a geometrical picture. In the 15 dimensional vector space of Hermitian unit trace matrices, the transformation of partial transposition is a reflection in an 11 dimensional plane. Every entangled state will be reflected to somewhere outside the set of density matrices, where at least one eigenvalue is negative. In Paper I [35] the set of separable states in certain three dimensional cross-sections through the set of 2×2 states are found. The four states of any orthonormal basis form a tetrahedron, in a three dimensional subspace. The convex sets of separable states within these tetrahedra are studied for a family of bases with varying 'amount of entanglement'. Figure 6-1 shows what it looks like for the maximally entangled Bell basis.

Partial transposition is always a reflection in the vector space of Hermitian unit trace matrices. For two N-dimensional systems the dimension of this vector space is $N^4 - 1$. Under partial transposition $\frac{1}{2}N^3(N-1)$ dimensions are reflected and a subspace of dimension $\frac{1}{2}N^3(N+1) - 1$ is fix. Even though transposition is a basis dependent transformation the set of partial transposed density matrices does not depend on the basis chosen, nor does it depend on the whether partial transposition is defined with respect to subsystem A or B [59].

There is another view on partial transposition, leading to an alluring reformulation of Peres criterion. Transposition of an Hermitian matrix is the same as taking the complex conjugate of the matrix elements. Complex conjugation is an anti-unitary operation, and the only (known) anti-unitary operation with a physical interpretation is time reversal. So transposition can be thought of as time reversal, up to unitary transformations. Hence, up to local unitary transformations, partial transposition is like time reversal of one of the subsystems. Of course, we cannot realize time-reversal, anyway it is inviting to restate Peres criterion in such terms: If a state is separable we can reverse time for one of the subsystems and still have a physical state. But if the state is entangled,

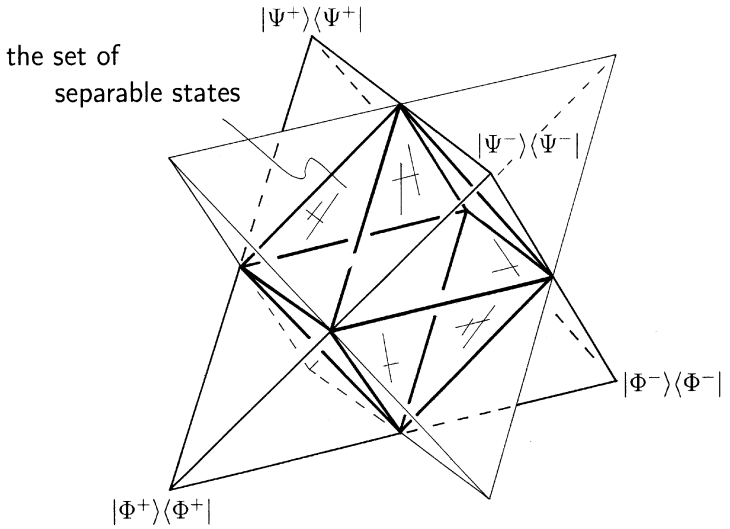


Figure 6-1: *"Stella octangula". This is a picture of the tetrahedron spanned by the Bell basis, the reflected tetrahedron from partial transposition, and the intersection of the two tetrahedra, which is an octahedron of separable states. For further explanation, see Paper I.*

reversing time for one of the subsystems will not give a physical state.[§] For entangled states it is as if the time arrows for the two subsystems are correlated.

For separable composite systems of higher dimension ($N > 6$) it is still true that there is no such correlation between the time arrows for the subsystems in an entangled state. A more general statement is the following [71]: *"This is what characterizes separable states: that any local symmetry transformation, which obviously transforms local physical states into local physical states, also transforms the global physical state into another global physical state".* For entangled states this is not always true.

Besides checking the partial transpose, there are other criteria that can be applied to investigate whether a state is entangled or not. One is stated in terms of majorization. For every undistillable state (separable or bound entangled) the vectors of eigenvalues of the state is majorized

[§] It might be tempting to write $\rho^{\mathrm{T}_B} = (\mathbb{1}_A \otimes K_B)\,\rho\,(\mathbb{1}_A \otimes K_B)$, where K stands for complex conjugation. But this does not make sense, as have been clarified by Leinaas et. al. [59]—$(\mathbb{1}_A \otimes K_B)$ is not an operator on Hilbert space.

by the eigenvalues of the reduced states [49]:

$$\vec{\lambda}(\rho_{AB}) \prec \vec{\lambda}(\rho_A) \qquad \text{and} \qquad \vec{\lambda}(\rho_{AB}) \prec \vec{\lambda}(\rho_B) \tag{6-7}$$

An extreme case is when one of the reduced states is pure, thus having only one non-zero eigenvalue, which gives a vector majorizing any other vector.

The dividing line between those states that are entangled and those that are separable is clear from the definition. But entangled states can be more or less entangled. Several measures of entanglement are in use, relevant in different contexts. (One example is the von Neumann entropy of the reduced states, which is refered to in Paper I.) In this thesis entanglement has been cosidered only for bipartite systems, but entanglement can be defined also for multipartite states. When there are several subsystems many partitions can be done and distinct types of entangled states exists. Lots of work is being done to classify entanglement in the multipartite case.

6.3 Useful entanglement

We saw in section 6.1, following Schrödinger, how it is possible to "steer" a system without interacting with it directly. The *quantum teleportation* protocol can be seen as a clever modification this idea [14]. For teleportation (in the standard version) the entangled state in use is the singlet state, equation (6-1), of two qubits. Let's assume that Alice and Bob have access to one qubit each. Instead of just measuring her qubit, Alice measures it together with another qubit in some state $|\psi\rangle$, perhaps unknown. She measures in a basis of four maximally entangled states. It can be shown that Bob's qubit will then be "steered" into the state $|\psi\rangle$ or one of three possible states $U_k|\psi\rangle$, $k = 1, 2, 3$, depending on what outcome Alice's measurement yields. Alice calls Bob to tell him the outcome and Bob can perform the unitary transformation U_k^{-1} if needed, to obtain the state $|\psi\rangle$. The state $|\psi\rangle$ has thus been teleported from Alice to Bob.[¶]

In quantum information processing, teleportation can be used as a quantum channel. Quantum information is stored in the states of quantum systems. Instead of actually move the system from one place to another the information can be teleported with the help of pre-established quantum corrrelations.

[¶] Details can be found in the seminal paper [14] or, for instance, in ref. [65].

Not only maximally entangled states can be used for teleportation. Every distillable entangled state[‖] can be used as a quantum channel, although the fidelity of the channel might be low. But using only bound entangled states will do no better than a classical channel. Still it is true that every entangled state is useful for quantum information processing. Masanes has shown that bipartite bound entangled states can enhance the teleportation power of other states [60]: Every entangled state can be employed so that it together with some other state will give a channel with higher fidelity than would be possible only with this other state. (For multipartite entangled states a similar result is true. Every entangled state can increase the "quality" of the entanglement distilled from other states [61].)

A related problem that has been studied is how one can transform entangled states of composite systems using only *local transformations and classical communication* (LOCC). Nielsen found a condition in terms of majorization, for when a bipartite state $|\Psi\rangle$ can be transformed into another state $|\Phi\rangle$, by LOCC [63]. The proof relies on Schrödingers mixture theorem. Let p_i and q_i be the eigenvalues of the reduced density matrices of $|\Psi\rangle$ and $|\Phi\rangle$. Then the transformation can be done if and only if the vector $\vec{p}$ is majorized by the vector $\vec{q}$, $\vec{p} \prec \vec{q}$. Note that the direction here is towards reduced states whose eigenvalues are "less even"; in particular it is always possible to get pure reduced states, which means that all entanglement is lost. The algorithm for how the transformations should be done utilizes the fact that there exists a bistochastic matrix B, such that $\vec{p} = B\vec{q}$ (equation (2-17)). Another algorithm uses a special way to write the bistochastic matrix as a convex combination of permutation matrices, that is, of the extreme elements of Birkhoff's polytope [54]. With this algorithm only a single measurement on one of the subsystems, followed by local unitary rotation, are needed to accomplish the transformation.

6.4 Maximally entangled bases

We just saw that majorization is useful within entanglement theory. Also other concepts dealt with in this thesis are employed there: Hadamard matrices and Latin squares are used in the most general construction known for bases of maximally entangled states.

For two qubits a basis consisting only of maximally entangled states

[‖] Distallable states are those from which it is possible to obtain entanglement in the form of singlet states only by means of local operations and classical communication.

is the so called Bell basis, containing the singlet state and three triplet states:

$$\begin{aligned} |\Phi^+\rangle &= \tfrac{1}{\sqrt{2}}(\,|00\rangle + |11\rangle\,)\;, \qquad |\Phi^-\rangle = \tfrac{1}{\sqrt{2}}(\,|00\rangle - |11\rangle\,)\;, \\ |\Psi^+\rangle &= \tfrac{1}{\sqrt{2}}(\,|01\rangle + |10\rangle)\;, \qquad |\Psi^-\rangle = \tfrac{1}{\sqrt{2}}(\,|01\rangle - |10\rangle\,)\;. \end{aligned} \tag{6-8}$$

($|xy\rangle$ stands for the state $|x\rangle_A \otimes |y\rangle_B$.) In many quantum information protocols measurements in this basis are essential, one example being the measurement to be done for quantum teleportation.

Generally a bipartite state is maximally entangled if and only if it is pure and if the reduced states are proportional to the identity. An equivalent requirement is that the number of terms in the Schmidt decomposition (equation (2-11)) equals the dimensions N of the subsystems and that all coefficients have the same modulus, $1/\sqrt{N}$. An orthonormal basis of N^2 such entangled states can be obtained from any set of N Hadamard matrices $H^{(\beta)}$, $\beta = 1, \ldots, N$ (the same Hadamard may be used several times), together with a Latin square L of size $N \times N$ [84]. The state vectors are

$$|\Psi^{(\alpha\beta)}\rangle = \sum_{k,l}^{N} H^{(\beta)}_{\alpha k}\; \delta_{l,L_{k\beta}}\; |l\rangle_A \otimes |k\rangle_B\;, \quad \alpha, \beta = 1, \ldots, N\;. \tag{6-9}$$

Every coefficient have the same modulus, since they are elements of Hadamard matrices. Of all possible products $|l\rangle_A \otimes |k\rangle_B$ exactly N appear in each state, because of the Dirac delta $\delta_{l,L_{k\beta}}$ (for every column β in the Latin square the number l occurs in exactly one row k). Orthogonality of the states follows from the properties of the Latin square and the orthogonality of the rows in the Hadamard matrices.

This construction does not only give bases of maximally entangled states. It has been established that there is a one-to-one correspondence between all of the following (with certain rather general definitions) [86]: orthonormal bases of maximally entangled vectors, quantum teleportation schemes, dense coding schemes, orthonormal bases of unitary operators and depolarizing channels. Thus, there is a connection between mathematical concepts dealt with in this thesis and several of the most central ideas in quantum information theory.

Chapter 7

Concluding remarks

In this thesis properties of quantum mechanical states have been studied, along with mathematical concepts that finds their use within quantum information theory. Geometrical descriptions have had a central place.

The set of quantum states as described with the Hilbert-Schmidt geometry has been the scene for an investigation of complementarity for states in a finite dimensional Hilbert space: Mutually unbiased bases are placed in orthogonal higher dimensional planes and form a polytope with an interesting combinatorial property. The scene is the same when the sets of separable states and entangled states have been illustrated for two qubit systems, although we are restricted to three dimensional subspaces to draw concrete illustrations.

Also the understanding of the possible ways a density matrix can be mixed from pure states is enhanced by thinking of the set of density matrices as a convex set in Hilbert-Schmidt space. The different ensembles are like different "mass distributions" giving the same "center of mass". With this picture in mind we could see that some probability distributions must be excluded, although they are majorized by the eigenvalues of the density matrix—a condition which for most cases is sufficient.

Majorization is directly linked to bistochastic matrices, and all bistochastic matrices sit in another similar vector space. They form Birkhoff's polytope, which has been given a detailed description in the three and four dimensional cases. As a subset we have all unistochastic matrices—a set more difficult to describe geometrically, but we have made some progress.

The point in the middle of Birkhoff's polytope is the van der Waerden matrix. It is always unistochastic and the corresponding unitaries—Hadamard matrices—lead us back to mutually unbiased bases. We have

studied the set of Hadamard matrices in six dimensions since they actually give us bases unbiased with the standard basis. From these we could find several inequivalent triplets of mutually unbiased bases. Whether there are larger sets of mutually unbiased bases in six dimensions is still an open question, in particular it is not known if there exists a complete set of seven mutually unbiased bases.

A somewhat related topic is that about distinguishability of quantum states—or rather, limited distinguishability owing to complementarity. I have described how the Bures-Uhlmann geometry is obtained from the concept of purifications and proven that the geodesics determine the optimal measurements, in a statistical sense—for distinguishing between states.

Several questions are left open. The one I am most eager to learn the answer to is the question about the existence of complete sets of mutually unbiased bases. Whenever an analytical proof or disproof of the existence of a complete set of mutually unbiased bases in dimension six—and other non-prime power dimensions—will be found, I think this will tell us interesting things about the set of quantum states. It might help us answer questions like: What does this set of quantum states look like? What is the difference between prime power and non prime power dimensions? What does this mean for quantum information processing? What is complementarity and how large uncertainties can there be about measurement outcomes in different bases?

Although the underlying physical theory of quantum information science is nothing but quantum mechanics, as it was formulated already in the mid nineteen-twenties, there is still much more to be explored. I believe that among the huge variety of phenomena arising from the few basic laws of quantum mechanics, new applications of quantum properties within information theory will be discovered, and this will also give new insights into the foundations of quantum mechanics.

Bibliography

*Note: References followed by "***Paper ...***" are included in this thesis.*

[1] P. M. Alberti and A. Uhlmann, *Stochasticity and partial order*, D. Reidel Publishing Company (1982).

[2] T. Ando, *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Lin. Alg. Appl. **118** (1989) 163-248.

[3] T. Ando, *Majorization and inequalities in matrix theory*, Lin. Alg. Appl. **199** (1994) 17-67.

[4] K. Ball, *An elementary introduction to modern convex geometry*, in "Flavors of Geometry", edited by S. Levy, Cambridge Univ. Press (1997) 1-58.

[5] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury and F. Vatan, *A new proof of the existence of mutually orthogonal bases*, Algorithmica **34** (2002) 512-528.

[6] H. N. Barnum, *Quantum information theory*, PhD Thesis, University of New Mexico (1998).

[7] H. Bechmann-Pasquinucci and A. Peres, *Quantum cryptography with 3-state systems*, Phys. Rev. Lett. **85** (2000) 3313.

[8] I. Bengtsson, W. Bruzda, Å. Ericsson, J.-Å. Larsson, W. Tadej and K. Życzkowski, *Mutually unbiased bases and Hadamards of Order Six*, J. Math. Phys. **48**, 052106 (2007). **Paper VI**

[9] I. Bengtsson and Å. Ericsson, *How to mix a density matrix*, Phys. Rev. A **67** (2003) 012107. **Paper II**

[10] I. Bengtsson and Å. Ericsson, *Mutually unbiased bases and the complementarity polytope*, Open Sys. & Information Dyn. (2005) **12**: 107-120. **Paper IV**

[11] I. Bengtsson, Å. Ericsson, M. Kuś, W. Tadej and K. Życzkowski, *Birkhoff's polytope and unistochastic matrices, $N = 3$ and $N = 4$*, Commun. Math. Phys. **259**, 307-324 (2005). **Paper III**

[12] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States*, Cambridge University Press, (2006).

[13] C. H. Bennett and G. Brassard, *Quantum Cryptography: Public-Key Distribution and Coin Tossing*, Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India (IEEE, New York, 1984), 175-179.

[14] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. K. Wootters, *Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels*, Phys. Rev. Lett. **70** 1895-1899 (1993).

[15] M. K. Bennett, *Affine and projective geometry*, Wiley, New York (1995).

[16] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York (1997).

[17] D. Birkhoff, *Tres observaciones sobre el algebra lineal*, Univ. Nac. Tucuman Rev. Ser. A 5:147-151 (1946).

[18] G. Björck and R. Fröberg, *A faster way to count the solutions of inhomogeneous systems of algebraic equations, with applications to cyclic n-roots*, J. Symbolic Computation **12** (1991) 329-336.

[19] G. Björck and B. Saffari, *New classes of finite unimodular sequences with unimodular Fourier transforms. Circulant Hadamard matrices with complex entries*, C. R. Acad. Sci. Paris, Sér. I **320** (1995) 319-324.

[20] N. Bohr, *Introductory survey* (1929), in "Atomic Theory and the Description of Nature", Cambridge University Press, Cambridge, (1934), pp. 1-24. Reprinted in Niels Bohr Collected Works, Vol. 6: Foundations of Quantum Physics I (1926-1932), edited by J. Kalckar, North-Holland, Amsterdam (1985), pp. 279-302.

[21] R. H. Bruck and H. J. Ryser, *The nonexistence of certain finite projective planes*, Canadian Journal of Math. **1** (1949) 88-93.

[22] C. Brukner and A. Zeilinger, *Information and fundamental elements of the structure of quantum theory*, in "Time, Quantum, Information", edited by L. Castell and O. Ischebeck, Springer (2003).

[23] D. Bruss, *Optimal eavesdropping in quantum cryptography with six states*, Phys. Rev. Lett. **81** (1998) 3018.

[24] D. J. C. Bures, *An extension to Kakutani's theorem on infinite product measures to the tensor product of semidefinite w^* algebras*, Trans. Am. Math. Soc. **135**, 199-212 (1969).

[25] P. Butterley and W. Hall, *I Numerical evidence for the maximum number of mutually unbiased bases in dimension six*, `arXiv:quant-ph/0701122`.

[26] A. R. Calderbank, P. J. Cameron, W. M. Kantor and J. J. Seidel, $\mathbf{Z}_4$-*Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets*, Proc. London Math. Soc. **75**, 436-480 (1997).

[27] N. J. Cerf, M. Bourennane, A. Karlsson and N. Gisin, *Security of quantum key distribution using d level systems*, Phys. Rev. Lett. **88** (2002) 127902.

[28] W. H. L. Clement, *The Search for a Finite Projective Plane of Order 10*, Am. Mathematical Monthly **98** (1991) 305318.

[29] J. H. Conway, R. H. Hardin and N. J. A. Sloane, *Packing lines, planes, etc: Packings in Grassmannian spaces*, Exp. Math. **5** (1996) 139159.

[30] P. A. M. Dirac, *The principles of quantum mechanics*, fourth edition, Clarendon Press, Oxford (1958).

[31] T. Durt, *If* $1 = 2 \oplus 3$*, then* $1 = 2 \odot 3$*: Bell states, finite groups, and mutually unbiased bases, a unifying approach*, `arXiv:quant-ph/0401046`.

[32] A. Einstein, in Proceedings of the Fifth Solvay Conference, p. 253, Gauthier-Villars, Paris (1928).

[33] A. Einstein, B. Podolsky and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47** (1935) 777-780.

[34] B.-G. Englert and Y. Aharonov, *The mean king's problem: Prime degrees of freedom*, Phys. Lett. A **284** (2001) 1-5.

[35] Å. Ericsson, *Separability and the stella octangula*, Phys. Lett. A **295** (2002) 256-258. **Paper I**

[36] Å. Ericsson, *Geodesics and the best measurement for distinguishing quantum states*, J. Phys. A: Math. Gen. **38** (2005) L725-L730. **Paper V**

[37] R. A. Fisher, *Theory of statistical estimation*, Proc. Cambridge. Phil. Soc. **22** (1925) 700-725.

[38] G. Fubini, *Sulle metriche definite da una forma Hermitiana*, Att Instituto Veneto **6** (1903) 501-513.

[39] C. A. Fuchs and C. M. Caves, *Mathematical Techniques for Quantum Communication Theory*, Open Sys. & Information Dyn. (1995) **3**: 345-356.

[40] W. Gerlach and O. Stern, *Der experimentelle Nachweiss der Richtungsquantelung im Magnetfeld*, Zeits. Phys. **9**, 349-355 (1922).

[41] K. S. Gibbons, M. J. Hoffman and W. K. Wootters, *Discrete phase space based on finite fields*, Phys. Rev. A **70**, 062101 (2004).

[42] A. M. Gleason, *Measures on the closed subspaces of a Hilbert space*, J. Math. Mech. **6**, 885-893 (1957).

[43] C. Godsil and A. Roy, *Equiangular lines, mutually unbiased bases, and spin models*, `arXiv:quant-ph/0511004`.

[44] M. Grassl, *On SIC-POVMs and MUBs in dimension 6*, `arXiv:quant-ph/0010082`.

[45] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür and R. Blatt, *Scalable multiparticle entanglement of trapped ions*, Nature **438**, 643-646 (2005). See also `http://heart-c704.uibk.ac.at/`.

[46] G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, 2nd edition, Cambridge University Press, New York (1952).

[47] J. E. Harriman, *Geometry of density matrices. I. Definitions. N matrices and* ***1*** *matrices*, Phys. Rev. A **17**, 1249-1268 (1978).

[48] W. Heisenberg, *The physical content of quantum kinematics and mechanics: The principle of indeterminism*, in "Quantum theory and measurement", edited by J. A. Wheeler and W. H. Zurek, Princeton University Press, (1983 (1927)), pp. 62-84.

[49] T. Hiroshima, *Majorization criterion for distillability of a bipartite quantum state*, Phys. Rev. Lett. **91**, 057902 (2003).

[50] M. Horodecki, P. Horodecki and R. Horodecki, *Separability of mixed states: necessary and sufficient conditions*, Phys. Lett. A **233** (1996) 1-8.

[51] M. Hübner, *Explicit computation of the Bures distance for density matrices*, Phys. Lett. A **163** (1992) 239-342.

[52] L. P. Hughston, R. Jozsa and W. K. Wootters, *A complete classification of quantum ensembles*, Phys. Lett. A **183** (1993) 14-18.

[53] I. D. Ivanović, *Geometrical description of quantal state determination*, J. Phys. A **14**, 324-2246 (1981).

[54] J. G. Jensen and R. Schack, *A simple algorithm for local conversion of pure states*, Phys. Rev. A **63** (2001) 062303.

[55] R. Jozsa, *Fidelity for mixed quantum states*, J. Mod. Optics **41**, 2315-2323 (1994).

[56] R. Jozsa, *Quantum information and its properties*, in "Introduction to quantum computation and information", edited by H.-K. Lo, S. Popescu and T. Spiller, World Scientific, Singapore (1998), p. 49-75.

[57] A. Klappenecker and M. Rötteler, *Constructions of mutually unbiased bases*, Proc. International Conference on Finite Fields and Applications (2003) pp. 137-144.

[58] J. Lawrence, Č. Brukner and A. Zeilinger, *Mutually unbiased binary observable sets on N qubits*, Phys. Rev. A **65** (2002) 032320.

[59] J. M. Leinaas, J. Myrheim and E. Ovrum, *Geometrical aspects of entanglement*, Phys. Rev. A **74**, 012313 (2006).

[60] Ll. Masanes, *All bipartite entangled states are useful for information processing*, Phys. Rev. Lett. **96** (2006) 150501.

[61] Ll. Masanes , *Useful entanglement can be extracted from all nonseparable states*, `arXiv:quant-ph/0510188`.

[62] C. A. Mead, *Mixing character and its application to irreversible processes in macroscopic systems*, J. Chem, Phys. **66**, 459-467 (1977).

[63] M. A. Nielsen, *Conditions for a class of entanglement transformations*, Phys. Rev. Lett. **83**, 436-439 (1999).

[64] M. A. Nielsen, *Probability distributions consistent with a mixed state*, Phys. Rev. A **62** (2000) 052308.

[65] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Cambridge University Press (2000).

[66] W. Pauli, in a letter to W. Heisenberg, 1926, in "Wolfgang Pauli, Wissenschaftlicher Briefwechsel", edited by K. v. Meyenn, Springer-Verlag, New York (1979-2000).

[67] A. Peres, *Separability criterion for density matrices*, Phys. Rev. Lett. **77**, 1413-1415 (1996).

[68] John Preskill, *Quantum Information and Computation*, Lecture notes at `www.theory.caltech.edu/people/preskill/ph229/#lecture`.

[69] C. R. Rao, *Information and accuracy attainable in the estimation of statistical parameters*, Bull. Calcutta Math. Soc. **37** (1945) 81-91.

[70] J. J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley Publishing Company, (1994).

[71] A. Sanpera, R. Tarrach and G. Vidal, *Quantum separability, time reversal and canonical decompositions*, `arXiv:quant-ph/9707041`.

[72] E. Schrödinger, *Discussion of probability relations between separated systems*, Proc. Cambridge Philos. Soc. **31**, 555-563 (1935).

[73] E. Schrödinger, *Probability relations between separated systems*, Proc. Cambridge Philos. Soc. **32**, 446-452 (1936).

[74] O. Schulz, R. Steinhübl, M. Weber, B.-G. Englert, C. Kurtsiefer and W. Weinfurter, *How to ascertain the values of σ_x, σ_y, and σ_z of a polarization qubit*, Phys. Rev. Lett. **90**, 177901 (2003).

[75] E. Study, *Kürzeste Wege in komplexen Gebiet*, Math. Annalen **60** (1905) 321-378.

[76] J. J. Sylvester, *Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers*, Phil. Mag. **34**, 461-475 (1867).

[77] W. Tadej and K. Życzkowski, *A concise guide to complex Hadamard matrices*, Open Sys. Information Dyn. **13** (2006) 133. For updated version of the catalogue of Hadamard matrices, see `http://chaos.if.uj.edu.pl/~karol/hadamard`.

[78] A. Uhlmann, *On the Shannon entropy and related functionals on convex sets*, Rep. Math. Phys. **1** (1970) 147-159.

[79] A. Uhlmann, *The "transition probability" in the state space of a *-algebra*, Rep. Math. Phys. **9** (1976) 273-279.

[80] A. Uhlmann, *Parallel transport and "quantum holonomy" along density operators*, Rep. Math. Phys. **24** (1986) 229-240.

[81] A. Uhlmann, *Density operators as an arena for differential geometry*, Rep. Math. Phys. **33** (1993) 253-263.

[82] A. Uhlmann, *Geometry of State Spaces,* Lecture notes available at http://www.mpipks-dresden.mpg.de/~issqui05 (2005).

[83] L. Vaidman, Y. Aharonov and D. Z. Albert, *How to ascertain the values of σ_x, σ_y, and σ_z of a spin-$\frac{1}{2}$ particle*, Phys. Rev. Lett. **58**, 1385-1387 (1987).

[84] K. G. H. Vollbrecht and R. F. Werner, *Why two qubits are special*, J. Math. Phys. **41**, 6772-6782 (2000).

[85] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin (1932).

[86] R. F. Werner, *All teleportation and dense coding schemes*, J. Math. Phys. A **34**, 7081-7094 (2001).

[87] H. Weyl, *The theory of groups and quantum mechanics*, Dover Publications (1950). (First published in German 1931.)

[88] S. Wiesner, *Conjugate coding*, SIGACT News **15**, 78-88 (1983).

[89] E. P. Wigner, *On the quantum correction for thermodynamic equilibrium*, Phys. Rev. **40** (1932) 749-759.

[90] P. Wocjan and T. Beth, *New construction of mutually unbiased bases in square dimensions*, Quant. Inf. & Comp. **5**, 93-101 (2005).

[91] W. K. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D **23**, 357-362 (1981).

[92] W. K. Wootters, *A Wigner-function formulation of finite-state quantum mechanics*, Ann. Phys. **176** (1987) 1-21.

[93] W. K. Wootters and B. D. Fields, *Optimal state-determination by mutually unbiased measurements*, Ann. Phys. **191** (1989) 363-381.

[94] W. K. Wootters, *Picturing Qubits in Phase Space*, `arXiv: quant-ph/0306135`.